

Can Digital Aid Deliver During Humanitarian Crises?

Michael Callen^{1†}, Miguel Fajardo-Steinhäuser^{1†}, Michael G. Findley^{2†},
Tarek Ghani^{3*†}

¹Department of Economics, London School of Economics, 32 Lincoln's Inn Fields,
London, WC2A 3PH, United Kingdom.

²Department of Government; McCoombs School of Business; LBJ School of Public
Affairs, University of Texas Austin, 3.102 Batts, Austin, 78712, TX, USA.

³Olin Business School, Washington University St. Louis, 1 Brookings Dr, St. Louis,
63130, Missouri, USA.

*Corresponding author(s). E-mail(s): tghani@wustl.edu;

Contributing authors: m.j.callen@lse.ac.uk; m.fajardo-steinhauser@lse.ac.uk;
mikefindley@utexas.edu;

†These authors contributed equally to this work.

Abstract

Global hunger levels have set new records in each of the last three years, outpacing aid budgets [1, 2]. Most households experiencing food insecurity crises are now in fragile states [3], making it difficult to support vulnerable, hard-to-reach populations without interference from oppressive governments and non-state actors [4, 5]. Despite growing interest in using digital payments for crisis response [6–8], there is limited causal evidence on their efficacy [9, 10]. We show that digital payments can address basic needs during a humanitarian crisis. We conducted a randomized evaluation among very poor, mostly tech-illiterate, female-headed households in Afghanistan with digital transfers of \$45 every two weeks for two months. Digital aid led to significant improvements in nutrition and in mental well-being. We find high usage rates, no evidence of diversion by the Taliban despite rigorous checks, and that 80% of recipients would prefer digital aid rather than pay a 2.5% fee to receive aid in cash. Conservative assumptions put the cost of delivery under 7 cents per dollar, which is 10 cents per dollar less than the World Food Program's global figure for cash-based humanitarian assistance. These savings could help reduce hunger without additional resources. Asked to predict our findings, policymakers and experts underestimated beneficiaries' ability to use digital transfers and overestimated the likelihood of diversion and the costs of delivery. Such misperceptions might impede the adoption of new technologies during crises. These results highlight the potential for digital aid to complement existing approaches in supporting vulnerable populations during crises.

Keywords: Economic development, Social protection, Fragile states, Digital Payments, Food Insecurity

1 Introduction

As global hunger outstrips available resources, humanitarian organizations are pioneering new approaches and are transitioning from in-kind food assistance to more cost-effective delivery models.¹ For example, the share of the World Food Program’s (WFP) assistance delivered as cash-based transfers rose from 2% in 2010 to 34% in 2021 [11], a move supported by research evidence [12]. While many development agencies deliver aid digitally via mobile money or e-banking [13], humanitarians have been slower to adopt these same digital transfers [14], especially in fragile states.

Hunger is clustering in states governed by oppressive regimes and violently contested by non-state actors [3],² raising concerns about aid diversion [4, 5], and highlighting both the promise and perceived challenges of implementing digital aid [7, 9]. Oppressive regimes and non-state actors restrict access to the most vulnerable populations to control resource flows. In such settings, humanitarians face a dilemma: either deliver aid and risk supporting hostile actors and exacerbating conflict [15] or suspend operations in the face of urgent needs. Indeed, the availability of support for aid programs is partly shaped by donor country concerns about diversion to hostile actors.³ As a potentially more cost-efficient alternative to physical cash distribution, digital transfers offer an attractive route to assisting vulnerable, hard-to-reach populations if they can utilize the technology. Digital aid may also prove harder to divert than traditional models of delivering assistance in the most fragile contexts, leading to growing policy interest [cf. 17, 18].

Here we develop, implement, and evaluate an approach to delivering digital transfers to a vulnerable, hard-to-reach population in the midst of Afghanistan’s humanitarian crisis. We utilize community-based institutions and an interoperable payments platform to address common issues facing digital transfers. Locally-elected Community Development Councils identified very poor, female-headed households as beneficiaries.⁴ Working with an Afghan commercial digital payments platform, we transferred digital vouchers to beneficiaries’ mobile phones that could be exchanged for goods from local merchants.⁵ Our approach builds on a rich economics literature on the adoption of financial technologies [cf. 6, 24, 25].⁶

The intervention transferred 4,000 AFN (approximately \$45 USD) every two weeks for two months to households in 16 neighborhoods in Kabul, Herat, and Mazar-i-Sharif. From our experimental sample of 2,409 households, we randomly assigned 1,208 households to an “early” group, which received benefits from November 6, 2022 until January 1, 2023. The remaining 1,201 households formed the “late” group, which received benefits from January 1, 2023 (two weeks after the “early” group stopped receiving payments) until February 28, 2023. We conducted monthly phone surveys with the beneficiaries to track outcomes. We identify causal impact by using the Late Group as the counterfactual for the Early Group during the two months in which the Early Group was receiving its payments. The study design is depicted in Figure E1.

Our experiment yields three core results. First, despite low levels of education and literacy, 99.75% of our sample successfully used their digital payments, and about 80% would not be willing to pay a 2.5% fee to have a cash out option. Second, digital payments improved all prespecified measures of nutritional well-being, with an index of these measures increasing by 0.5σ ($SE = 0.032$; $p < 0.0001$) and all prespecified measures of mental and financial health, with an index of these measures improving by 1.5σ ($SE = 0.042$; $p < 0.0001$). Last, rigorous tests reveal no evidence of diversion either from beneficiaries or from merchants accepting digital payments. This result is particularly

¹Humanitarian food security budgets fell 40% from USD 85 per person in 2018 to USD 51 per person in 2021 [2].

²In 2022, 11 countries – Afghanistan, the Democratic Republic of the Congo, Nigeria, South Sudan, Sudan, Yemen, Somalia, the Central African Republic, Syria, Chad and Haiti – accounted for over 50 percent of the quarter billion facing acute hunger and potential starvation [those with a food insecurity score of 3 or higher according to 2, among the roughly 1.350B people analysed].

³For example, in testimony to the U.S. House Committee on Oversight and Accountability, which has considerable discretion over U.S. international assistance, the Special Inspector General for Afghanistan Reconstruction stated “The Taliban generate income from U.S. aid by imposing customs charges on shipments coming into the country and charging taxes and fees directly on NGOs” [16].

⁴These institutions ensured the safety of staff and beneficiaries by informing local authorities and maintaining independence from external interference in the beneficiary selection process. The CDCs are the subject of several high-profile studies, including [19–22].

⁵Specifically, we transferred digital *value* vouchers, which had a denominated currency value redeemable at participating merchants for any available goods. WFP [23] treats value vouchers as an official form of a cash-based transfer, unlike *commodity* vouchers, which are both tied to specific merchants and goods.

⁶See SM D.1, which reviews relevant literature and demonstrates the uniqueness of the study’s focus on digital aid for vulnerable, hard-to-reach populations during a humanitarian crisis.

important given growing calls in the U.S. Congress – the primary funder of humanitarian assistance in Afghanistan – to cut aid levels given concerns that the Taliban are diverting resources, a relevant issue in other fragile countries as well.⁷ While digital aid enables scalable checks on diversion, we discuss potential risks that might accompany scaling in Section 4 below.

More broadly, digital payments yield several advantages, which are relevant in both fragile and repressive states. First, the system facilitates decentralized aid distribution as beneficiaries do not appear at a common location to receive physical cash or in-kind food, which lowers delivery costs and reduces opportunities for diversion. Second, payments offer end-to-end traceability based on blockchain records, enabling greater transparency for donors while preserving the privacy of beneficiaries. Third, even accounting for all aspects of facilitation, the cost of delivery at 6.7 cents per dollar is less than 40% of the WFP’s global figure of 17 cents per dollar for cash-based transfers, helping to address the growing humanitarian funding gap. Fourth, following one-time registration, digital payments provide purchasing power instantly, enabling more dynamic delivery approaches. In our sample, participants used their funds shortly after receiving payments.

To our knowledge, this is the first randomized evaluation of an aid delivery method designed to reach a vulnerable population despite the host government’s desire to tightly restrict interactions between beneficiaries and humanitarians.⁸ Research during humanitarian crises with populations facing extreme cultural and political constraints is not common. In SM D, we review a substantial literature and find 23 RCTs of cash for humanitarian applications, 10 such RCTs in contexts categorized as “not free” by Freedom House, three of those 10 having some digital component, and none, that we can find, targeting populations that the government actively oppresses.⁹ Our study is also unique in its careful focus on measuring the extent to which this modality for humanitarian aid avoids diversion and in documenting a cost of delivery below previous comparable studies.

2 Results

2.1 Can the extreme poor use digital vouchers?

Can those in need use digital payments? This is a fundamental question for less technology-literate samples such as ours, where 63.3% of the women in our sample have no schooling and 33.9% have at most primary education.¹⁰ The results indicate high levels of usage. Nearly all of the women in our treatment group (99.75%) used their digital payments to buy goods, and 94.14% of them continued to use the payments after the initial four transfers.¹¹ 98.2% of the total value transferred in the four payments was spent in the first eight weeks. It is important to note that the degree of success in utilization was largely the result of an emphasis on human-centered design and extensive piloting.¹²

Beyond the high levels of usage, two additional results support the argument that this tech-illiterate population can use digital vouchers. First, 20.9% of the funds were spent at different merchants than those who facilitated an initial “test transaction” during onboarding, indicating that beneficiaries understood they could use the payment at any participating merchant. Second, following our pre-analysis plan, we also checked whether impacts on nutritional outcomes vary by pre-intervention need, city, marital status, age, household size, education, and whether women were the primary household financial decision-maker. We found no heterogeneity in impacts, consistent with the technology being roughly equally useful for the different groups in our study (Figure E3). Finally, we asked participants in a hypothetical, non-incentivized exercise whether they would prefer to receive their full 4000 AFN payments digitally, or 4000 – X AFN in cash, where $X \in \{100, 300, 500\}$ reflects the costs that would be associated with receiving the aid in cash.¹³ With a conservative fee of 2.5%, 80% of participants preferred digital aid over cash; with a higher fee of 7.5%, the share choosing digital aid over cash

⁷Cliffe et al. [5] provide in-depth discussion; for media coverage, see O’Donnell [26], UN News [27], Politico [28].

⁸SM C details the ethical considerations addressed in the design and implementation of the study.

⁹We coded the Freedom House rating for the first year of the respective program if the program occurred in 2013 or later given that ratings only go to 2013. For programs occurring before 2013 (Ecuador 2011; Philippines 2010; Sri Lanka 2005; Uganda 2009; Niger 2009 and 2011; DRC 2011; Yemen 2011), we used the Freedom House 2013 value [29].

¹⁰Using the 2015 DHS, we compare our sample to a representative sample of similarly-aged women in urban areas of Kabul, Balkh and Mazar. The DHS sample has higher educational attainment: 56.8% have no schooling and 13.9% have at most primary education.

¹¹Three women in the treatment group never used the funds transferred to their accounts. Tracking efforts indicate that each migrated to new cities during the period between enrollment and the start of payments.

¹²See SM A.1 for full details of the piloting process.

¹³See SM A.2 for details on this exercise.

increased to $\sim 95\%$ (Figure E4).¹⁴ Collectively, the results indicate that digital payments are a viable option even with marginalized populations.

2.2 Do digital transfers improve nutritional and mental well-being?

The payments reduced all four pre-specified measures of hunger (Table 1, Panel A). Beneficiaries, on average, skipped 0.76 fewer meals per week (SE = 0.051, $p \leq 0.0001$), children skipped 0.117 fewer meals (SE = 0.012, $p \leq 0.0001$), the share of households where everyone was able to eat during the prior week increased by 0.093 percentage points (SE = 0.015, $p \leq 0.0001$), and beneficiaries reduced meals of only bread and tea by 1.608 (SE = 0.121, $p \leq 0.0001$). This reflects a more nutritionally diverse diet¹⁵ and an increased ability to purchase medicine when needed.¹⁶ An index of these four measures improved by 0.5 standard deviations (SE = 0.032, $p \leq 0.0001$). These measures kept improving throughout the two months of payments (Table 2), and more modest nutritional improvements remain for at least two months after payments conclude (Figure E2). We do not find any increases in outside income, employment, or agency over financial decisions (Table 1 and Table E4).

To assess if these estimated impacts are in line with what we should expect given the size of the payments, we collected data from households on what they were purchasing (e.g., wheat flour, cooking oil, and sugar) and price data from merchants. We calculate that the cost of a basket of popular goods to support a family for two weeks is consistent with the aid payment size plus the survey participation incentive. While not eliminated, skipped meals declined after each payment (Figure E2, Panel A) and follow-up surveys indicate a range of explanatory factors for the persistence of basic needs including survey timing, household size, and substitution behaviors.¹⁷

Turning to mental well-being, since the Taliban takeover in 2021, Afghanistan consistently ranks as the country with the lowest levels of happiness [see, e.g., 31, 32]. Our participants report extremely low levels of happiness and financial health.¹⁸ Treatment improved all four prespecified measures of mental and financial health, with an index of these measures improving by 1.5 standard deviations (Table 1, Panel B), although these impacts disappear as soon as payment stops (Figure E2). Specifically, beneficiaries are 33.5 percentage points more likely to report that they feel the economic situation of their household has improved compared to 30 days ago (SE = 0.011, $p \leq 0.0001$), from a base of just 4.8% in the control group. They are also 26.3 percentage points more likely to report being satisfied with their current financial situation (SE = 0.012, $p \leq 0.0001$). The intervention also improved their mental health:¹⁹ Beneficiaries are 28 percentage points more likely to report being very or quite happy (SE = 0.014, $p \leq 0.0001$), and report a score 1.96 higher on the Cantril Self-Anchoring Striving Scale, an established measure of life satisfaction that runs from 1 to 10 (SE = 0.068, $p \leq 0.0001$). Table 2, Panel B shows that these effects are increasing over time, with the effects after receiving 3-4 aid payments between 40% to 100% larger than after 1-2 aid payments.

Our results are robust to common concerns in RCTs. First, we find no evidence of contamination from borrowing and spillovers (due to this population being credit constrained and not exchanging resources with other participating households, see SM A.3). Second, survey bias in the absence of objective measures of well-being (such as anthropometric or biometric measures) is a concern, especially since participants might respond strategically if they believe this could influence their benefits. To check for potential strategic reporting by participants, we primed half of our respondents with an explicit description of the study’s purpose, finding no evidence that the prime influenced responses

¹⁴Outside the study, the digital payment platform applied a 0.5% commission for customers to withdraw cash from centralized agents, and a 3-5% commission for community-based cash-out events (similar to a WFP cash distribution, but with individual digital vouchers) depending on the remoteness of the site. By contrast, purchases at registered merchants did not involve a fee.

¹⁵Participants report eating rice 0.6 days (SE = 0.035, $p \leq 0.0001$), beans 0.49 days (SE = 0.029, $p \leq 0.0001$), chicken 0.01 days (SE = 0.016, $p = 0.035$), and dairy 0.05 days (SE = 0.013, $p \leq 0.0001$) more in the previous week. These represent increases between 60% to 95% of the mean in the Late Group.

¹⁶This question is asked only to those who said that they had had any medical needs to purchase medicine in the past 30 days. Given that attrition could be differential by treatment group (e.g. if the Early Group is doing better health-wise due to the aid payments), we also calculate Lee bounds for this variable following Lee [30]. The 95% confidence interval for the Lee bounds goes from 0.014 to 0.068.

¹⁷See SM A.5 for further discussion of interpreting the magnitudes of needs results.

¹⁸In the control group, just 15.4% of respondents answer that they are either very happy or quite happy, all things considered, 4.8% report that the overall economic situation of their household is either slightly or much better than 30 days ago, and 13.3% agree with the statement that they are highly satisfied with their present financial condition. See column 1 of Table 1.

¹⁹While these results may seem unrealistically large, this is attributable to low base levels, as is seen in other surveys in contemporary Afghanistan [see, e.g., 31, 32].

(see Table E5.)²⁰ Finally, given that participants were given the option of skipping questions, and to keep the sample comparable, we re-estimate the results restricting the sample to those who answered all questions relevant for the analysis. Results using this restricted sample are almost identical to baseline results (see Table E6).

Our nutrition results complement a broader literature demonstrating mostly positive impacts, whereas our mental well-being results contribute newer insights to a recently emerging literature with mixed results. As discussed in SM D.2, our review of 23 cash-based programs in humanitarian contexts found that 21 included nutritional outcomes, with 16 of those 21 finding any positive effect. Less attention has focused on the psychological impacts of cash-based programs, with only seven of the 23 studies considering a relevant outcome and six finding any positive effect.

2.3 Were digital transfers diverted?

An important concern about distributing aid in fragile settings is diversion, especially to regimes with poor human rights records and to those that might sponsor international terror. Reports of aid diversion across multiple fragile countries have emerged in the past year [26–28]. This is especially a concern in Afghanistan. In July of 2023, the US House of Representatives passed a bill that would bar the Secretary of State and the USAID Administrator from giving any funds to Afghanistan if it directly or indirectly supported the Taliban [33]. Reports of Taliban aid requirements and infiltration of UN assistance [16, 34] are deepening these concerns, adding further pressure to cut assistance. Therefore, the humanitarian mandate to address hunger depends partly on avoiding diversion.

Several pieces of evidence suggest that this intervention avoided capture by the Taliban. First, we directly asked beneficiaries whether they have been asked for informal assistance. Specifically, we asked them whether local community leaders or government officials have asked them for any kind of assistance, such as food or money. Given that participants may be reluctant to disclose doing this themselves, we first asked them whether they know someone in their community who has been asked to do so [similar to the approach in 35], and then whether they themselves had been asked. Results are presented in Table 3, Panel A, which show that the Early Group does not report informal payments to authorities in larger proportion than the Late Group (all coefficients are insignificant at traditional levels, and precisely estimated). Overall, only 27 beneficiaries in the Early Group answered yes to any of these four questions since payments started, compared to 21 beneficiaries in the Late Group, with the difference statistically insignificant as well (column 5).²¹ Table 3, Panel B, shows that the results are also not increasing over time, when beneficiaries could have become more visible to local authorities.

Even so, given the authoritarian context these are questions that are sensitive for participants [40] and thus could lead to a lack of positive responses due to fear of reporting inappropriate behavior by local authorities. We conducted a list experiment, an established method to measure the presence of sensitive behaviors, with beneficiaries to gauge the extent of informal taxation in this population.²² Table 3, Panel C, shows the results of the list experiment. Regardless of whether we analyze the whole sample or either of the two treatment groups, individuals who receive the longer list including the informal taxation statement do not report experiencing significantly more relevant experiences than those receiving the shorter list without the informal taxation statement.²³ The effects are precisely estimated and small in magnitude.

We also have access to the beneficiaries’ transaction data from the mobile payments provider, which coupled with the fact that they could not cash out their payments and that they spent almost all the money they received, provides a clear picture of how beneficiaries spent their funds. Beneficiaries

²⁰For more details on this experimenter demand exercise, see SM A.3. We remain agnostic as to whether participants would under or overreport across measures. A priori, this could go either way, and could differ across treatment and control groups. Thus, we present results across the whole population and divided by treatment group.

²¹Anecdotally, one of the reasons for this is that participants are too poor for local authorities to ask them for payments. Evidence from a diverse set of global contexts [36, 37] and the pre-2021 Afghanistan NSP program specifically [38, 39] also indicates that most diversion occurs at national and regional levels and primarily through bureaucratic and partisan channels because community development councils develop transparency and monitoring mechanisms to prevent diversion. Note that this dummy variable for whether the respondent answered yes to any of the four diversion questions was *not* pre-specified in our Pre-Analysis Plan. In Table E3 we show results using different indices, including the KLIK index we did pre-specify in the PAP. In SM A.4 we describe this deviation from the PAP.

²²See SM A.2 for details.

²³For the exact situations included in the list experiment, see SM A.2.

spent 74.2% of their funds at the merchants they visited during the beneficiaries’ onboarding sessions, 21.2% at other registered merchants, 3.5% at individually-registered accounts, and .1% on airtime purchases. We attempted to contact all onboarding merchants, and none reported being asked for any informal assistance by local authorities.

Despite the importance of measuring diversion in humanitarian crisis response, rigorous evidence is limited.²⁴ Policy reports indicate that cash is no more prone to diversion than other types of aid during humanitarian crises [60–62]. But the studies these reports point to do not have diversion as a central research focus. In our review of 23 studies on crisis response, only two attempted to measure diversion.²⁵ Our study tests for diversion both directly in self-reports and indirectly using a list experiment and finds no evidence in either case.

2.4 Do experts understand the costs and benefits of digital transfers?

Humanitarians face urgent challenges and limited time for experimentation. This may lead to reluctance to adopt new technologies, especially if they are untested or perceived as risky. To benchmark the beliefs of key experts, we surveyed 55 international analysts, practitioners and policymakers working on global development issues (including many engaged specifically on hunger in Afghanistan) and 36 academics, many with expertise related to digital payment systems.²⁶ The survey provided a precise description of the intervention.²⁷

Table E9 compares the actual values of the intervention’s components (column 1) to the respondent’s predicted values (column 2), and shows the p-value of a test of equality (column 3). Across all four components, the predictions are statistically significantly different from the actual values. These respondents predicted, on average, that only 43% of the women in our study would be able to use digital payments (standard deviation = 25.67pp). In practice, as described above, 99.75% of women purchased goods at least once. They also predicted, on average, that 40% of beneficiaries would be taxed by local authorities. In practice, less than 2% (standard deviation = 27.28) of the treatment group reported any diversion, with no statistically significant difference in the control group. Moreover, beliefs were highly diffuse, indicating considerable uncertainty about the technology’s efficacy. Such beliefs, even if weakly held, might plausibly explain an unwillingness to adopt innovations. Results for cost-effectiveness were more accurate, with experts on average predicting delivery costs to be 10.65 cents on the dollar, while the actual delivery costs were 6.7 cents on the dollar, indicating a more common belief that mobile aid delivery is relatively cost-efficient.²⁸

3 Benefits of Digital Aid

Digital aid is both cost-efficient and cost-effective. The total *cost per beneficiary* (CPB) of delivering the entire \$180 (across four disbursements) ranges from \$2.44 USD (without recruitment) to \$12 USD (with recruitment). This information can be represented as a *total cost transfer ratio* (TCTR), which captures the cost to deliver a single dollar (including the transferred dollar). The TCTRs for our digital aid intervention range from 1.014 (\$182.44/\$180.00) to 1.067 (\$192.00/\$180.00), which correspond

²⁴Although few studies in the humanitarian cash space rigorously test for diversion [41–43], the broader international development literature long identified diversion as one of the most pernicious forms of corruption [44, 45]. Through extensive implementation of Public Expenditure Tracking Surveys (PETS) and seminal evaluations of specific public development programs, high levels of diversion across numerous contexts are well-documented [46–49]. In response, programs began to include interventions such as government audits and community-based monitoring, which effectively reduced diversion [50, 51]. A wave of digital interventions followed, including biometric smartcards, beneficiary ID cards, e-invoicing, and phone-based monitoring of distribution agents, which also reduced diversion in a variety of contexts [52–59].

²⁵Blattman et al. [41] asked participants how much of the grant they had to give to other household and community members, and participants reported less than 1% of the grant. Lehmann and Masterson [43] asked a general question of whether households had been robbed during the period in which the program was running and none reported any occurrence. Separately, Aker [42] notes that if diversion had occurred, then there should be indirect evidence in the levels of uptake between cash and vouchers. And Aker et al. [63] mentioned that the implementer feared diversion, and therefore hired security to guard cash during transport, which comprised the largest share of their costs, but did not report measuring diversion.

²⁶The policy organizations represented include the USAID, the UN World Food Program, UNICEF, the World Bank, GiveDirectly, Brookings and BRAC, and the academic institutions included Brown, Duke, Georgetown, the London School of Economics, Princeton, Stanford, the University of Chicago, UC Berkeley, and UC San Diego.

²⁷Specifically, respondents were told: “We are piloting a direct aid program to Afghan women using a commercial platform called HesabPay. Local partners have identified 2500 highly vulnerable women in three major cities (Kabul, Herat, and Mazar), each of whom will receive four semi-monthly digital payments of 50 USD. All beneficiaries complete an in-person onboarding that includes identity verification, registration of a digital wallet linked to a unique mobile phone number, and a test transaction using the digital wallet to purchase goods from a registered merchant. Aid payments are unconditional and can be used for purchases at any HesabPay-registered local merchant.” See Table E10 for the full text of the prompt.

²⁸Results are similar when looking at predictions by academics and practitioners separately and are available on request.

to cost-transfer ratios (CTRs) from 1.4 cents to 6.7 cents, when onboarding costs are excluded or included respectively.²⁹ (See SM B for details of the cost-efficiency estimation.) These cost-efficiency estimates compare favorably to other cash-based transfer programs in humanitarian crises. Although cost analysis is rarely performed for humanitarian cash programs [12, 64], we were able to identify three organization-specific global estimates and 12 individual studies with estimates, *all* of which had higher TCTRs (when cost categories are comparable). SM B details these comparative cost-efficiency estimates and also reports cost-effectiveness analyses, which incorporate information about program effectiveness, resulting in estimates of the number of dollars required to achieve a defined effect, specifically a reduction of a day with skipped meals and the corresponding change to calorie intake.

Digital aid offers further advantages to donors and humanitarian agencies in terms of increased decentralization and transparency compared to status quo modalities of assistance. Delivery of in-kind food or physical cash incurs transport and mobilization costs associated with each distribution, and is subject to potential diversion or disruption at key chokepoints, such as airports and border crossings. Furthermore, beneficiaries bear the time and travel costs associated with reaching distribution points and may be subject to harassment or extortion during this process.

By contrast, digital payments platforms enable instantaneous, private transfers of value, which can then be exchanged for goods at local merchants – or potentially “cashed-out” for physical currency – without need to visit a central distribution location each time.³⁰ This is particularly important when considering women in food-insecure settings as existing research documents substantial benefits to channeling resources to women directly [65]. Beneficiaries in our study used their digital vouchers at many locations across the merchant acceptance network, increasing convenience and privacy while also ensuring competitive pressure on any single merchant who might otherwise seek to exercise market power.³¹ As digital payments automatically generate transaction data, donors and humanitarian agencies gain increased transparency into the delivery and utilization of cash-based assistance programs that can be used for auditing and real-time programming adjustments.³² Finally, the distribution of digital payments is instantaneous, with recipients able to access their funds immediately. Figure 1, Panel A displays the share of the (cumulative) funds participants have received that remains unspent over time, and shows that participants spent over 70% of the funds they received the day they received their first payment. This drops quickly over the next few days to below 5% of the received funds. The same pattern holds every time participants receive their funds. Panel B displays how participants spent their money, with the vast majority of funds going to the original merchants.³³

4 Methodological Safeguards and Limitations

Humanitarian operations are complex and expensive, and the stakes can involve life or death. If a new approach fails, the consequences can be severe, both financially and in terms of human suffering. Particularly in fragile contexts like Afghanistan, humanitarian aid is under constant and intense scrutiny, especially from donors concerned that it is either not effective or being diverted. As such, it is appropriate for humanitarians to demand the highest standard of evidence before adopting a new approach. In addition, humanitarian innovators may have strong incentives to oversell the efficacy of new approaches [66]. In the language of statistics, false positives (i.e., Type I errors) whereby the null hypothesis that an innovation has no impact is incorrectly rejected in favor of the alternative that it is effective are both plausible and extremely costly.

Modern experimental social science provides at least three valuable methodological safeguards against false positives. To increase the credibility of our findings, we utilized these safeguards in our study, even though they are costly financially, in terms of statistical power, and in terms of the ability

²⁹For comparison, WFP’s global figure for providing cash-based humanitarian aid is 17 cents per dollar [10].

³⁰With the digital payments platform in this study, such transfers could be initiated either locally or internationally, and recipients with smartphones could also initiate peer-to-peer digital transfers. When payments are delivered as digital vouchers like in this study, there is less need to accumulate cash at intermediate locations to facilitate beneficiary cashout, further reducing delivery costs and opportunities for diversion.

³¹Table E7 reports the total number of merchants serving beneficiaries in the transaction data, which is consistently larger than the number from onboarding sessions. Figure E5 maps the locations of these merchants for each of the three cities in which we conducted the intervention, visually demonstrating the decentralized distribution network.

³²The digital payments platform in this study utilizes the Algorand blockchain, so all payments and purchases were automatically recorded on an immutable public ledger that facilitates external auditing. In addition, beneficiary accounts can be automatically screened for compliance with sanctions lists, further reducing the potential for diversion to malign actors.

³³See SM D.3 for a discussion of related work on the advantages of cash-based transfers, which highlights how few existing studies have examined these possible benefits or delivered aid as cost efficiently.

to conduct exploratory analyses. First, RCTs provide unbiased causal evidence by establishing an independent control group. This is particularly important in humanitarian settings, where the circumstances on the ground can change quickly and often worsen. Simpler analyses, such as comparing outcomes before and after an intervention, can therefore be highly misleading [67]. Second, while strict adherence to detailed Pre-Analysis Plans (PAPs) can place limits on learning, doing so helps ensure that p-values are correctly calculated, providing additional protection against false positives [68, 69].³⁴ In our case, our goal was to write a sufficiently detailed PAP that it would enable us to provide the data and the PAP to a third party and ensure that they come back with identical conclusions, which would be viewed as conservative by modern standards in economics [69].³⁵ Third, survey bias, especially in the absence of anthropometrics or biometric markers (which are challenging to measure for groups that host governments seek to isolate), is a concern as respondents may strategically misrepresent their status to increase current or future assistance. The literature on testing and avoiding survey response bias is helpful in this regard [70]. In our case, we took the most extreme approach we could think of to test for survey bias as a potential confound by randomly priming respondents with an explicit description of the purpose of our study, and found that the prime did not affect participants' responses.

Our analysis has limitations. First, the program only ran for two months and we delivered aid to our control group immediately after. Given the exigency of the situation and that it would have been unethical to onboard but not deliver assistance to the Late Group, we opted for a short, staggered program where every participant received some assistance during the lean season. We therefore cannot estimate impacts over a longer period. Second, because we could not visit our participants, and to avoid survey fatigue over the phone, we focused on a narrow set of questions and could not take measurements that would yield deeper insights (such as anthropometrics). Thus, other relevant topics (e.g. within household dynamics, validated mental health measures, and impacts on other household members) were not covered. Third, though this trial proved relatively successful, the digital approach might face additional challenges at scale. The cities in the study are less conservative than cities in Afghanistan's South and East. We only worked with female-headed households and so cannot speak to potential complications that would arise for other households. Moreover, if scaled up, the Taliban might be more likely to identify ways to divert more substantial sums. However, the digital platform builds in protections against potential abuse with automated know your customer verification, sanctions-list screening, and fixed limits on the size of digital transfers. Finally, funds are channeled directly to end users, generating an automatic administrative blockchain record and leaving no way for the Taliban to divert funds during transmission. Risks of diversion will remain; in particular, the government has the legal authority to tax both merchants and digital payment platforms.

Our analysis shows that digital aid can be a valuable tool for inexpensively providing aid, including to vulnerable, hard-to-reach groups in fragile states, which can complement existing approaches. First, we find that digital aid assists women at a facilitation cost that is 10 cents less per dollar than the WFP's global figure for providing humanitarian cash-based transfers. This is despite extreme social restrictions on women and the severe challenges of working under an oppressive regime. In 2022, had this approach been used to deliver the \$188M that was delivered as cash for humanitarian aid in Afghanistan, agencies would have saved \$18.8M. This is enough to support an additional 48,958 families or 342,708 people for the four months of the lean season. More ambitiously, if this approach could save 10 cents on the dollar for 25% (roughly the share of Afghans who live in cities) of the roughly \$1.5B USD of food security humanitarian aid delivered, 98,574 families or about 690,000 people could be supported through the lean season.³⁶ In 2022, 3.2B was delivered in total – a level of need likely to be sustained for years to come – indicating substantial potential benefits to increasing digital adoption. Second, this approach helps limit diversion, which will be important to sustain donors' willingness to support humanitarian response. Encouragingly, and informed by the results of this study, at the time of writing the WFP is piloting digital payments in Afghanistan. The existence of key enablers for the success of a program like this in other fragile settings (e.g., the presence of CDCs, mobile phone availability, and digital payments networks), as shown in Table E8, suggests

³⁴In addition, selecting multiple hypothesis test corrections that explicitly minimize the likelihood of false positives, such as correcting for the Family Wise Error Rate, can also help in this regard, as we do here.

³⁵We are not advocating for this approach in general, but do highlight the usefulness in applications where false positives are costly and where there is likely to be minimal benefit from more exploratory analysis.

³⁶The Afghanistan Cash and Voucher Working Group estimates that it costs \$96 USD to feed a family of 7 for a month.

that this approach could provide a useful complement to existing delivery mechanisms not only in Afghanistan but also in other fragile states with high levels of food insecurity.

Declarations

Author contributions

Authors are listed in the author list in alphabetical order, with their contributions as follows. All authors contributed to RCT design and supervision. TG, MC and MGF conceived the project. MFS supervised data collection and conducted data analysis. MF conducted literature review and cost-effectiveness analysis. All authors edited and revised the manuscript.

Acknowledgments

We thank Ilaha Eli Omar, Hila-Nawa Alam and the staff at Uplift Afghanistan Fund, Mukhtar Sabri and the staff at CDDO, Sanzar Kakar, Nigel Pont and the staff at HesabPay, and Kathleen McGowan for collaboration in program design and implementation. Cedric Antunes and Michael Cowan provided excellent research assistance. Shahim Kabuli provided excellent field management. Noam Angrist, Michael Denly, Daniel Nielson, Rohini Pande, Jacob Shapiro, Abu Siddique, Garima Siwach, Rebecca Wolfe, Noam Yuchtman and seminar participants at ADBI/Yale, Brookings, the IFS, King's College, and WUSTL shared insightful feedback. Research funding was provided by the UK Foreign, Commonwealth & Development Office, awarded through the J-PAL Crime and Violence Initiative.

Competing interests

The authors declare no competing interests.

Ethics approval

The research protocol was approved by the London School of Economics' Institutional Review Board (study number 89546).

5 Tables & Figures

Table 1: Summary Table – Treatment Effects

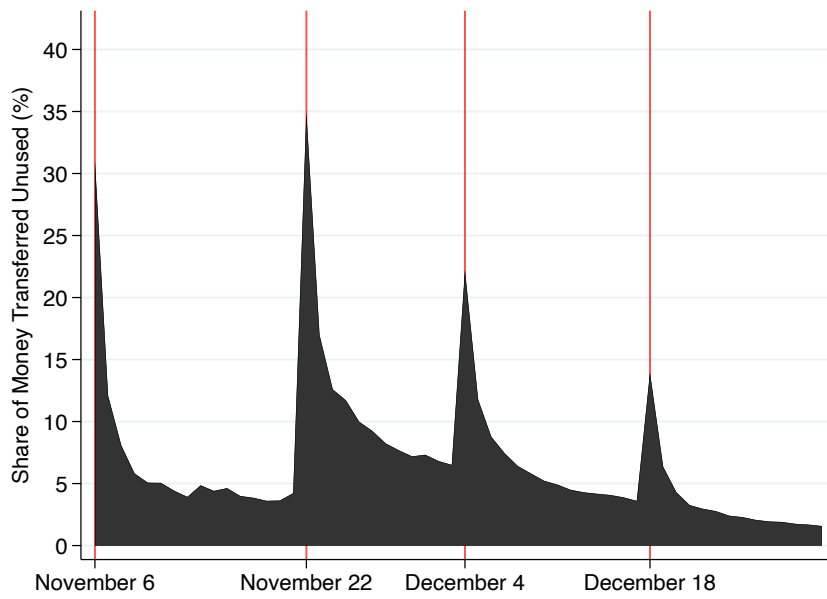
| | Control Mean (1) | Control SD (2) | Treatment Effect (3) | Standard Error (4) | Naive <i>p</i> -value (5) | Adjusted <i>p</i> -value (6) | N (7) |
|--|------------------------|----------------------|----------------------------|--------------------------|---------------------------------|------------------------------------|----------|
| Panel A. Primary Outcomes | | | | | | | |
| Days skipping meals (past week) | 2.569 | 1.717 | -0.76 | 0.051 | 0 | 0.0002 | 4761 |
| Children skipping meals (=1) | 0.873 | 0.333 | -0.117 | 0.012 | 0 | 0.0002 | 4750 |
| Regularly eat twice a day | 0.501 | 0.5 | 0.093 | 0.015 | 0 | 0.0002 | 4736 |
| Total bread and tea meals (past week) | 13.639 | 3.549 | -1.608 | 0.121 | 0 | 0.0002 | 4763 |
| <i>Food Security - KLK Index</i> | 0 | 1 | 0.501 | 0.032 | 0 | | 4763 |
| Better economic situation | 0.048 | 0.213 | 0.335 | 0.011 | 0 | 0.0002 | 4762 |
| Satisfied with fin. situation | 0.133 | 0.34 | 0.263 | 0.012 | 0 | 0.0002 | 4755 |
| Happy | 0.154 | 0.361 | 0.28 | 0.014 | 0 | 0.0002 | 4732 |
| Life satisfaction | 3.179 | 1.707 | 1.963 | 0.068 | 0 | 0.0002 | 4763 |
| <i>Economic/Wellbeing - KLK Index</i> | 0 | 1 | 1.498 | 0.042 | 0 | | 4763 |
| Panel B. Secondary Outcomes | | | | | | | |
| Days eating rice (past week) | 0.698 | 1.058 | 0.595 | 0.035 | 0 | 0.001 | 4763 |
| Days eating beans (past week) | 0.518 | 0.836 | 0.493 | 0.029 | 0 | 0.001 | 4763 |
| Days eating vegetables (past week) | 1.394 | 1.346 | -0.003 | 0.041 | 0.936 | 0.454 | 4763 |
| Days eating chicken (past week) | 0.021 | 0.147 | 0.012 | 0.006 | 0.035 | 0.037 | 4763 |
| Days eating dairy (past week) | 0.074 | 0.383 | 0.047 | 0.013 | 0 | 0.001 | 4763 |
| Able to buy medicine | 0.051 | 0.221 | 0.035 | 0.01 | 0.001 | 0.002 | 3582 |
| Involved in fin. decisions | 0.666 | 0.472 | 0.017 | 0.015 | 0.271 | 0.184 | 4757 |
| Total household income (past month) | 876.683 | 1581.531 | 129.119 | 91.029 | 0.156 | 0.117 | 4763 |
| Household's head employed (past month) | 0.192 | 0.394 | -0.008 | 0.014 | 0.59 | 0.306 | 4741 |

Notes: Control for stratification fixed effects, survey round fixed effects, and baseline value of dependent variable, if available. Standard errors clustered at individual level. Primary outcomes show FWER-adjusted *p*-values within each family outcome (following Romano & Wolf, 2005, using 5000 repetitions), while secondary outcomes show FDR-adjusted *p*-values (following Anderson, 2008). The KLK Index is created following Katz, Kling, & Liebman (2007), and is the equally-weighted sum of the standardised component variables. Better economic situation is an index that equals 1 if the respondent answered that her economic situation compared to 30 days ago is slightly or much better, and 0 otherwise. Satisfied with financial situation is a dummy that equals 1 if the respondent answered that she agrees a lot or somewhat with the statement that she is highly satisfied with her current financial condition, and 0 otherwise. Happy is a dummy that equals 1 if respondent said that she was very happy or quite happy, and 0 otherwise. Life satisfaction is the score from 1 (dissatisfied) to 10 (satisfied) in terms of how satisfied the respondent is with her life as a whole these days. Total household income excludes the aid payments.

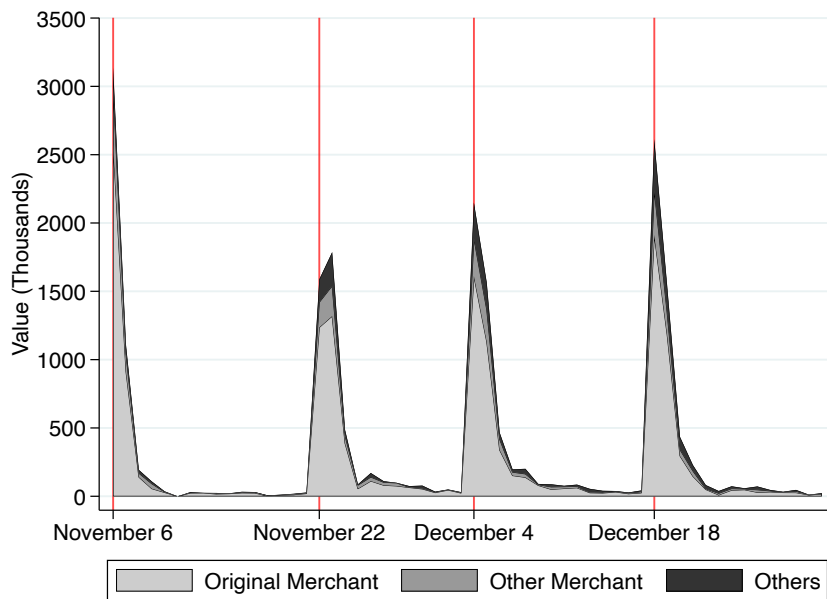
Table 2: Short vs. Long Treatment Effects – Nutrition & Wellbeing Outcomes

| | (1) | (2) | (3) | (4) | (5) |
|---|---------------------------|-------------------------------|--------------------------------|-------------------------|---------------------|
| Panel A. Food Security | Days Skipping Meals | Children Skipping Meals (=1) | Regularly Eat Twice a Day (=1) | Total Bread & Tea Meals | KLK Index |
| β_1 : Treated \times Round 2 | -0.580*** (0.094) | -0.096*** (0.021) | -0.006 (0.028) | -0.949*** (0.219) | 0.309*** (0.056) |
| β_2 : Round 2 | -0.194*** (0.071) | 0.025* (0.013) | 0.072*** (0.020) | -0.321** (0.151) | 0.096** (0.038) |
| β_3 : Treated | -0.470*** (0.072) | -0.069*** (0.015) | 0.096*** (0.020) | -1.133*** (0.163) | 0.346*** (0.042) |
| Control Mean | 2.569 | 0.873 | 0.501 | 13.639 | 0.000 |
| Observations | 4,761 | 4,750 | 4,736 | 4,763 | 4,763 |
| $(\beta_1 + \beta_2 + \beta_3) / \beta_3$ | 2.650 | 2.040 | 1.700 | 2.120 | 2.170 |
| p-value $\beta_1 + \beta_3 = 0$ | 0 | 0 | 0 | 0 | 0 |
| Panel B. Wellbeing | Better Economic Situation | Satisfied with Financial Sit. | Happy | Life Satisfaction | KLK Index |
| β_1 : Treated \times Round 2 | 0.219*** (0.021) | 0.200*** (0.024) | 0.102*** (0.026) | 0.842*** (0.127) | 0.837*** (0.078) |
| β_2 : Round 2 | -0.013 (0.009) | -0.036*** (0.014) | 0.030* (0.017) | -0.223*** (0.075) | -0.074* (0.040) |
| β_3 : Treated | 0.225*** (0.015) | 0.163*** (0.017) | 0.229*** (0.019) | 1.542*** (0.093) | 1.079*** (0.055) |
| Control Mean | 0.048 | 0.133 | 0.154 | 3.179 | 0.000 |
| Observations | 4,762 | 4,755 | 4,732 | 4,763 | 4,763 |
| $(\beta_1 + \beta_2 + \beta_3) / \beta_3$ | 1.910 | 2.010 | 1.580 | 1.400 | 1.710 |
| p-value $\beta_1 + \beta_3 = 0$ | 0 | 0 | 0 | 0 | 0 |

Notes: This table reports estimated impacts of treatment separately for the first and second survey round. Households were surveyed once per month for two months. Each of these months constitutes a survey round. All specifications control for stratum fixed effects and the baseline value of the dependent variable, if available. Standard errors are clustered at individual level. The outcome variables follow the primary outcomes shown in Table 1.



(a) Share of Money Unused



(b) Transactions Flow

Fig. 1: Account Usage Over Time

Notes: Panel A shows the cumulative share of money participants in the treated group that they have not spent. Panel B shows where participants spent their funds at. “Original merchant” is the merchant a participant visited during the onboarding session to conduct their test purchase, “other merchant” is another account that has been identified to belong to a merchants (either a different test merchant or another merchant that did not participant in the onboarding sessions).

Table 3: Are Digital Payments Diverted?

| | Gov. Off. Others (1) | Comm. Leader Others (2) | Gov. Off. You (3) | Comm. Leader You (4) | KLK Index (5) | Yes to Any Question (6) |
|---|----------------------------|-------------------------------|-------------------------|----------------------------|----------------------|-------------------------------|
| Panel A. Baseline | | | | | | |
| Treated | 0.001 (0.001) | 0.002 (0.002) | 0.002 (0.001) | 0.002 (0.003) | 0.074* (0.042) | 0.003 (0.003) |
| Observations | 4,558 | 4,596 | 4,631 | 4,626 | 4,648 | 4,509 |
| Control Mean | 0.003 | 0.002 | 0.001 | 0.005 | 0.000 | 0.009 |
| Panel B. Long-Run | | | | | | |
| β_1 : Treated \times Round 2 | -0.000 (0.003) | 0.006* (0.003) | 0.002 (0.003) | -0.001 (0.005) | 0.092 (0.084) | 0.004 (0.006) |
| β_2 : Round 2 | -0.002 (0.002) | -0.003 (0.002) | -0.003* (0.001) | -0.008** (0.003) | -0.138*** (0.045) | -0.010** (0.004) |
| β_3 : Treated | 0.001 (0.005) | -0.007 (0.005) | -0.001 (0.005) | 0.003 (0.009) | -0.064 (0.139) | -0.003 (0.011) |
| p-value $\beta_1 + \beta_3 = 0$ | 0.699 | 0.698 | 0.721 | 0.585 | 0.661 | 0.857 |
| $(\beta_1 + \beta_2 + \beta_3) / \beta_3$ | -0.850 | 0.500 | 1.810 | -1.670 | 1.700 | 2.600 |
| Observations | 4,558 | 4,596 | 4,631 | 4,626 | 4,648 | 4,509 |
| Panel C. List Experiment | | | | | | |
| | All Sample | Late Sample | Early Sample | | | |
| Treated List | -0.006 (0.030) | -0.019 (0.050) | 0.010 (0.032) | | | |
| Observations | 2,359 | 1,172 | 1,187 | | | |
| Mean Items Mentioned | 0.756 | 0.711 | 0.800 | | | |

Notes: In Panels A and B, control for stratification fixed effects, survey round fixed effects, and baseline value of dependent variable, if available. These are answers to questions of the type "Have you/someone in your community been asked to provide informal assistance (for example money or food) to local community leaders/government officials in the past month?". The outcome in column 6 was not pre-specified. Panel C shows the results of a list experiment where the treatment group received the following additional statement: "I have been approached by government officials or community leaders to provide them with any kind of assistance, like food or money, in the past month". Includes control for surveyor fixed effects. Standard errors clustered at individual level.

References

- [1] WFP: A global food crisis. World Food Programme Report (2023)
- [2] FSIN: 2023 global report on food crises (2023)
- [3] Townsend, R., Verner, D., Adubi, A., Saint-Geours, J., Leao, I., Juergenliemk, A., Robertson, T., Williams, M., Preneuf, F., Jonasova, M., et al.: Future of food: Building stronger food systems in fragility, conflict, and violence settings. World Bank Report (2021)
- [4] Kurtzer, J.D.: Denial, delay, diversion: Tackling access challenges in an evolving humanitarian landscape. Center for Strategic and International Studies (2019)
- [5] Cliffe, S., Dwan, R., Wainaina, B., Zamore, L.: Aid strategies in ‘politically estranged’ settings. Chatham House Research Paper (2023)
- [6] Suri, T., Jack, W.: The long-run poverty and gender impacts of mobile money. *Science* **354**(6317), 1288–1292 (2016)
- [7] Pazarbasioglu, C., Mora, A.G., Uttamchandani, M., Natarajan, H., Feyen, E., Saal, M.: Digital Financial Services (2020)
- [8] Aiken, E., Bellue, S., Karlan, D., Udry, C., Blumenstock, J.E.: Machine learning and phone data can improve targeting of humanitarian aid. *Nature* **603**, 864–870 (2022) <https://doi.org/10.1038/s41586-022-04484-9>
- [9] Gentilini, U.: Cash transfers in pandemic times: Evidence, practices, and implications from the largest scale up in history. World Bank Report (2022)
- [10] WFP: Cash-based transfers empower people while building resilience. World Food Programme Report (2023)
- [11] McDonough, S.: Giving people cash is usually better than shipping them food. *Vox* (2022)
- [12] World Bank: Cash and in-kind transfers in humanitarian settings (2022)
- [13] GiveDirectly: Cash evidence explorer. *Vox* (2023)
- [14] WFP: Wfp evidence summary, cash-based transfers, lessons from evaluations (2021)
- [15] Nunn, N., Qian, N.: Us food aid and civil conflict. *American economic review* **104**(6), 1630–1666 (2014)
- [16] CBS News: U.s. taxpayers helping fund afghanistan’s taliban? aid workers say they’re forced “to serve the taliban first”. CBS News (2023)
- [17] Byrd, W.: Afghanistan’s crisis requires a coherent, coordinated international response. United States Institute of Peace (2023)
- [18] De Waal, A.: The crisis in sudan calls for a new model of humanitarian aid. *Financial Times* (2023)
- [19] Beath, A., Christia, F., Enikolopov, R.: Empowering women through development aid: Evidence from a field experiment in afghanistan. *American Political Science Review* **107**(3), 540–557 (2013)
- [20] Beath, A., Christia, F., Egorov, G., Enikolopov, R.: Electoral rules and political selection: Theory and evidence from a field experiment in afghanistan. *The Review of Economic Studies* **83**(3), 932–968 (2016)
- [21] Beath, A., Christia, F., Enikolopov, R.: Can development programs counter insurgencies?:

Evidence from a field experiment in afghanistan (2017)

- [22] Beath, A., Christia, F., Enikolopov, R.: Direct democracy and resource allocation: Experimental evidence from afghanistan. *Journal of Development Economics* **124**, 199–213 (2017)
- [23] WFP: World Food Programme Cash Policy
- [24] Aker, J.C., Mbiti, I.M.: Mobile phones and economic development in africa. *Journal of economic Perspectives* **24**(3), 207–232 (2010)
- [25] Suri, T., Aker, J., Batista, C., Callen, M., Ghani, T., Jack, W., Klapper, L., Riley, E., Schaner, S., Sukhtankar, S.: Mobile money. *VoxDevLit* **2**(2), 3 (2023)
- [26] O’Donnell, L.: The taliban have “infiltrated” u.n. deliveries of aid. *Foreign Policy* (2023)
- [27] UN News: Un suspends food assistance as ethiopia wrestles with aid diversions. *UN News* (2023)
- [28] Politico: Ukraine uncovers corruption scheme implicating top officials. *Politico* (2023)
- [29] Freedom House: Freedom in the World 2023: Marking 50 Years in the Struggle for Democracy. *Freedom House Report* (2023)
- [30] Lee, D.S.: Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies* **76**(3), 1071–1102 (2009)
- [31] Gallup: Afghans lose hope under the taliban (2022)
- [32] Gallup: Taliban plunge afghans into economic ruin (2022)
- [33] Binesh, B.: Sopko speaks of us legislation restricting aid to afghanistan. *Tolo News* (2023)
- [34] Foreign Policy: The taliban have ‘infiltrated’ u.n. deliveries of aid. *Foreign Policy* (2023)
- [35] Reinikka, R., Svensson, J.: Using micro-surveys to measure and explain corruption. *World Development* **34**(2), 359–370 (2006)
- [36] Guggenheim, S., Petrie, C.: *Alternative Aid Modalities: Community Development* (2022)
- [37] Samii, C.: Wider working paper 2023/26-revisiting community-driven reconstruction in fragile states (2023) <https://doi.org/10.35188/UNU-WIDER/2023/334-5>
- [38] Beath, A., Christia, F., Enikolopov, R.: Do elected councils improve governance? experimental evidence on local institutions in afghanistan (2018)
- [39] Burde, D., Middleton, J., Rahnama, R., Samii, C.: Can Communities Take Charge? A Randomized Controlled Trial on Sustaining Schools in Afghanistan (2023)
- [40] Blair, G., Coppock, A., Moor, M.: When to worry about sensitivity bias: A social reference theory and evidence from 30 years of list experiments. *American Political Science Review* **114**, 1297–1315 (2020) <https://doi.org/10.1017/S0003055420000374>
- [41] Blattman, C., Green, E.P., Jamison, J., Lehmann, M.C., Annan, J.: The returns to microenterprise support among the ultrapoor: A field experiment in postwar uganda. *American Economic Journal: Applied Economics* **8**, 35–64 (2016) <https://doi.org/10.1257/app.20150023>
- [42] Aker, J.C.: Comparing cash and voucher transfers in a humanitarian context: Evidence from the democratic republic of congo. *World Bank Economic Review* **31**, 44–70 (2017) <https://doi.org/10.1093/wber/lhv055>

- [43] Lehmann, M.C., Masterson, D.T.R.: Does aid reduce anti-refugee violence? evidence from syrian refugees in lebanon. *American Political Science Review* **114**, 1335–1342 (2020) <https://doi.org/10.1017/S0003055420000349>
- [44] Svensson, J.: Eight questions about corruption. *Journal of Economic Perspectives* **19**, 19–42 (2005)
- [45] Olken, B.A., Pande, R.: Corruption in developing countries. *Annual Review of Economics* **4**, 479–509 (2012) <https://doi.org/10.1146/annurev-economics-080511-110917>
- [46] Reinikka, R., Svensson, J.: Local capture: Evidence from a central government transfer program in uganda. *Quarterly Journal of Economics*, 679–705 (2004)
- [47] Dutta, P., Murgai, R., Ravallion, M., Walle, D.V.D.: Does India’s Employment Guarantee Scheme Guarantee Employment? (2012)
- [48] Niehaus, P., Sukhtankar, S.: The marginal rate of corruption in public programs: Evidence from india. *Journal of Public Economics* **104**, 52–64 (2013) <https://doi.org/10.1016/j.jpubeco.2013.05.001>
- [49] Suri, T., Bhogale, S.: Digital Identification & Finance Initiative Africa: An Overview of Research Opportunities (2019)
- [50] Olken, B.A.: Monitoring corruption: Evidence from a field experiment in indonesia. *Journal of Political Economy* **115**, 200–249 (2007)
- [51] Bjorkman, M., Svensson, J.: Power to the people: Evidence from a randomized field experiment on community-based monitoring in uganda. *Quarterly Journal of Economics*, 735–769 (2009)
- [52] Cash at your fingertips: Biometric technology for transfers in developing countries. *Review of Policy Research* **29**, 91–117 (2012) <https://doi.org/10.1111/j.1541-1338.2011.00539.x>
- [53] Muralidharan, K., Niehaus, P., Sukhtankar, S.: Building state capacity: Evidence from biometric smartcards in indi. *American Economic Review* **106**, 2895–2929 (2016) <https://doi.org/10.1257/aer.20141346>
- [54] Banerjee, A., Hanna, R., Kyle, J., Olken, B.A., Percepatan, T.N., Kemiskinan, P., Khadijah, N., Kodir, C.-D., Marliani, L., Nugroho, P., Salame, H.S., Siregar, F.: Tangible Information and Citizen Empowerment: Identification Cards and Food Subsidy Programs in Indonesia Sudarno Sumarto (2018)
- [55] WDR: The Changing Nature of Work (2019)
- [56] Banerjee, A., Dufo, E., Imbert, C., Mathew, S., Pande, R.: E-governance, accountability, and leakage in public programs **12**, 39–72 (2020) <https://doi.org/10.2307/27087073>
- [57] Lowe, C., Mccord, A., Beazley, R.: National cash transfer responses to Covid-19 Operational lessons learned for social protection system-strengthening and future shocks (2021). www.odi.org/en/publications/national-cash-transfer-
- [58] Muralidharan, K., Niehaus, P., Sukhtankar, S., Weaver, J.: Improving last-mile service delivery using phone-based monitoring†. *American Economic Journal: Applied Economics* **13**, 52–82 (2021) <https://doi.org/10.1257/app.20190783>
- [59] Information bypass: Using low-cost technological innovations to curb leakages in welfare programs. *Journal of Development Economics* **164** (2023) <https://doi.org/10.1016/j.jdevec.2023.103137>
- [60] Sossouvi, K.: E-transfers in Emergencies: Implementation Support Guidelines

- [61] Gordon, L.: Risk and Humanitarian Cash Transfer Programming: Background Note for the High Level Panel on Humanitarian Cash Transfers
- [62] ODI: Doing Cash Differently: How Cash Transfers Can Transform Humanitarian Aid. www.odi.org/twitter
- [63] Aker, J.C., Boumniel, R., McClelland, A., Tierney, N.: Payment mechanisms and antipoverty programs: Evidence from a mobile money cash transfer experiment in niger. *Economic Development and Cultural Change* (2016)
- [64] Gentilini, U.: Revisiting the “cash versus food” debate: new evidence for an old puzzle? *The World Bank Research Observer* **31**(1), 135–167 (2016)
- [65] Field, E., Pande, R., Rigol, N., Schaner, S., Moore, C.T.: On her own account: How strengthening women’s financial control impacts labor supply and gender norms. *American Economic Review* **111**(7), 2342–75 (2021)
- [66] Ghani, T., Gordon, G.: Predictable disasters. *Breakthrough: The Promise of Frontier Technologies for Sustainable Development*, 103 (2022)
- [67] Wolfe, R.: The benefits and challenges of randomized control trials in conflict environments: Reflections from a scholar-practitioner. In: Acar, Y.G., Moss, S.M., Uluğ, Ö.M. (eds.) *Researching Peace, Conflict, and Power in the Field: Methodological Challenges and Opportunities*, pp. 259–280 (2020). Chap. 14
- [68] Olken, B.: Promises and perils of pre-analysis plans. *Journal of Economic Perspectives* **29**(3), 61–80 (2015)
- [69] Duflo, E., Banerjee, A., Finkelstein, A., Katz, L., Olken, B., Sautmann, A.: In praise of moderation: Suggestions for the scope and use of pre-analysis plans for rcts in economics. *National Bureau of Economic Research Working Paper* (2020)
- [70] De Quidt, J., Haushofer, J., Roth, C.: Measuring and bounding experimenter demand. *American Economic Review* **108**(11), 3266–3302 (2018)
- [71] United Nations: ‘the world cannot abandon the people’: Top humanitarian official in afghanistan. *United Nations News* (2023)
- [72] WFP: Cash-based transfers and commodity vouchers. *World Food Programme* (2023)
- [73] United Nations Capital Development Fund: Interoperability of financial services providers afghanistan. *UNCDF Report* (2023)
- [74] World Bank: Afghanistan’s economy faces critical challenges as it adjusts to new realities. *World Bank Report* (2023)
- [75] WFP: Wfp afghanistan: Situation report 13 october 2022. *World Food Programme Afghanistan Report* (2023)
- [76] United Nations: Security council emphasizes that punitive restrictions on women’s rights, escalating hunger, insecurity taking devastating toll in afghanistan. *UN News* (2023)
- [77] Roberts, L.: Taliban decree deepens afghanistan’s health crisis. *Science (New York, NY)* **379**(6629), 222–223 (2023)
- [78] United Nations: Afghanistan: Security council condemns taliban’s ban on women working for un. *UN News* (2023)
- [79] UNAMA: Cash shipments to the un in afghanistan - info sheet. *United Nations Assistance Mission in Afghanistan* (2023)

- [80] O'Donnell, L.: The taliban are abusing western aid. *Foreign Policy* (2022)
- [81] SIGAR: 2023 high-risk list. Special Inspector General for Afghanistan Reconstruction Report (2023)
- [82] Senate Armed Services Committee: Senate armed services committee hearing on posture of uscentcom and usafricom in review of the defense authorization request for fy24 and the future years defense program. *Foreign Policy* (2023)
- [83] World Bank: Mobile cellular subscriptions - afghanistan. World Bank Database (2023)
- [84] Asia Foundation: A survey of the afghan people. Asia Foundation Report (2019)
- [85] Broadcasting Board of Governors: Media use in afghanistan. Broadcasting Board of Governors Gallup Report (2015)
- [86] Gallup: Digital freedom out of reach for most afghan women (2023)
- [87] Blumenstock, J.E., Callen, M., Ghani, T., Koepke, L.: Promises and pitfalls of mobile money in afghanistan: evidence from a randomized control trial. In: Proceedings of the Seventh International Conference on Information and Communication Technologies and Development, pp. 1–10 (2015)
- [88] Blumenstock, J., Callen, M., Ghani, T.: Why do defaults affect behavior? experimental evidence from afghanistan. *American Economic Review* **108**(10), 2868–2901 (2018)
- [89] Blumenstock, J., Callen, M., Ghani, T., Gonzalez, R.: Violence and financial decisions: Evidence from mobile money in afghanistan. *Review of Economics and Statistics*, 1–45 (Forthcoming)
- [90] Blumenstock, J.E., Callen, M., Faikina, A., Fiorin, S., Ghani, T.: Strengthening fragile states: Evidence from mobile salary payments in afghanistan. Available at SSRN 4473387 (2023)
- [91] Carril, A.: Dealing with misfits in random treatment assignment. *The Stata Journal* **17**(3), 652–667 (2017)
- [92] Gallup: Suffering reaches unprecedented levels in afghanistan (2022)
- [93] Kling, J.R., Liebman, J.B., Katz, L.F.: Experimental analysis of neighborhood effects. *Econometrica* **75**(1), 83–119 (2007)
- [94] Boudreau, L.E., Chassang, S., Gonzalez-Torres, A., Heath, R.M.: Monitoring harassment in organizations. National Bureau of Economic Research Working Paper (2023)
- [95] Anderson, M.L.: Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association* **103**(484), 1481–1495 (2008)
- [96] Bete, G.: Afghanistan Cash & Voucher Working Group (CVWG) Minimum Expenditure Basket (MEB) and Setting the Transfer Value: Guidance Document
- [97] Pople, A., Hill, R., Dercon, S., Brunckhorst, B.: Anticipatory cash transfers in climate disaster response * csae working paper wps/2021-07 (2021) <https://doi.org/10.1257/rct.6576-1.1>
- [98] Bedoya, G., Coville, A., Haushofer, J., Isaqzadeh, M., Shapiro, J.P., School, W.W.: No household left behind: Afghanistan targeting the ultra poor impact evaluation. National Bureau of Economic Research (2019)
- [99] Alik-Lagrange, A., Buehren, N., Goldstein, M., Hoogeveen, J.: The Welfare Impacts of Public Works in Fragile Conflict and Violent Economies: The Londö Public Works in the Central African Republic *. <http://hdr.undp.org/en/composite/HDI>

- [100] Chowdhury, R., Collins, E., Ligon, E., Sulaiman, M.: Valuing Assets Provided to Low-Income Households in South Sudan
- [101] Quattrochi, J., Bisimwa, G., Windt, P., Voors, M.: Cash-like vouchers improve psychological well-being of vulnerable and displaced persons fleeing armed conflict. *PNAS Nexus* **1** (2022) <https://doi.org/10.1093/pnasnexus/pgac101>
- [102] MercyCorps: Beyond Meeting Immediate Needs: The Impact of Electronic Cash Transfer Approaches on Disaster Recovery and Financial Inclusion
- [103] White, P., Hodges, A., Greenslade, M.: Guidance on measuring and maximising value for money in social transfer programmes—second edition. Technical report, UK Department for International Development (2013)
- [104] Riegg Cellini, S., Edwin Kee, J.: Cost-effectiveness and cost-benefit analysis. *Handbook of practical program evaluation*, 636–672 (2015)
- [105] Clarke, P.K., Darcy, J.: Insufficient Evidence: The Quality and Use of Evidence in Humanitarian Action. www.alnap.org/pool/files/discussion-starter-evidence-alnap-2014.pdf
- [106] Burton, J.: “doing no harm” in the digital age: What the digitalization of cash means for humanitarian action. *International Review of the Red Cross* **102**, 43–73 (2020) <https://doi.org/10.1017/S1816383120000491>
- [107] WFP: Cash-based Transfers: Lessons from Evaluations
- [108] UN: UN Principles for Responsible Digital Payments
- [109] HHS: The Belmont Report
- [110] Wood, E.J.: The ethical challenges of field research in conflict zones. *Qualitative Sociology* **29**, 373–386 (2006) <https://doi.org/10.1007/s11133-006-9027-8>
- [111] Campbell, S.: Ethics of research in conflict environments. *Journal of Global Security Studies* **2** (2019)
- [112] Puri, J., Aladysheva, A., Iversen, V., Ghorpade, Y., Brück, T.: Can rigorous impact evaluations improve humanitarian assistance? *Journal of Development Effectiveness* **9**, 519–542 (2017) <https://doi.org/10.1080/19439342.2017.1388267>
- [113] Hidrobo, M., Peterman, A., Heise, L.: The effect of cash, vouchers, and food transfers on intimate partner violence: Evidence from a randomized experiment in northern ecuador. *American Economic Journal: Applied Economics* **8**, 284–303 (2016) <https://doi.org/10.1257/app.20150048>
- [114] Buller, A.M., Peterman, A., Ranganathan, M., Bleile, A., Hidrobo, M., Heise, L.: A mixed-method review of cash transfers and intimate partner violence in low and middle-income countries. *World Bank Research Observer* **33**, 218–258 (2018) <https://doi.org/10.1093/wbro/lky002>
- [115] Bastagli, F., Hagen-Zanker, J., Harman, L., Barca, V., Sturge, G., Schmidt, T.: The Impact of Cash Transfers: A Review of the Evidence from Low- and Middle-income Countries. <https://doi.org/10.1017/S0047279418000715>
- [116] Baranov, V., Cameron, L., Suarez, D.C., Thibout, C.: Theoretical Underpinnings and Meta-analysis of the Effects of Cash Transfers on Intimate Partner Violence in Low- and Middle-Income Countries. <https://doi.org/10.1080/00220388.2020.1762859>
- [117] Blofield, M., Knaul, F.M., Calderón-Anyosa, R., Peterman, A., Franzoni, J.M., O’Donnell, M., Bustreo, F.: A Diagonal and Social Protection Plus Approach to Meet the Challenges of the

COVID-19 Syndemic: Cash Transfers and Intimate Partner Violence Interventions in Latin America. [https://doi.org/10.1016/S2214-109X\(21\)00444-7](https://doi.org/10.1016/S2214-109X(21)00444-7)

- [118] Hidrobo, M., Fernald, L.: Cash transfers and domestic violence. *Journal of Health Economics* **32**, 304–319 (2013) <https://doi.org/10.1016/j.jhealeco.2012.11.002>
- [119] Angelucci, M.: Love on the rocks: Domestic violence and alcohol abuse in rural Mexico. *BE Journal of Economic Analysis & Policy* **8** (2008)
- [120] Bobonis, G.J.: The impact of conditional cash transfers on marriage and divorce. *Economic Development and Cultural Change*, 281–312 (2011)
- [121] Bobonis, G.J., González-Brenes, M., Castro, R.: Public transfers and domestic violence: The roles of private information and spousal control. *American Economic Journal: Economic Policy* **5**, 179–205 (2013) <https://doi.org/10.1257/pol.5.1.179>
- [122] McDermott, R., Hatemi, P.K.: Ethics in field experimentation: A call to establish new standards to protect the public from unwanted manipulation and real harms. *Proceedings of the National Academy of Sciences* **117**, 30014–30021 (2020) <https://doi.org/10.1073/pnas.2012021117>
- [123] APSA: Principles and Guidance for Human Subjects Research
- [124] Baele, S.J.: The ethics of new development economics: is the experimental approach to development economics morally wrong? *The Journal of Philosophical Economics* **7** (2013)
- [125] Humphreys, M.: Reflections on the ethics of social experimentation. *Journal of Globalization and Development* **6**, 87–112 (2015) <https://doi.org/10.1515/jgd-2014-0016>
- [126] Urquhart, A., Girling-Morris, F., Mason, E., Nelson-Pollard, S.: Global Humanitarian Assistance Report 2023. *Development Initiatives* (2023)
- [127] Tappis, H., Doocy, S.: The effectiveness and value for money of cash-based humanitarian assistance: a systematic review. *Journal of Development Effectiveness* **10**, 121–144 (2018) <https://doi.org/10.1080/19439342.2017.1363804>
- [128] Jeong, D., Trako, I.: Cash and In-Kind Transfers in Humanitarian Settings A Review of Evidence and Knowledge Gaps
- [129] Kurtz, J., Krishnan, V., Leape, V., Jovanovic, V., Ross, M., Macaranas, R.: A Million-Dinar Question: Can Cash Transfers Drive Economic Recovery in Conflict-Driven Crises?
- [130] WFP: Harnessing the power of money to help people survive and thrive. *World Food Programme Report* (2023)
- [131] WFP: Meeting Needs and Empowering People: WFP’s Cash-based Food Assistance in Uganda
- [132] UN: WFP Financial Inclusion Through Mobile Money
- [133] IRC: Seven Steps to Scaling Cash Relief Driving Outcomes and Efficiency International Rescue Committee —
- [134] Smith, G., McCormack, R., Jacobs, A., Chopra, A., Gupta, A.V., Abell, T.: The State of the World’s Cash Report: Cash Transfer Programming in Humanitarian Aid
- [135] Agur, I., Peria, S.M., Rochon, C.: Digital Financial Services and the Pandemic: Opportunities and Risks for Emerging and Developing Economies
- [136] Mikulak, M.: Cost-effectiveness in Humanitarian Work: Cash-based Programming
- [137] DFID: The Effects of DFID’s Cash Transfer Programmes on Poverty and Vulnerability: An

- [138] Maunder, N., Dillon, N., Smith, G., Truelove, S., Bauw, V.D.: Evaluation of the Use of Different Transfer Modalities in ECHO Humanitarian Aid Actions 2011–2014. <https://doi.org/10.2795/00448>
- [139] WorldBank: Cash Transfers in Humanitarian Contexts
- [140] Bailey, S., Harvey, P.: State of Evidence on Humanitarian Cash Transfers Background Note for the High Level Panel on Humanitarian Cash Transfers
- [141] ALNAP: Alnap: Strengthening humanitarian action through evaluation and learning (2023)
- [142] IRC: Cost Efficiency Analysis: Unconditional Cash Transfer Programs
- [143] OECD: States of Fragility 2022. <https://doi.org/10.1787/c7fedf5e-en> . https://www.oecd-ilibrary.org/development/states-of-fragility-2022_c7fedf5e-en
- [144] Hoop, J., Morey, M., Seidenfeld, D.: No lost generation: Supporting the school participation of displaced syrian children in lebanon. *Journal of Development Studies* **55**, 107–127 (2019) <https://doi.org/10.1080/00220388.2019.1687875>
- [145] Peachey, K.: Home Blog The 90Bn Question: Can We Reach 700 Million People in Response to COVID-19?
- [146] Bossuroy, T., Goldstein, M., Karimou, B., Karlan, D., Kazianga, H., Parienté, W., Premand, P., Thomas, C.C., Udry, C., Vaillant, J., Wright, K.A.: Tackling psychosocial and capital constraints to alleviate poverty. *Nature* **605**, 291–297 (2022) <https://doi.org/10.1038/s41586-022-04647-8>
- [147] Doocy, S., Tappis, H.: Systematic Review 28 Cash-based Approaches in Humanitarian Emergencies A Systematic Review
- [148] Attah, R., Barca, V., Kardan, A., MacAuslan, I., Merttens, F., Pellerano, L.: Can social protection affect psychosocial wellbeing and why does this matter? lessons from cash transfers in sub-saharan africa. *Journal of Development Studies* **52**, 1115–1131 (2016) <https://doi.org/10.1080/00220388.2015.1134777>
- [149] Hagen-Zanker, J., Ulrichs, M., Holmes, R., Nimeh, Z.: Cash Transfers for Refugees The Economic and Social Effects of a Programme in Jordan. www.odi.org/twitter
- [150] Ecker, O., Maystadt, J.-F., Guo, Z.: Can unconditional cash transfers mitigate the impact of civil conflict on acute child malnutrition in yemen? evidence from the national social protection monitoring survey. International Food Policy Research Institute (2019)
- [151] Lyall, J., Zhou, Y.Y., Imai, K.: Can economic assistance shape combatant support in wartime? experimental evidence from afghanistan. *American Political Science Review* **114**, 126–143 (2020) <https://doi.org/10.1017/S0003055419000698>
- [152] Grossman, G., Humphreys, M., Sacramone-Lutz, G.: I Wld Like U WMP to Extend Electricity 2 Our Village: On Information Technology and Interest Articulation. <https://doi.org/10.1017/S0003055414000331>
- [153] Davies, S., Pettersson, T., Öberg, M.: Organized violence 1989–2022, and the return of conflict between states. *Journal of Peace Research* (2023) <https://doi.org/10.1177/00223433231185169>
- [154] ITU: Measuring digital development Facts and Figures 2022. UN International Telecommunication Union Report (2022)
- [155] Andersson-Manjang, S.K.: The Mobile Money Prevalence Index (MMPI) A Country-Level

- Indicator for Assessing the Adoption, Activity and Accessibility of Mobile Money. GSMA (2021)
- [156] Hidrobo, M., Hoddinott, J., Peterman, A., Margolies, A., Moreira, V.: Cash, food, or vouchers? evidence from a randomized experiment in northern ecuador. *Journal of Development Economics* **107**, 144–156 (2014) <https://doi.org/10.1016/j.jdeveco.2013.11.009>
- [157] Kandpal, E., Alderman, H., Friedman, J., Filmer, D., Onishi, J., Avalos, J.: A conditional cash transfer program in the philippines reduces severe stunting. *Journal of Nutrition* **146**, 1793–1800 (2016) <https://doi.org/10.3945/jn.116.233684>
- [158] Crost, B., Felter, J.H., Johnston, P.B.: Conditional cash transfers, civil conflict and insurgent influence: Experimental evidence from the philippines. *Journal of Development Economics* **118**, 171–182 (2016) <https://doi.org/10.1016/j.jdeveco.2015.08.005>
- [159] Salti, N., Chaaban, J., Moussa, W., Irani, A., Mokdad, R.A., Jamaluddine, Z., Ghattas, H.: The impact of cash transfers on syrian refugees in lebanon: Evidence from a multidimensional regression discontinuity design. *Journal of Development Economics* **155** (2022) <https://doi.org/10.1016/j.jdeveco.2021.102803>
- [160] Moussa, W., Irani, A., Salti, N., Mokdad, R.A., Jamaluddine, Z., Chaaban, J., Ghattas, H.: The impact of cash transfers on syrian refugee children in lebanon. *World Development* **150** (2022) <https://doi.org/10.1016/j.worlddev.2021.105711>
- [161] Sandström, S., Tchatchua, L.: Do Cash Transfers Improve Food Security in Emergencies? Evidence from Sri Lanka
- [162] Hoddinott, J., Sandström, S., Upton, J.: The impact of cash and food transfers: Evidence from a randomized intervention in niger. *American Journal of Agricultural Economics* **100**, 1032–1049 (2018) <https://doi.org/10.1093/ajae/aay019>
- [163] Premand, P., Barry, O.: Behavioral change promotion, cash transfers and early childhood development: Experimental evidence from a government program in a low-income setting. *Journal of Development Economics* **158** (2022) <https://doi.org/10.1016/j.jdeveco.2022.102921>
- [164] Langendorf, C., Roederer, T., Pee, S., Brown, D., Doyon, S., Mamaty, A.A., Touré, L.W.M., Manzo, M.L., Grais, R.F.: Preventing acute malnutrition among young children in crises: A prospective intervention study in niger. *PLoS Medicine* **11** (2014) <https://doi.org/10.1371/journal.pmed.1001714>
- [165] Bonilla, J., Carson, K., Kiggundu, G., Morey, M., Ring, H., Nillesen, E., Erba, G., Michel, S.: Humanitarian Cash Transfers in the Democratic Republic of the Congo: Evidence from UNICEF’s ARCC II Programme. www.air.org
- [166] Schwab, B.: In the form of bread? a randomized comparison of cash and food transfers in yemen. *American Journal of Agricultural Economics* **102**, 91–113 (2020) <https://doi.org/10.1093/ajae/aaz048>
- [167] Schwab, B., Margolies, A., Hoddinott, J.: Impact evaluation of cash and food transfers for the seasonal emergency safety net in hajjah and ibb governorates, yemen endline report. International Food Policy Research Institute (2013)
- [168] Kurdi, S.: The nutritional benefits of cash transfers in humanitarian crises: evidence from yemen. *World Development* **148** (2021) <https://doi.org/10.1016/j.worlddev.2021.105664>

Appendix A Intervention & Analysis Details

The research protocol was approved by the London School of Economics’ Institutional Review Board (study number 89546) and preregistered in the AEA RCT Registry (study number 0010189). In SM C we discuss ethical considerations. The pre-analysis plan is registered at: <https://www.socialscisceregistry.org/versions/160809/docs/version/document>.

A.1 Research Design

Local context

Afghanistan is entering its third year of a complex humanitarian crisis combining economic contraction, political repression and ongoing concerns over stability and security. With a humanitarian appeal of \$4.6 billion in 2023, it is currently the world’s largest aid operation [71]. It is also the WFP’s largest recipient by value of cash-based transfers and commodity vouchers [72] and though multiple financial service providers exist in the country, digital aid programs are still nascent [73]. Afghanistan’s GDP contracted by 30-35% since the Taliban took over in 2021 and ever since it has been under the threat of “winters of famine” [74, 75].

Given the Taliban’s draconian restrictions on education, employment and freedom of movement, Afghan women are particularly vulnerable. Taliban edicts prevent women from travelling more than 75 kilometers without a male guardian and ban them from visiting public baths, restaurants and parks [76]. In December 2022, the Taliban banned female Afghan employees from working in non-governmental organizations, prompting major foreign aid groups to suspend operations.³⁷ The Afghan central bank’s reserves have been frozen since 2021, requiring the United Nations to fly in approximately 40 million U.S. dollars per week to support humanitarian operations [79]. Western donors remain concerned that diversion of these substantial flows is enabling the world’s most gender-repressive regime [80, 81] and that the Taliban are providing safe haven to international terrorist groups [82].

Mobile phone ownership has grown rapidly over the past two decades, from approximately 25,000 subscribers in 2002 to over 22 million subscribers in 2021 [83]. In a nationally-representative survey, 91% of respondents reported at least one member of their household owned a mobile phone (66% of respondents report personally using a mobile phone), while 46% of that subgroup reported having an internet connection [84].³⁸ For an overview of empirical research on digital payments in Afghanistan over the last decade, see [87–90].

Piloting Activities

Prior to starting the experiment, we conducted three small pilots ($N < 50$) to i) refine our survey instruments, ii) work out logistical processes including how to enroll beneficiaries and iii) identify patterns that needed to be taken into account before the full scale up of the intervention.

The first pilot involved around 30 women in Kabul. Our initial idea was to conduct the experiment without any face-to-face interaction. Thus, participants were contacted over the phone, invited to participate, and explained how to open accounts with the mobile payments provider. They received smaller payments (800 AFA) than in the actual experiment. A second, similar pilot was conducted a few weeks after the first one. These two initial pilots were intended to evaluate the survey instruments and sort out the logistics for the eventual scale up. From these pilots, it was apparent that participants were struggling to open accounts with the mobile payments provider and use their funds, as almost all participants had never used mobile money services or apps similar to the one used in this program, had never been part of the formal banking system and mostly had feature phones. While the mobile payments platform can be used with a feature phone, the process of creating an account is more complicated than when using a smartphone.

Due to these pilots, we decided to organize in-person registration sessions with around 50 women each, where potential participants would be introduced to the program, helped by the CDDO and representatives from the mobile payments provider. We conducted a third pilot with 52 women in

³⁷See, for example, Roberts [77]. In April 2023, the Taliban expanded the ban to include female Afghan employees of the United Nations [78].

³⁸Other studies have documented a substantial gender gap in women’s access to mobile phones [85] and access to the internet [86].

Kabul to test the logistics of the full scale up and revise the last versions of the survey instruments before conducting the registration sessions with all participants. This included conducting the in-person registration session and several rounds of phone follow-up surveys. This also allowed us check whether congregating women in a given place would cause problems. We observed much higher rates of usage of the funds sent to women and no meaningful problems during the onboarding process.

Beneficiary identification.

Our goal was to identify ~ 2400 vulnerable women in three Afghan cities (Kabul, Herat and Balkh) to be part of the intervention. To do so, we worked with the Community Driven Development Organization (CDDO), an Afghan organization that assists Community Development Councils (CDCs) in a wide-array of local activities.³⁹ Local CDCs identify potential beneficiaries through a community-based exercise (“Well-Being Analysis”) in which community members, elders and mullahs together categorize all community households into different socioeconomic groups (e.g. well-off, middle income, poor, very poor). Our participants come from the lowest group. Thus, participants are identified through a process relying on the community’s consensus of who is most vulnerable.

Onboarding and training.

Once the CDDO had identified enough potential participants, they were invited to onboarding sessions in each of the three cities. These took place during September 2022, with each session having between 24 and 80 participants (most had around 50). During these onboarding sessions, four activities took place. First, the program was described to the potential participants, and their informed consent was collected if they wanted to participate. Second, a baseline survey was collected. Third, each woman was helped to open an account with the mobile money platform we worked in this project. During pilots we conducted prior to the actual intervention, it became clear that participants needed help at first using the mobile money platform, as most of them were illiterate and had no prior experience with mobile money. Thus, after opening their mobile payment accounts, women completed a test purchase with a nearby merchant using the mobile payment platform.

Timeline and randomisation.

During the onboarding session, 2,422 women agreed to participate. Due to technical and logistical issues, the final experimental sample consisted of 2,409 women.⁴⁰ These women were then assigned to one of two groups. The “Early Group” received four bi-weekly 4,000 Afghani (roughly 45 USD) aid payments via the mobile payments platform first, between November 6, 2022 and December 20, 2022. The “Late Group” received the same payments between January 1, 2023 and February 12, 2023, after the Early Group had received all four of its payments. Figure E1 shows the timeline of the project. We believe that given the abject situation of participants, the ethical thing to do was for all participants to receive the aid payments eventually, and that is why we settled for this staggered intervention. During the onboarding sessions, all participants were told that they would receive the payments eventually.

Participants were randomly assigned to the Early and Late Groups, stratifying on two variables: First, the nahia (neighborhood) in which they registered. There were 16 nahias in total across Kabul, Mazar and Herat. Second, a measure of vulnerability. During the baseline, we asked participants in the past seven days, how many of their meals (breakfasts, lunches and dinners) had been only bread and tea, a measure of vulnerability and poor nutrition. We then created a categorical variable that indicated whether the participant was above or below the median number of bread and tea meals, which was used as the second stratification variable. We assign “misfits” independently across strata. Given that we have few strata and a single treatment arm, the number of misfits will be relatively small, so the risk of harming treatment fractions by independently assigning misfits across strata is low [see 91, for more details]. We pre-specified this as one of our main (nutrition) outcomes in the

³⁹The CDCs were established through local elections as part of the National Solidarity Program starting in 2004 [19–22, 38], where their primary job was to oversee block grants of development funding, and they were given a much broader range of local administrative authorities under the Citizens’ Charter, starting in 2016.

⁴⁰The 13 remaining women had issues with their phone numbers, including mismatches between phone numbers in the survey and transaction data or multiple women registered with the same phone number. While these 13 women were dropped from the experimental sample, we contacted each of them, resolved the issues, and transferred them the humanitarian payments as we did with the women in the experimental sample.

pre-analysis plan. As can be observed in our pre-analysis plan, the treatment groups are balanced in 17 of the 18 outcome and heterogeneity variables we collected at baseline.⁴¹

Sample characteristics.

Table E1 shows summary statistics of the participants. Most participants are widows (66%), have no education (63%) and are the main financial decision-maker in the household (66%), which on average has 6.31 members, indicating that the women in the study are highly vulnerable. The women are also poor and food-insecure: Out of 21 possible meals in the last 7 days, 13.76 consisted of just bread and tea. The household’s mean income in the past 30 days was just 357.97 Afghanis (roughly 4 USD), and any kind of employment was basically non-existent (only 3 reported that the head of the household had worked in the past 30 days). Among those who had a medical emergency in the last month, only 1% could afford the medicine needed. Unsurprisingly, these facts translated into extremely low levels of happiness (just 1% report being very or quite happy) and life satisfaction (mean score 3.53 on the Cantril Self-Anchoring Striving Scale which runs from 1 to 10).⁴²

In addition to being poor, vulnerable and mostly uneducated, our sample also had almost no experience with mobile money or other financial mechanisms. Only 2 women reported having transferred airtime in the past month, none reported having transferred money via a mobile money platform, and only 2 already had accounts with the mobile payments provider before the start of the program. Just four of them report that anyone in their household had ever had a bank account.

Data collection.

Data were collected through three ways. First, a baseline survey was completed during the onboarding session (see above). Second, we had access to participants’ transaction data from the mobile payments provider, which we could link to participants’ survey data. We obtained permission to do so during the consent process. Third, we conducted four rounds of follow-up surveys over the phone, from after the Early Group received its first payment to a couple of weeks after the Late Group received its last payment (see Figure E1).

A team of 12 female enumerators were tasked with contacting each participant once every month. Participants were randomly assigned a date to be contacted. Enumerators contacted participants on said date, and if the survey was not completed (because the participant did not pick up or was busy), they would attempt again on a different date and time. Participants received a 350 AFA payment as compensation for their time for completing their survey. Overall, for the two survey rounds that correspond to the experimental sample, completion rates were around 99%. In the first survey round, 29 interviews could not be completed (17 treatment, 12 control). In the second round, 26 interviews could not be completed (10 treatment, 16 control). Attrition is not differential by treatment status, as seen in Table E2, but there is a slightly higher non-response rate among control participants in the second round of surveys (although the difference is only 6 surveys and unlikely to meaningfully affect the results).

We pre-specified all analyses, including how the outcome variables would be constructed, and what our primary outcomes variables were going to be. We divided our outcomes variables in three families: Basic needs, wellbeing and informal taxation. In doing so, we provided a level of detail consistent with that articulated in [69], that two separate research assistants could take the data, and the pre-analysis plan, and provide back an identical analysis.

⁴¹The only unbalanced variable, whether the individual has been asked for any kind of assistance by local community leaders in the past month, is because only four individuals answered yes to this question, and they all ended up in the control group by chance. We do not believe this is a cause of concern.

⁴²A nationwide Gallup survey conducted just after the Taliban took control in August 2021 indicates that 94% of Afghans rate life satisfaction below 4 [92]. In addition, the mobile payments provider conducted a concurrent program in which women who already had an account with them or clicked on an advertisement in social media received four 800 AFA payments over a two month period. This helps benchmark the vulnerability and poverty of our sample. As expected given that this other sample is more tech-savvy, participants were significantly more likely to be married (42%), have some education (88%), and younger (28.44 years) than those in our study. They were also less likely to be the main financial decision-maker of the household (31%), had much higher incomes (5,325 AFA) and employment levels, and only 5.31 of their meals in the past week had been bread and tea only.

A.2 Analysis

Estimation strategy of baseline results.

We estimate intent-to-treat treatment effects for pre-specified outcomes based on the following specification:⁴³

$$Y_{itn} = \gamma_0 + \gamma_1 \mathbb{1}[\text{Early Group}]_{in} + \gamma_2 X_{i0n} + \gamma_3 Y_{i0n} + \gamma_4 \mathbb{1}[t = 2] + \varepsilon_{itn} \quad (\text{A1})$$

where Y_{itn} is the outcome of woman i in nahia n in survey round t .⁴⁴ Only the first two months of intervention are used for all primary analyses as those are the months in which we have clear experimental variation. X_{i0n} are the stratification variables (nahia fixed effects, and baseline needs). $\mathbb{1}[t = 2]$ is a dummy for the second survey round period (round fixed effect). For variables for which we have values at baseline, we control for the baseline values Y_{i0n} . Standard errors are clustered at the individual level, the unit of randomisation.

We conduct our analysis in accordance with our pre-analysis plan. For each of the three families of outcomes we have, we specified four different primary outcome variables. For needs, the four primary outcomes are the number of days in which the person skipped meals in the past week, an indicator for whether children in the household skipped meals in the past week, an indicator for whether household members have regularly eaten at twice a day for the past week, and the total number of meals that have consisted solely of bread and tea in the past week. For the informal taxation outcomes, the four primary outcomes are indicators for whether the participant has been approached by government officials to provide them with any kind of assistance (such as food or money) in the past month, whether the participant has been approached by local community leaders to provide them with any kind of assistance (such as food or money) in the past month, and the same two questions but for anyone else in the community. For wellbeing, the four primary outcomes are an indicator for whether the participant feels that the overall economic situation of her household is slightly or much better than 30 days ago, an indicator for whether the participant agrees a lot or somewhat with the statement “I am highly satisfied with my present financial condition”, an indicator for whether the participant says that taking all things together she is very or quite happy, and a score on how satisfied she is with her life as a whole these days (from 1, dissatisfied, to 10, satisfied). For each of these variables, participants were allowed to not answer the question if they did not want to.

The results for needs and wellbeing are presented in Table 1. These also include other variables that we denoted as secondary, which include measures on participants’ diets (whether they consumed rice, beans, vegetables, chicken, or dairy in the past week), whether they were able to buy medicine if they had an emergency, household’s income and employment status of the head of the household. To increase power and precision, for each family of outcomes, we also summarized the primary measures into a single summary index following [93], which we denote “KLIK Index”.⁴⁵ Given that our main experimental hypothesis is that direct aid payments will reduce immediate humanitarian needs and improve wellbeing, we control for the Family Wise Error Rate (FWER) for each family of primary outcomes, as our primary concerns relate to falsely rejecting the null that the program had no impact on humanitarian needs or wellbeing. We therefore take the more conservative approach of controlling the FWER rather than the False Discovery Rate (FDR). For secondary outcomes, we control instead for the FDR. These results using these multiple hypothesis corrections are shown in column 6 of Table 1.

The results for informal taxation are presented in Table 3. Panel A shows the corresponding results for the informal taxation questions as in Table 1. Column 5 includes a dummy for whether the participant responded yes to any of the four informal taxation questions, which better captures the

⁴³While in theory we could have estimated treatment-on-the-treated treatment effects, we did not have non-compliance issues and non-response rates were very low. Hence we only present intent-to-treat results.

⁴⁴Note that t can be either 0 (baseline survey), 1 (first round of follow-up survey, after the Early Group has received 1 or 2 aid payments) or 2 (second round of follow-up survey, after the Early Group has received 3 or 4 aid payments).

⁴⁵The index is created as follows: First, each measure is standardized by the pre-intervention values of the variable in the control group. Second, for those observations with missing values, these are imputed as the mean in the participant’s treatment group. For missing values in the baseline data, the imputation is done with the values at baseline, and for missing values in the follow-up data, the imputation is done with the values at the follow-up rounds. Third, all variables are aligned in the same direction, such that higher values indicate “better” outcomes. The final index is the equally-weighted average of z-scores of the index’s individual component variables. The final measure is then standardized (relative to the control group) to assist interpretation.

prevalence of informal taxation in these communities. However, note that this outcome was not pre-specified. Controlling for the FWER, none of the adjusted p-values for these four outcomes is below 0.1 (results not shown for brevity).

Given that participants were allowed to skip certain questions if they wanted, this means that the sample in Table 1 changes in each regression. We also provide results restricting the sample to only those individuals who answered all questions necessary to construct our primary and secondary outcomes, ensuring a constant sample across regressions. These results are presented in Table E6, which follows the same structure as Table 1. Results using this alternative approach are consistent with the baseline results in Table 1.

Estimation strategy of results over time.

We also evaluate how the results change in the second round of follow up surveys, after the Early Group has received 3-4 aid payments, compared to the first round of follow up surveys, after the Early Group has received only 1-2 aid payments. We do so by estimating the following specification:

$$Y_{itn} = \beta_0 + \beta_1 \mathbb{1}[Early\ Group]_{in} \times \mathbb{1}[t = 2] + \beta_2 \mathbb{1}[t = 2] + \beta_3 \mathbb{1}[Early\ Group]_{in} + \beta_4 X_{i0n} + \beta_5 Y_{i0n} + u_{itn} \quad (A2)$$

and testing (H_1) whether the effect of the treatment in the second round is statistically different from that of the control group, with $H_0 : \beta_1 + \beta_3 = 0$. Panel B of Table 3 shows the results for the informal taxation outcomes, while Table 2 shows the results for the needs (Panel A) and wellbeing (Panel B) primary outcomes.

List experiment.

A key question, given the context, is whether the Taliban government managed to capture any of the aid payments. Avoiding diversion of this sort is naturally important at scale. Participants might feel uncomfortable answering questions about informal taxation by community leaders and government officials, and so answer our informal taxation questions falsely. To provide additional evidence that the lack of informal taxation we observe is not due to fears of answering yes to our questions, we conducted a list experiment with participants. List experiments have been used extensively in the political science literature to gauge the prevalence of sensitive behaviors or situations, without individuals having to disclose that they have indeed done so. The basic idea is to create two lists of statements which are identical but for the fact that one of the list (“treated list”) has an additional statement that is the one the researchers want to learn about. Participants are asked how many of the different statements apply to them, and then by comparing the number of reported statements in the short and long lists, this allows estimation of the share many respondents have experienced the situation/behavior of interest. The basic idea is that respondents are provided cover as they are only signaling how many behaviors they undertook, not which specific behaviors [94].

We implemented a list experiment to gauge the prevalence of informal taxation in the population. There are two things to note. First, we decided to conduct the list experiment as an additional check *after* we had analyzed some of the follow up data and thus this analysis was not pre-specified. However, we follow the standard approach in the literature for analyzing list experiments. Second, we conducted the list experiment in the very last round of surveys, when the Late Group was receiving their 3rd and 4th aid payments. Thus, the timeframe is different from that of all the other analysis.

Concretely, we asked participants how many of 4 (control) or 5 (treatment) situations have happened to them in the past month. Both lists include the same initial four statements: “I have received some form of financial support from local authorities”, “I have borrowed money from a friend or family member”, “I have participated in an informal savings group” and “I have borrowed money from an informal loan provider”. The treatment list included in addition the statement “I have been approached by government officials or community leaders to provide them with any kind of assistance, like food or money”. Participants were explained that they only needed to give the total number of these situations that had happened to them in the past month, not which of the situations had happened to them.

We randomly assigned individuals to the treatment and control lists, stratifying the randomisation by surveyor and treatment status. We estimate the following specification to evaluate the results of

the list experiment:

$$\text{Number of Statements}_{isn} = \rho_0 + \rho_1 \text{Treated List}_{isn} + \rho_2 X_{isn} + v_{isn} \quad (\text{A3})$$

where s corresponds to the surveyor in charge of the survey. We control for strata fixed effect, X_{isn} . We do this combining the whole sample, for the Late Group only (which has been receiving aid payments for over a month) and for the Early Group only (which has not received any aid payments in over a month). Panel C of Table 3 shows the results of this exercise.

Checking treatment impact heterogeneity.

We look for differential treatment effects on outcomes by subgroup by estimation the following specification:

$$Y_{itn} = \mu_0 + \mu_1 \mathbb{1}[\text{Early Group}]_{in} \times \mathbb{1}[\text{Heterogeneity}]_{in} + \mu_2 \mathbb{1}[\text{Heterogeneity}]_{in} + \mu_3 \mathbb{1}[\text{Early Group}]_{in} + \mu_4 X_{i0n} + \mu_5 Y_{i0n} + \mu_6 \mathbb{1}[t = 2] + v_{itn} \quad (\text{A4})$$

where $\mathbb{1}[\text{Heterogeneity}]_{in}$ is a dummy for whether woman i in Nahia n belongs to a certain subgroup. We check for heterogeneity along the following dimensions: basic needs (total meals composed of only bread and tea), city of residence (Kabul vs. Mazar and Herat), whether the woman is able to leave the house at baseline, whether the woman is married, whether the woman is Pashtun, whether the woman has some education, whether the woman is above the median age, whether the woman is the household’s financial decision-maker, whether the household is above the median household size. For brevity, we only present results for the KLK indices of the three main outcome families in Figure E3.

Hypothetical cash versus digital aid

During the fourth survey round, we asked participants a hypothetical question to measure their willingness to pay to receive their aid payments in cash rather than digital. The question asked “We are hoping to use what we have learned from these surveys and from your experience with these payments to try to expand the program. While we do not have funding to do so at the moment, we are working to find it. In the future, we are also considering whether to give recipients the option to exchange the voucher for cash, rather than for goods at merchants. If we provide a cash out option, however, the fortnightly payments would be smaller because we have to pay a fee to make physical cash available.” and then proceeded by asking participants “If the fee was X AFN, would you prefer $4000 - X$ AFN in physical cash or 4000 AFN in HesabPay credit?”, where $X \in \{100, 300, 500\}$.

Results of this exercise are presented in Figure E4. Even with a fee as small as 100 AFN (2.5% of the total payment), almost 80% prefer receiving the aid digitally rather than in cash. A fee of 300 AFN that more accurately reflects the true costs of delivering the aid in cash results in 94% of respondents preferring digital aid. The Early Group has a slightly higher preference for digital aid than the Late Group, potentially because they have had more experience receiving the aid digitally, although responses are very similar across the two groups. While this choice was not incentivised, we believe that the results likely *underestimating* the true extent of preference for digital aid, as the benefits of cash are clear but the costs of distributing and collecting the aid in cash are not salient the way the statement is framed (other than the facilitation fee). For example, the cash option would entail mobilizing to a cash collection center, potentially waiting for hours to receive the payment, and being more easily identifiable as the recipient of the aid payments.

A.3 Identification Threats

Experimenter demand effects.

Our primary outcomes are self-reported survey data. Because of social restrictions at the time of the study in Afghanistan, we could not send (even female) enumerators to interview beneficiaries in-person. Moreover, subjects cannot be blinded to their treatment status. As such, there is potential for experimenter demand effects (i.e. the participants answering what they believed we wanted to hear, not their true answers). In particular, it is plausible that respondents might either indicate that they are doing worse than they are in fact in order to influence the experimenter to send more aid.

It is also plausible that respondents might want to overstate their wellbeing in order to provide more favorable evidence that might encourage policymakers to scale the program.

To assess whether this is a problem in this setting, in the last round of follow-up surveys ($t = 2$) we “primed” participants by telling them the purpose of the study quite explicitly and checking whether that information affects their responses. This exercise is similar in spirit to the work by [70]. More specifically, we randomly assigned individuals into two groups: a “primed” group hears the following statement just before the questions related to needs: “I would now like to ask you a few questions about how you and your family are doing. The goal of the CDDO and HesabPay program is to help you and your family meet basic needs, such as buying food, and we would like to see how you are doing in this regard. We will share what we learn from interviewing participants like yourself, with international organizations who are trying to help Afghans deal with these difficult times.” Thus, this group is explicitly told what we are expecting to find. The “not primed” group hears this placebo statement instead: “I would now like to ask you a few questions about how you and your family are doing.” We stratified the random assignment by treatment status and the enumerator that will conduct the survey.

We run two types of specifications. First, to evaluate whether primed individuals give different answers than not primed individuals, we estimate the following specification:

$$Y_{isn} = \psi_0 + \psi_1 \text{Primed}_{isn} + \psi_2 X_{isn} + \psi_3 Y_{is0n} + \omega_{isn} \quad (\text{A5})$$

where we control for strata fixed effects (surveyor and Early/Late Group status), X_{isn} and the baseline value of the dependent variable (when available), Y_{is0n} . Standard errors are clustered at the individual level, the unit of randomisation.

In addition, we test whether the prime affected participants’ responses differently depending on whether they belonged to the Early or Late Group, by estimating the specification:

$$Y_{isn} = \eta_0 + \eta_1 \text{Primed}_{isn} + \eta_2 \text{Primed}_{isn} \times \mathbb{1}[\text{Early Group}]_{isn} + \eta_3 \mathbb{1}[\text{Early Group}]_{isn} + \eta_4 X_{isn} + \eta_5 Y_{is0n} + \omega_{isn} \quad (\text{A6})$$

The results of the experimenter demand effect analysis are shown in Table E5. Column 2 shows the baseline estimates, $\hat{\gamma}_1$, column 3 shows the estimates for the overall experimenter demand effects, $\hat{\psi}_1$, column 4 shows the experimenter demand effects only for the Late Group, $\hat{\eta}_1$, and column 5 shows the (overall) experimenter demand effects for the Early Group, $\hat{\eta}_1 + \hat{\eta}_2$.

Borrowing and spillovers.

There are two additional concerns that might affect our results. First, given that we told participants that they would eventually receive the treatment (however we did not tell them when this would happen), it is possible that they could have borrowed money at the time, alleviating needs in the short run. However, it is unlikely that this affects our results. When asked at baseline how difficult it would be for them to raise 1,500 Afghans within a month in case of an emergency, only 7 women answered that this would be somewhat or very easy. Moreover, as explained above, our sample seems to have very little experience with financial instruments, and only one participant in the Early Group reported using her aid payments to repay debt.

The second concern is that there might be spillovers from treated to control households. Given that our sample comes from only 16 Nahias and they met other participants during the onboarding sessions, it could be that the Early Group women, when they were receiving the aid payments, helped the Late Group women. If anything, this would mean that we are *underestimating* the true effect of the intervention. However, our data indicate this is not likely to be a problem. Only around 30% of our sample reports knowing another woman receiving aid payments from our program. Among the subsample that reports knowing another participating woman, just 2% reports receiving any aid from another participating household (17 women in the control group).

A.4 Deviations from pre-analysis plan

As mentioned before, our pre-analysis plan (PAP) is registered at <https://www.socialscisearch.org/versions/160809/docs/version/document>. All the analysis was conducted after the submission of this PAP.

We stuck to the PAP as closely as possible, although there were a few instances in which we deviated. We do not report Treatment on the Treated (ToT) estimates because there were no issues with non-compliance and survey response rates were extremely high. Given that attrition was so low and not different across treatment groups, we also don't show Lee Bounds results. We don't report results on participants' satisfaction with the mobile money platform, HesabPay, but there were only 2 instances of participants stating that they were "somewhat dissatisfied" with HesabPay (out of 2393 responses in the Early Group), with 2358 reporting being "very satisfied". In addition to testing for experimenter demand effects for the whole primed sample (pre-specified), we also show results disaggregated by treatment group (not pre-specified). We also don't report results on the "Response timing" section of the PAP, as there were no significant results (results are available upon request). Initially, we had described a broader set of questions for the experts' survey, but we decided to cut the number of questions from six to four in order to reduce the burden on respondents. The four questions we did ask match the topics in the PAP. We created a table summarizing all the main (pooled) results, in which the sample was restricted to the participants who answered all the relevant questions across the two months of surveying, to allay concerns of comparability (see Table E6), including an alternative way of creating summary indices following [95], another popular approach for constructing indices. This was also not pre-specified.

We deviate slightly from the PAP in the diversion results. Initially, we had pre-specified presenting results for each of the four measures of diversion, as well as an index created by combining the four measures following [93], the KLK index. Upon further consideration, and after computing the results, we decided that combining the four different measures in a single index is not sensible given that it can lead to double counting if participants get confused being asked about others first and whether that should include them or not. Therefore, to avoid this double counting, in Table 3 we report instead results using a dummy for whether the respondent answered yes to any of the four diversion tables (column 5), which was not pre-specified. For completeness and robustness, Table E3 shows the results of the original, pre-specified KLK index (column 3), as well as results in which the two measures of diversion of others are combined into a single KLK index (column 1), the two measures of diversion of the participants themselves are combined into a single KLK index (column 2), and an index of the four measures together following [95] (column 4). None of these alternative indices was pre-specified. As can be observed, the results using the pre-specified KLK index in column 3 are only significant at the 10% level, while the other three alternative indices are all statistically insignificant at conventional levels. We also don't report the p-values adjusted for multiple hypothesis testing for the individual diversion measures, but are all above 0.25.

A.5 Interpreting magnitudes of needs results

The Afghanistan Cash and Voucher Working Group (CVWG) estimates a Basic Food Basket for a family of seven costs approximately 96 USD per month [96], roughly equivalent of two biweekly direct aid transfers of 4000 AFA and a monthly survey incentive of 350 AFA.⁴⁶ In mid-December 2022, we contacted 25 merchants serving beneficiaries to better understand spending patterns and realized prices. We confirmed that the most popular purchases were wheat flour, cooking oil and sugar, and using merchant-specific estimates of prices and volumes, verified that the cost of a typical basket of goods matched the total value of the aid payments and survey incentive.

It is natural to ask whether the estimated impacts on basic needs are in line with what we should expect given the size of payments – and more specifically, why hunger appears to persist throughout the length of the intervention. As noted above, we observe large reductions in each of the four primary outcomes under needs, and these measures kept improving throughout the two months of payments (Table 2, Panel A). While not eliminated, the number of days in which a person skipped meals in the

⁴⁶This basket was composed of 89 kg wheat flour, 21 kg domestic rice, 7 liters vegetable oil, 9 kg pulses, and 1 kg salt at prevailing exchange rates in August 2022. For reference, the Minimum Expenditure Basket including food, healthcare, shelter, and other components totaled almost twice as much at 181.36 USD.

past week declined after each payment (Figure E2, Panel A). To better understand the explanatory factors for this persistence, we completed a brief survey on January 2 & 3, 2023 with 52 randomly selected Early Group participants who reported any skipped meals in the second month of surveys (ie. December 2022). Consistent with the gradual phase-in of impacts, 31 respondents (60%) reported no adults had skipped meals in the last week. The remaining respondents provided explanations for persistent needs consistent with a wide range of potential economic factors including larger-than-average needs, inter-temporal substitution, inter-household risk-sharing, and non-consumption substitution. 9 respondents (17%) indicated the payment size was insufficient for everyone in their household to eat at least bread and tea at each meal, with several highlighting the difficulty of feeding larger families. 12 respondents (23%) indicating saving some food for future consumption, but typically only mentioned having one or two week’s worth on hand. 4 respondents (8%) reported sharing food with individuals outside their household, and 3 respondents (6%) reported bartering food for other expenses like medicine or rent.

For a comparison of magnitudes, albeit in a very different context, in 2013–14 UNHCR gave 100 USD per month for nearly six months (575 USD total) to Syrian refugee households in Lebanon and found that cash reduced the number of days/week that adults decreased their daily meal intake by 0.6 relative to the control mean of 3.2 [43].⁴⁷ In four other programs, cash did not result in any improvements in similar measures of food security. Unconditional cash in South Sudan and cash-for-work in Central African Republic did not reduce the number of days with skipped meals [99, 100], vouchers in D.R.C. did not reduce days relying on a variety of coping strategies that included meal skipping [101], and varying the number of cash disbursements (lump sum vs. three disbursements along with financial education and nudging) in Philippines did not produce any differential impacts on similar coping strategies [102].

Appendix B Cost analysis

B.1 Cost-Efficiency

Delivering aid digitally can be highly cost-efficient. We estimated three cost-efficiency quantities: the total cost-transfer ratio (TCTR), the cost-transfer ratio (CTR), and the the cost-per-beneficiary (CPB). The TCTR is the ratio of total program costs to the transfer value, the CTR is the ratio of distribution/administrative costs to the transfer value, and the CPB is the total program costs for at household.⁴⁸ All estimates are based on the variable operational costs of the mobile payments provider, NGO, & community development organization.⁴⁹ Table E11, Panel A, reports raw costs (a) disaggregated by category, (b) including/excluding onboarding, and (c) by treatment and control group which both received transfers but at different times. Panel B reports the CPB, including/excluding the transfer value, and including/excluding onboarding. All in, the total CPB is 192.00 USD consisting of four approximately 45 USD transfers amounting to 180 USD and the variable operations costs totaling 12 USD. Excluding the costs of onboarding, the total CPB is 182.44 USD. Here, the transfer amount is the same (180 USD) but the administrative costs are 2.44 USD. Panel C reports the TCTR and CTR, first including onboarding, and then excluding onboarding costs. When including all costs, the TCTR is 1.067, meaning that 6.7 cents is required for each dollar transferred to a beneficiary (CTR). When excluding the costs of onboarding, the TCTR is 1.014, meaning that 1.4 cents is required for each dollar transferred (CTR).

B.2 Cost-Effectiveness

Digital aid is also cost-effective. We compute cost-effectiveness ratios (CERs) dividing costs by treatment effects aggregated over participants and time [104], which are interpreted as the number of dollars required to achieve a defined impact (e.g., a reduction in days with skipped meals).⁵⁰ We first

⁴⁷Two other programs found positive effects, but the differing outcome measurements complicate effect size comparisons. Households reporting a day of skipped meals decreased by 10 percentage points for cash recipients in Bangladesh [97] and 23 percentage points for a bundled program that included cash in Afghanistan [98].

⁴⁸See White et al. [103] for the FCDO/DFID guidance on this approach.

⁴⁹Fixed costs, which consist of the mobile provider’s platform origination and standard maintenance costs, are excluded. Consistent with standard practice, we also excluded evaluation costs given that future operation of the program would not depend on evaluation costs. Due to the security environment and the vulnerability of our subjects, we were not able to collect additional data on indirect costs.

⁵⁰By improving the nutrition and well-being of beneficiary households, it is likely that program beneficiaries experienced related improvements, including higher household productivity, and that the community experienced related improvements,

estimated cost-effectiveness ratios with one of our key preregistered outcomes, days with skipped meals.⁵¹ Using the treatment effect estimates, we estimate the costs required to achieve a reduction of a day with skipped meals, which is computed as: treatment group costs / (# beneficiaries * treatment effect * # weeks). All CER estimates appear in Table E12. Most conservatively, the average cost to achieve a reduction of a day with skipped meals is \$31.58 USD: \$231,936.08 / (1208 beneficiaries * 0.760 fewer days with skipped meals per week * 8 weeks). Panel A shows that across total possible days (67,648), program participants skipped 7,344.64 fewer days of meals, a 10.857pp decrease. Because there were four bi-weekly transfers, estimates from the program’s second half (Weeks 5–8) and final week (Week 8) may better approximate a steady-state program effect. When based on estimates from Weeks 5–8 (Panel B), \$22.91 USD results in a reduction of a day with skipped meals, aggregating to 14.964pp fewer days of skipped meals for treated households. When based on estimates from Week 8, arguably the most appropriate given that the beneficiaries had now received the full \$180 transfer, \$16.90 USD results in a reduction of a day with skipped meals, which aggregates to 20.286pp fewer days of skipped meals for treated households.

Reducing days with skipped meals is important for many reasons, especially improving the nutrition and health of household members. Given the security environment, we were unable to measure nutrition and health directly. Instead we draw on the the Afghanistan Cash and Voucher Working Group (CVWG) estimates of calorie consumption per person-household-day [96] to estimate the increase in calorie intake associated with a reduction of a day with skipped meals, which allows the CER to be represented as the cost required to increase calorie consumption by a given amount. CVWG estimates that a day of meals amounts to 2,100 calories/person/day. In our sample, the average household has 6.31 members, making 13,251 the total possible calories/household/day. Based on these inputs and assumptions, a household of 6.31 should consume 92,757 calories per week, but on average skips 34,042 each week (Table 1, control mean). With this benchmark, we can estimate the cost required to increase the number of calories by a fixed amount. We scale the CER so that it is interpreted as the cost required to increase calories by 1,000 (Table E12, Panel B). Averaging across the full eight-week program period, we estimate that 2.38 USD in digital aid results in an increase of 1,000 calories: \$231,936.08 / (1208 beneficiaries * (97,323,825 fewer forgone calories * 8 weeks)). Based on estimates from the second half of the program, \$1.73 USD results in a 1,000 calorie increase. Based on estimates only from the program’s final week, \$1.28 USD results in a 1,000 calorie increase.

Because digital aid continued to have an effect after the program ended (see Extended Figure E2), CERs can reasonably include longer-run effects in which costs remain the same, but are now scaled by the cumulative reduction in days with skipped meals across all sixteen weeks.⁵² Based on these inputs and assumptions, the approximately 180 USD in digital aid to 1208 beneficiaries resulted in a reduction of 15,844.13 days that households skipped meals (7,344.64 during the eight-week program period and 8,499.49 in the eight-week post-program period) out of a possible 135,296 possible days across the sixteen weeks. This is a 11.71 percentage point reduction relative to not receiving assistance, making 14.64 USD the cost to reduce a day with skipped meals (\$231,936.08/15,844.13). The corresponding, cumulative calorie intake estimate is 209,950,567 fewer forgone calories for the Early Group (15,844.13 fewer days * 13,251 calories per household-day), making 1.10 USD the cost to increase calorie intake by 1,000 (\$231,936.08/(209,950,567/1000)).

such as better market conditions. Cost-benefit analysis characterizes the relationship between *all* direct/indirect benefits and all direct/indirect costs, both for the present as well as the future, by monetizing all benefits and then calculating a benefit-cost ratio, net present value, or an internal rate of return. Given the security context and the vulnerability of our subjects, unfortunately we were unable to collect sufficiently detailed information to assess these possible benefits and any other indirect costs that would make possible a formal cost-benefit analysis. Studying a vulnerable demographic in an oppressive environment such as Afghanistan means that certain types of data will always be difficult to obtain [105].

⁵¹The estimates rely on the following survey question: “Over the past seven days, how many days did you or any other adults in your household skip meals because there were not enough resources for food?”. When a respondent reports the number of days, we assume that all meals are skipped for all adults in the household for those days. This corresponds to what we observe in the data. That is, there are 21 meals a week, 14 of which are bread and tea meals. Participants rarely eat any of the other food groups, which leaves seven meals unaccounted for. At baseline, they report skipping meals on average 2.5 days, which means that assuming all the meals in a given day are skipped by all adults translates roughly into the seven missing meals that we observe. Relaxing this assumption predictably decreases cost-effectiveness. All inputs, assumptions, and estimates are available with the replication materials.

⁵²For the post-program period, the Early Group is compared to the Late Group’s *baseline value* (2.615) because in weeks 9–16 the Late Group is now receiving transfers and no longer constitutes a control. The Early Group’s 12- and 16-week values could be compared to those of the Late Group just before receiving treatment (2.546), but by that point the Late Group could have experienced other changes, including updated expectations about entering the program or any effect of receiving the monthly 350 AFA survey incentive, which almost all households received given the high response rates. Alternatively, they could be compared to the Early Group’s own baseline value (2.626) or an average of the Early and Late Groups’ baseline values. In practice, all three are similar and do not qualitatively change the inferences.

Appendix C Ethical Considerations

We obtained Institutional Review Board approval on 4 May 2022 from the London School of Economics (#89546). There is no local IRB in Afghanistan. After submitting our initial application with accompanying consent/instrument forms (22 April 2022), the IRB asked for one round of revisions (27 April 2022), which we resubmitted (3 May 2022) and then received approval (4 May 2022).⁵³ The IRB asked us to provide a written update after beneficiary registration procedures were clarified (but before the program began), which we submitted on schedule (30 June 2022) along with results of two small-scale pilots. During the study, we submitted an amendment with plans to carry out a survey of experts (20 November 2022), which the IRB approved (Study #145636 on 29 November 2022).

A collaboration of practitioners, local grassroots organizations, the digital payments provider, and academics co-designed the study. Local- and internationally-based Afghans either led or worked with each of these collaborative organizations and fully participated in all decision-making, helping to ensure representation of the views of the participants, sensitivity to possible risks, and fair distribution of the program’s benefits and costs.⁵⁴ The study went through a due diligence phase of several months in which the team met weekly to assess the feasibility of implementing the program ethically. After launch, the entire team continued to meet every week to assess progress and implement any changes deemed necessary. The team was committed to early termination of the program, evaluation, or both, if adverse events were to occur. Notably, although the research team conducted multiple rounds of surveying, the main NGO partner also conducted its own internal evaluations, which also involved interviews and surveys with participants, which they reported as independent checks on the research team’s evaluation. In what follows, we discuss the ethical dimensions of our study, which we organized around HHS [109] principles, and further address considerations specific to insecure, humanitarian crisis environments [67, 110–112].

C.1 Respect for persons

All participants were adults and due to low literacy were verbally informed about the study (and provided their verbal consent) in either Dari or Pashtu. The consent process occurred during the onboarding/baseline survey as well as with each successive survey wave. Specifically, when beneficiaries were registered for the program during the onboarding sessions described in SM A.1, they were asked to participate in a baseline survey in which detailed information about the study was provided. In particular, the informed consent process included discussion of possible risks and benefits, their right to skip questions or opt out entirely without penalty or loss of benefits, our commitment to data confidentiality, and advance notice that they would be invited to participate in follow-up surveys conducted by phone. Moreover, each time beneficiaries were invited to participate in follow-up surveys during and after the program, they were reminded about opt-out, skipping questions, and data confidentiality.

When a program provides large financial resources to potential participants, subjects could feel pressure to participate even if they have reservations. If that is the case, then informed consent to participate in the program could be insufficient. Thus, it was important that the implementation and research teams fully commit themselves to early termination of the program if any possible signs of harm were to arise. The entire team met weekly throughout the design and execution of the project and evaluated this possibility in each session. Notably, the local community council liaison and the two local leaders of the enumeration team attended these calls and weighed in from their perspective as those closest to implementation and evaluation. No significant issues were identified during the study.

For the surveys, although subjects were informed that they could opt out of the survey or skip specific questions without any penalty related to the program (or survey incentives), we nonetheless designed

⁵³As part of the approval process, the IRB requested: clarifications about the digital payment provider, the qualifications of enumerators, the inclusivity of the recruitment process, privacy/confidentiality of the subjects, attention to a careful assessment of digital technologies in research outputs, COVID safety measures, a suggestion for the survey instrument, and a post-approval written update of how the registration process proceeded in practice. We provided a comprehensive response to each of the IRB’s questions and requests, including providing the enumerator non-disclosure agreement and updated instruments.

⁵⁴Because we were using a new technology in an insecure and uncertain operating environment, we also designed the study to adhere faithfully to key international principles governing digital engagement: fair treatment, protection and accessibility of funds, prioritization of women, safeguarding of data, designing for individuals, transparency, interoperability, responsive, and accountability in the value-chain [106–108].

the instrument to avoid objectionable material. We vetted the instrument with all partners and made important modifications based on input from all partners, especially the local Afghan contributors. In particular, the survey did not ask sensitive questions, with the exception of the diversion questions. To reduce researcher-subject power differentials, which can undermine subject autonomy, the enumerator team was made up entirely of local Afghan women. The all-women enumerator team helped ensure greater sensitivity to positionality concerns, which shapes whether and how well enumerators can respect subject autonomy or recognize diminished subject autonomy.

Early Group and Late Group beneficiaries who elected to participate in surveys received compensation for their time, which was transferred using the same digital payments platform described in the study. The in-person baseline survey and each of the first three monthly surveys entailed one-time transfers of 350 AFA, and the final phone survey entailed a one-time transfer of 800 AFA. The amount was calibrated to be economically meaningful, and corresponded roughly to the average reported monthly income of households in our surveys. At the same time, as 350 AFA represented only 5% of the size of the monthly direct aid transfers, the incentive payment was intended to be small enough that beneficiaries would not feel pressured to participate in surveys.

C.2 Beneficence

Risks: We prioritized the long-standing principle of *do no harm* [109, 110]. In addition to designing the study to minimize the risk of potential harm,⁵⁵ we also proactively monitored the program and evaluation for evidence of any realized harm. We identified several key risks.

First, we anticipated that the biggest risk stemmed from the recent rise of the Taliban (August 2021), including the possibility that the Taliban would try to divert aid to themselves, or threaten subjects or local implementation partners for their participation. We took several steps to mitigate the risk, which involved: (a) implementation through community development councils who are social workers, social organizers, and other community advocates (many of whom are women) and evaluation by local Afghan enumerators, which could ensure sensitivity to the context and improved ability to detect interference, (b) distribution of program benefits of sufficient size to meet humanitarian needs responsibly, but not so large that they would attract unnecessary attention, and (c) readiness to engage in public messaging including with the *de facto* authorities to address possible concerns. In the surveys, less than 1% of participants reported any diversion. We quickly followed up on samples of those reports and learned that they were negligible and, in some cases, likely misreported. We also contacted participating merchants, and they confirmed that they had not been approached for funds by government or community entities. Because we had full access to the participants' transaction data, we also confirmed where and how the funds were spent, which confirmed little risk of diversion.

Second, diversion, threats, and retaliation would be possible primarily if the privacy and confidentiality of the participants were compromised. We sought to maximize privacy by conducting surveys by phone so that door-to-door presence would not unnecessarily expose participants. We sought to maximize confidentiality by being extremely cautious about what information we collected in the surveys and how we protected that information. We developed a data management plan, which included provisions to fully anonymize all personally identifiable information for any use outside of the research team. Although merchants were not direct research subjects, we also took steps to preserve their anonymity, including how precisely they are represented in resulting data and maps. As noted above, we were careful not to collect other sensitive information.

Third, a frequently cited concern with cash-based programs is intimate partner violence (IPV), although we note that recent rigorous studies and research syntheses indicate that IPV risks are in practice extremely small and confined to certain types of subjects and types of IPV [113–117].⁵⁶ We deliberately targeted female-headed households as this would help mitigate the possibility of IPV by program design. (Two-thirds of participants reported not having a partner and two-thirds also reported that they are the household's financial decisionmaker.) Given the challenging context and vulnerability of our subjects, we could not ask about intimate partner violence directly. Instead, we

⁵⁵See SM A.1 for the design details referenced below.

⁵⁶There may be isolated negative impacts for women with similar or higher education than their partner [118] or when transfers are relatively large [119]. They may also be more pronounced for emotional rather than physical violence [120] or threats rather than actual violence [121].

looked to possible observable indicators, including self-reported measures of happiness and within-household diversion, which did not suggest reasons to be concerned. As with possible government- or community-level diversion, 99 per cent of women reported that they had not been approached by anyone for money, and 95% reported that they alone made the decision on how to spend the funds. We did not receive any (even informal) reports of intimate partner violence whether from the survey enumerators or from the NGO’s own internal evaluation team, which were conducted separately.

Fourth, although research ethics have largely focused on direct human subjects, we took seriously the minimization of risks to the implementers and research teams [122, 123]. The implementer was an established community development organization, which had a long-time presence in all of the study sites, fully understood how to work with community development councils and community members, and had full legal approval to operate. The enumeration team operated entirely remotely, reducing the security risks associated with having a physical presence in communities and at the homes of beneficiaries.⁵⁷ Given that enumerators were not going door to door, remote data collection also helped insure participants’ privacy.

From the study’s inception, we sought the input of all partners, including especially local implementers and enumerators, and verified that they did not see any other risks of harm. Throughout the study, we repeatedly and proactively evaluated any realization of harm. Across all discussions with implementers and enumerators, we received no reports of retaliatory activity by the regime or others in the community. We also see no patterns in any of the data analysis that indicate other forms of harm that may have befallen beneficiaries or merchants, or more broadly the local implementation or evaluation teams.

Benefits: Although minimizing the risk of harm is essential, the principle of beneficence entails maximizing possible benefits while at the same time minimizing the risks to subjects and society [109]. In practice, quantifying direct and indirect benefits can be challenging and perhaps even misleading [124, 125], necessitating extreme caution about relying on a benefit-risk “balance sheet approach”. As such, we were careful not to over promise, acknowledging in our informed consent communication that: “There may or may not be any direct benefits to you from participating in this study.” That said, because the program consisted of a cash-based value voucher, redeemable at nearby merchants selling food and household items, the program reasonably provides some valuable humanitarian relief directly to participants, which is also an important justification for conducting a randomized trial during a humanitarian crisis [67]. Although any subsequent scale up of digital aid should be considered prospective and, therefore, not weighed heavily (or at all) in calculations of potential societal benefits, we note that there could be downstream benefits to delivering aid in a decentralized, transparent, and cost-effective way to address the challenges of operating in such a difficult political and security environment.

C.3 Justice

Extreme need in Afghanistan far exceeded available program resources, making it critical that the program reach the most vulnerable. The program was implemented through community development councils, which were integrated in the communities, and had established practices for identifying households with extreme needs who were not already benefiting from other aid programs. Before launching the study we conducted two small-scale pilots and found that nearly 80 per cent had reported skipping a meal or cutting down on meals in the past 30 days. Data collected in the full study further validated these levels of need.

Beneficiaries were randomly assigned to treatment and control using a phase-in/waitlist design so that program benefits were not denied to anyone deemed eligible through the targeting and recruitment process. Although all study participants eventually received the program, a relevant question is whether half of those deemed eligible should be required to wait during a crisis. Put differently, is a randomized design appropriate in the midst of a humanitarian crisis? We addressed this in Section 4 and further note that in contrast to the relatively long timeline of many randomized programs,

⁵⁷The baseline survey was collected in person by representatives of the community development organization rather than the enumeration team.

the wait time was two months and both treated and control households received their assistance during the lean or hungry season. Given the high need in the entire sample, randomization was the most equitable approach for deciding on whether a household received assistance in the first versus the second two-month period. Regardless of whether a household was assigned to the Early Group (treatment) or Late Group (control), they all received the survey incentive each time that they participated. As such, those in the Late Group control did not accrue costs to their time that the Early Group did not, which is important among other reasons because if they were to attrit before receiving the program transfers then they would have paid higher costs.

Because researchers and society might derive benefit from the study, we also informed possible participants so that they were aware. In particular, after noting that they “may or may not benefit directly” we acknowledged that researchers might benefit saying “The investigator, however, may learn more about the usage of digital payment platforms in Afghanistan and how to deliver humanitarian aid” and then raised the possibility of broader benefits for society saying “and society may benefit from this knowledge.”

Finally, although not acknowledged to program participants, we note here that the funds provided by the donor were used entirely for programmatic purposes with the intended beneficiaries receiving nearly all of those program funds, as the cost-efficiency section discussed. The evaluation costs were funded entirely by external resources (in particular a grant from J-PAL’s CVI initiative, study/award GR-1969, with funding provided by UK FCDO).

Appendix D Related Work

D.1 Cash-based transfers in humanitarian contexts

Cash-based programs have become one of the most common foreign aid modalities, with nearly 17% of the world’s population (1.36 billion people) receiving some form of cash assistance during the pandemic period alone [9]. The evidence base for cash programming is extensive, with documented positive impacts in almost all categories of development outcomes, including poverty reduction, education, improving financial outcomes, promoting human development, empowering the vulnerable, and social cohesion [13, 115]. While the vast majority of cash-based programs are carried out in developing, but stable contexts, recent years have seen tremendous growth in humanitarian cash-based programming. The amount of cash- and voucher-based humanitarian assistance has more than doubled since 2017 (3.3 billion USD in 2017 to 7.9 billion USD in 2022), now comprises 20% of all humanitarian assistance globally, and appears to be growing [126]. To date, there has been very little evaluation of the impact of cash-based humanitarian assistance. In recent systematic reviews of thousands of humanitarian cash-based assistance programs, only a very small set included rigorous evaluations of impact [112, 127, 128], the evidence is especially thin in contexts facing the most severe and protracted humanitarian crises [129].

At the same time that humanitarians are increasing their use of cash-based programming, they are simultaneously shifting from physical to digital distribution [130–132]. Although physical cash-based programs are unlikely to disappear entirely, digital aid is clearly becoming the preferred delivery mechanism for cash and vouchers, with growing policy support [62, 130, 133–135]. The evidence on digital cash-based programs also comes primarily from retrospective performance evaluations or synthetic reviews, and appears primarily in policy reports rather than peer-reviewed outlets. According to GiveDirectly’s *Cash Evidence Explorer*, the most comprehensive public repository of systematic evaluations (N=332 as of 16 August 2023), there is extremely little evidence on digital aid in humanitarian contexts. In our review of the database, we found only four digital cash-based programs in humanitarian contexts.⁵⁸

⁵⁸We text-mined the abstracts and target beneficiary fields of the 328 papers for the following relevant keywords (and variations on these words): women, vulnerability, conflict, violence, war, humanitarian, crisis, emergency, repression, mobile, digital, phone, ATM, bank, e-card, and electronic. Restricting the sample to studies with any of those keywords (Boolean OR operator) resulted in 97 papers, of which 90 presented evaluation results, of which there were approximately 36 unique programs (i.e., there could be multiple evaluations of the same program, such as with Progresa / Oportunidades in Mexico, though we note that the database does not have a unique program identifier and so we also coded the abstracts for program titles to code uniqueness). We then hand-coded the resulting subset of studies and found a number of false positives, for example, the keyword *violence* returned some studies that addressed intimate partner violence as opposed to government, rebel, or structural violence.

To further map the evidence base specific to our study, we examined systematic academic reviews, policy evidence reviews, evaluation databases and clearinghouses, and broader policy discussions, which together covered thousands of cash-based transfer programs.⁵⁹ For purposes of coding relevant cash-based programs, our inclusion criteria consisted of (a) cash-based transfers (including vouchers) needed to comprise a core part of the intervention, (b) the program needed to be carried out during a humanitarian crisis,⁶⁰ and (c) the program needed to be evaluated using a method that established a credible counterfactual (i.e., randomized controlled trial, natural experiment, regression discontinuity, and difference-in-differences).

In all, we identified 23 relevant cash-based transfer programs summarized in Table E13. Sixteen of these studies are published in peer-reviewed academic outlets, five are academic working papers, and two are evaluation reports. The programs are sorted by the current level of fragility of the country in which they occur, moving from the least fragile (top) to most fragile (bottom) [143].⁶¹ Column 2 reports the country in which the cash-based intervention took place. Sometimes the intended beneficiaries were refugees from another country, but located in a host country, and sometimes the intended beneficiaries were local to a country. Regardless, target beneficiaries were high need or otherwise vulnerable. Column 3 reports the type of cash-based assistance, specifically identifying cash and vouchers, and the most relevant alternative, food assistance. Some of these programs include other/combined treatments, which we note in the extended discussions. Column 4 reports the delivery mechanism, whether digital, physical, or some combination.

Sixteen of the 23 programs occurred in conflict-affected countries (seven targeting displaced populations and nine targeting poor/vulnerable households) and seven in countries facing high natural disaster risk. Collectively, the programs cluster in 13 countries, but only three of which currently rank among the world’s most food insecure. Although the populations covered in these 23 programs are highly vulnerable, all have been located in contexts where host governments do not actively interfere in the provision of humanitarian assistance. Programs that address conflict-based humanitarian crises, for example, frequently occur in displacement camps in which host governments and humanitarians typically attempt to cooperate. Underscoring the relatively favorable donor-government environment, [144] found that 99% of their study sample (Syrian refugees in Lebanon) received other forms of humanitarian assistance. Our collective understanding of cash-based programming in humanitarian crises has little to say about protracted crisis contexts that sit at the confluence of extreme need, high insecurity, and oppression [129], where social protection systems are also extremely underdeveloped [145].

Across the 23 programs, 12 distributed cash physically, three distributed cash through ATM cards, three distributed through mobile money accounts, four distributed through some combination of digital and physical depending on the program mandate, and one program did not specify the delivery mechanism. In the 23 programs we reviewed, digital programs were implemented mostly in the *least* insecure/oppressive countries, whereas physical transfers occurred primarily in contexts characterized by *greater* insecurity and oppression.

D.2 Nutrition, psychological well-being, and diversion results

To contextualize our experimental findings on nutrition and psychological well-being, Table E13 reports whether the 23 reviewed studies empirically examined the same (or similar) outcome categories that we did, including whether they found positive, null, or negative effects. None of the studies carried out systematic tests for diversion, so we only include a ✓ if the study made any observation

⁵⁹We culled general information about cash-based humanitarian programming and also coded details of relevant cash-based humanitarian programs. These include academic review papers [64, 112, 115, 127], policy evidence review papers [107, 128, 136–140], program evaluation databases [9, 13, 141], and policy reports [62, 126, 130, 134, 142].

⁶⁰We focused on conflict- and natural-disaster-based humanitarian crises because rigorous evaluations of cash-based programs during health outbreaks have not been conducted [128, 3]. With the onset of COVID-19, many cash-based programs were implemented around the world, but overwhelmingly in contexts where governments and humanitarians alike had a shared interest in mitigating negative impact. Gentilini [9] documents that the overwhelming majority of nearly 4,000 cash-based programs occurred in high- and middle-income countries and, in most cases, as part of existing government response systems. Furthermore, although the evidence base is growing, rigorous evaluation appears to be following patterns in programming, overwhelmingly confined to wealthier (middle-income) countries.

⁶¹We list the programs in Ecuador and Philippines at the top because neither country appears in the OECD’s list of fragile countries. Lebanon and Sri Lanka are not on the core list either, but the OECD explicitly flags them as they show key early-warning signs.

about levels of diversion in their context. To supplement the summary table, Table E14 provides details about specific study context and findings.

Nutrition. Most of these studies (21 of 23) considered a nutritional outcome in some form, although there is considerable variation in the indicators each examines. Sixteen of the 21 find that cash-based transfers improved nutritional outcomes, five found no relationship, and none of the studies found any negative effects. Twenty of 21 examined at least one indicator of food quantity (e.g., food expenditures or calorie consumption). Eighteen of 21 examined at least one indicator of food quality (e.g., dietary diversity). And 16 of 21 considered at least one indicator of negative coping strategies (e.g., meal skipping). These patterns add support to a growing body of evidence documenting that humanitarian cash-based programs have positive impacts on a variety of nutrition outcomes. In a systematic review of humanitarian cash-based programs, [128] reported that existing studies most frequently consider basic needs outcomes, with eight of 20 studies considering food security (primarily food consumption) and 4 of 20 studies examining dietary diversity, most of which demonstrated positive impacts. Moreover, five studies examined reductions in negative coping strategies (e.g., meal skipping) and three found positive impacts.

Psychological well-being. In contrast to nutrition outcomes, Table E13 shows that limited attention has been given to the possible psychological benefits (or costs) of cash-based programs. Only seven of the 23 studies considered some psychological well-being outcomes and six found positive effects. Like our study, [101] found a positive impact (+0.32sd) on an overall psychological well-being index at the end of six weeks, which attenuated to some extent, but remained remarkably durable at the end of one year (+0.18sd and statistically significant). They also measured life satisfaction, where they found a positive impact at six weeks (18% higher in treatment; control: 3.29 on 10-pt scale), but the result disappeared by the one-year mark. [146] find what was perhaps the strongest impact of cash on psychological well-being, identifying short-, medium-, and long-term effects of cash relative to control. Importantly, like the other treatment arms, the cash was accompanied by coaching, savings, training, and market facilitation, making the overall bundle difficult to compare to other cash-only programs. In their review, [128] also report a much smaller set of studies that consider psychological well-being. They only identified two (of 20) studies reporting subjective well-being results: the [101] study (just discussed) and a study of Syrian refugees in Jordan (that did not meet our inclusion criteria) that found isolated improvement on a single measure of self-esteem, but notably no effect on a measure of satisfaction, the closest to our life satisfaction indicator. Similarly, in their systematic review, [127] found no rigorously evaluated studies that addressed psychological well-being.⁶²

Diversion. The extent of diversion in these studies is essentially unknown as four of the 23 studies addressed diversion, only one of which attempted to measure its extent. [42] measured the amount of the cash and vouchers actually received and found little leakage, [41] asked participants how much of the grant they had to give to other household and community members and participants reported less than 1% of the grant, and [43] asked about robbery specifically and report no occurrence. [63] mentioned that the implementer feared diversion, and therefore hired security to guard cash during transport, which comprised the largest share of their costs, but did not report measuring diversion. Many policy reports have downplayed concerns about diversion, suggesting that diversion risks are likely overstated and that cash is no less prone to diversion than other types of aid [60–62, 150]. As far as we are aware, no systematic tests of diversion in a cash-based context have been undertaken. In our study, we experimentally tested for diversion, both using direct and indirect questions, and found no evidence of its presence. We even interviewed key stakeholders, such as merchants, to understand alternative points where diversion could have occurred, and found little indication.

D.3 Additional advantages

The last four columns of Table E13 report whether the programs (a) had high usage (mobile delivery only), (b) decentralized their aid distribution, (c) could transparently track aid receipt and usage, and (d) reported on and delivered aid cost-efficiently or cost-effectively.

Usage. Because tech-illiterate populations are largely poor, rural, elderly, and female [145], it is important to assess whether digital transfers will reach those most vulnerable and be used effectively.

⁶²Their earlier review of evidence [147] identified several mixed-methods evaluations that suggested a positive impact. See, for example, Attah et al. [148] and Hagen-Zanker et al. [149].

If they cannot, then digital cash-based approaches may hurt intended beneficiaries relative to what may have occurred with physical cash or in-kind food distribution. In India’s PMGKY program intended to provide COVID relief, for example, the government transferred benefits to the accounts of over 200 million beneficiaries, but being female, illiterate, and living in a household without a smartphone resulted in extremely low use of digital payments, which ranged from 1% to 3.9% [9]. Of the ten studies with a digital component, five relied at least in part on mobile platforms, meaning that the beneficiary needed to use the phone for receipt/withdrawal of funds.⁶³ Of these five, only one reported on usage, noting that most of the time the funds were withdrawn immediately and, where not, they were used up eventually [151]. The remaining four did not report usage explicitly, but descriptions were suggestive of substantial usage, though we acknowledge that this assessment may be overly generous. Our own results in one of the most challenging environments, and with an arguably more vulnerable demographic than most other studies, are encouraging, though we emphasize the importance of user-centered design in our registration and training procedures. Near 100% usage by a largely illiterate and vulnerable population in a context of extreme oppression and insecurity suggests genuine potential for technology to “flatten access” to otherwise marginalized populations [152].

Decentralization. Because most cash-based assistance has been centralized with a single (or small set of) donors and can depend critically on a small set of distribution points, often run by external NGOs or temporary collectives, decentralizing assistance is another key advantage to consider. Only five of the 23 programs that we reviewed could be considered decentralized, capable of facilitating quicker access and use of funds. Of those five, three distributed cash on ATM cards that could be withdrawn at any ATM, any number of times, and in any amount, which were less constrained than ATM cards only redeemable at artificial NGO-operated mobile ATMs but still dependent on ATM penetration, which is unlikely to be high in most crisis contexts. It is unclear whether funds from other organizational or individual donors could add funds to these ATM cards. Two of the five programs were based on mobile money platforms that did not strictly require cash out, but in practice nearly all recipients cashed out quickly and fully, thus making the aid dependent on mobile agent penetration. In contrast, our digital voucher approach could be agnostic as to the donor, and could be redeemed at a variety of merchants relatively close to beneficiaries.

Transparency. All 23 studies we reviewed report only survey-based measures of activity. Still, some of the approaches could, in theory, lend themselves to greater transparency of receipt/usage of funds. We thus coded programs as transparent if their delivery technology could *potentially* capture non-survey-based receipt/usage data. Only three of the programs could provide transparency beyond the point of funds distribution or withdrawal. For physical cash, ATM cards, and mobile transfers with required cash-out, subsequent usage cannot be traced. Of the three approaches in which mobile transfers did not require cash out (though nearly all recipients still cashed out), activity could be tracked to understand the uses of assistance and also detect diversionary efforts. Policy makers have long cited the transparency gains of moving to digital cash [62], but in this context there remains almost no evidence of entirely mobile transfer systems, which can ensure end-to-end tracing from deposit to purchase.

Cost Analysis. 12 of the 23 reviewed programs report cost-efficiency or cost-effectiveness metrics. (Table E13 reports the derivable total cost transfer ratios (TCTRs) for nine of the 12; the remaining three, which we tag with a ✓, either reported on bundled programs for which we could not disaggregate the costs or did not include the precise estimates in their reports.) The cost efficiency of our digital aid program is the lowest of any of these programs when comparably counting cost categories. Although all of the reviewed studies focus on variable operational costs, they vary in the specific cost categories they count as relevant. In particular, very few studies include targeting/recruitment costs, and are thus comparable to our TCTR of \$1.014 (1.4 cents required to deliver 1 dollar of aid), though we think the inclusion of targeting/recruitment costs is more appropriate given that these costs are likely to be present in most crisis contexts.

The cost efficiency of our direct aid intervention is 40% of the World Food Programme’s global average, 48% of IRC’s (International Rescue Committee’s) most cost-efficient program, and 45% of ECHO’s (European Commission Directorate General for Humanitarian Aid and Civil Protection) most cost-efficient program. We do not have access to the disaggregated data used to compute

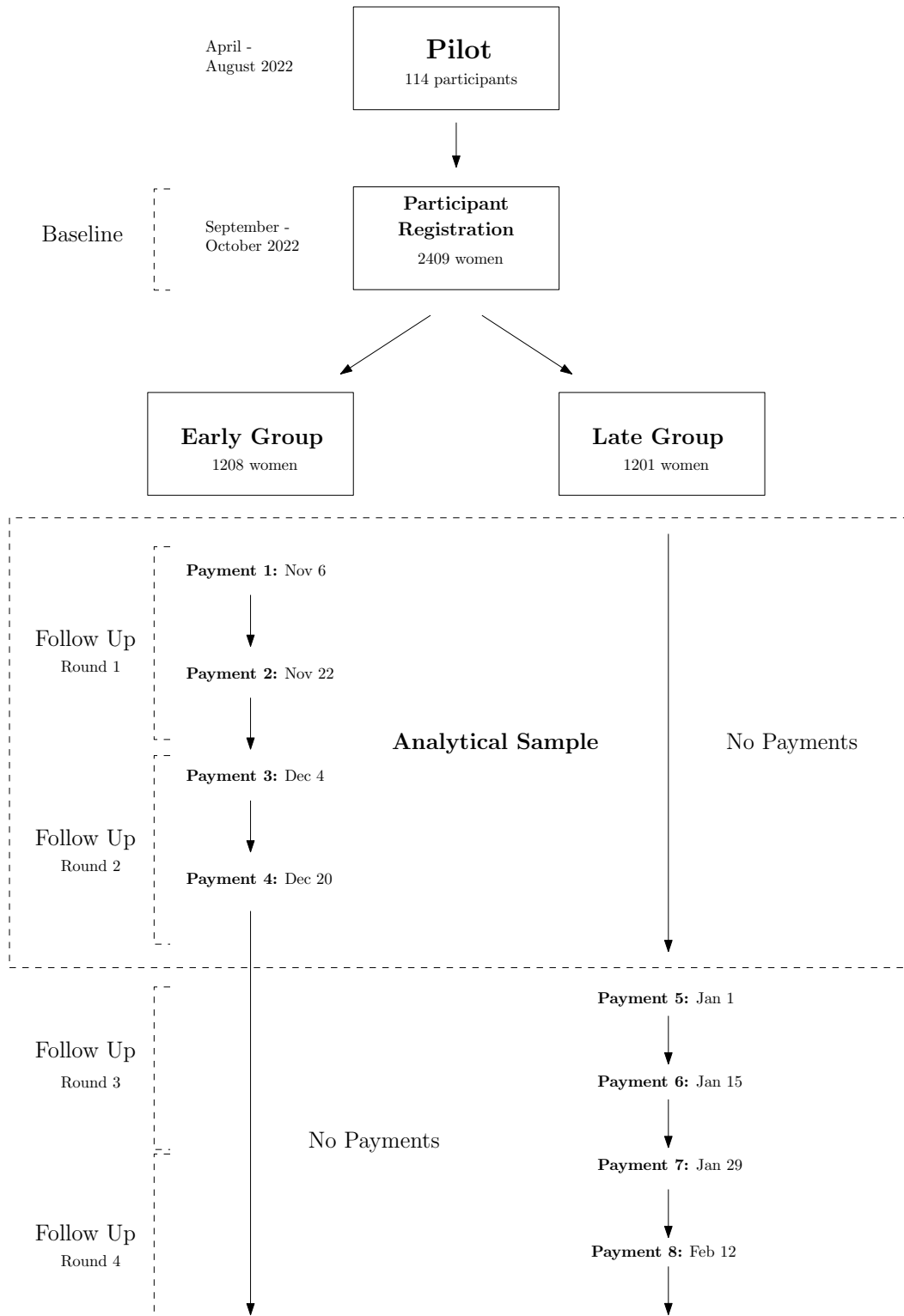
⁶³The remaining five relied primarily on ATM cards, which are less demanding technologically, and also confer fewer advantages such as decentralization and transparency.

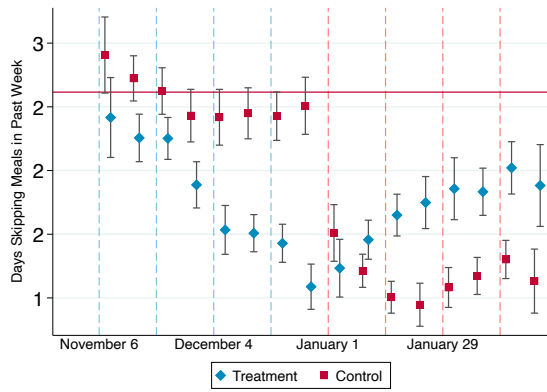
the WFP's or IRC's cost efficiency estimates, but [138] examined a large number of ECHO-funded projects and report a variety of disaggregated statistics. The average total cost transfer ratios range from \$1.15 (cash for refugee response) to \$2.81 (cash for complex emergencies). Their TCTR estimates of cash for slow-onset and sudden-onset crises are \$1.64 and \$1.39 respectively. Their estimates for voucher projects are \$1.54 (slow onset), \$1.81 (refugee response), \$2.11 (complex emergency), and \$2.72 (sudden onset). They disaggregate their cash projects by delivery mechanism, reporting average TCTRs of \$1.32 (ATM card; 2 projects), \$1.64 (mobile phone; 7 projects), \$1.66 (mix of bank transfer and cash in envelope; 3 projects), \$1.97 (cash in envelope; 22 projects), and \$2.03 (bank transfer; 13 projects). They also disaggregate vouchers and report TCTRs of \$1.31 (electronic), \$1.76 (paper), and \$2.25 (voucher fair). Their lowest TCTR is more than double that of our digital aid program in Afghanistan, and their most relevant categories are many times higher than ours (electronic voucher: \$1.31; mobile phone: \$1.64).

The cost-efficiency of digital aid in Afghanistan is considerably lower than all other humanitarian programs for which we could find estimates.⁶⁴ The cost-efficiency of our study was substantially lower, in part, because our approach was consistent with the *High-Level Panel on Humanitarian Cash Transfers* guidance to work with the private sector [62], which, among other benefits, reduced potentially significant start-up costs. In one of the most careful experimental comparisons, Aker et al. [63] contend that mobile transfers are more efficient than physical cash only when a mobile network infrastructure is available, convenient, and clear [128], which the private digital payments platform provided in our context.

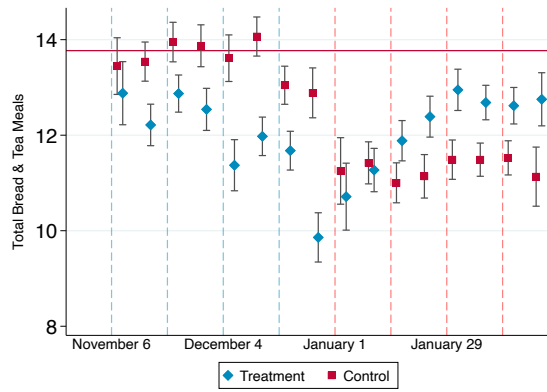
⁶⁴These cost-efficiency estimates also compare favorably to non-humanitarian social transfer programs — Kenya CT-OVC, Nigeria CDG, Mexico PROGRESA, and Kenya HSNP — which reported TCTRs at each year of 2–5 year programs [103], estimates that, from the beginning to the end of these programs, TCTRs ranged from 2.63 to 1.34 (Kenya CT-OVC), 2.04 to 1.40 (Nigeria CDG), 2.34 to 1.05 (Mexico PROGRESA), and 2.41 to 1.21 (Kenya HSNP). Only Mexico PROGRESA had a lower TCTR than our study, but this lower TCTR was only obtained in year 4 of operation, once the program had gone to scale. Furthermore, these four programs were carried out in more favorable implementation environments, where costs should be expected to be lower than in a context such as Afghanistan.

Appendix E Figures & Tables

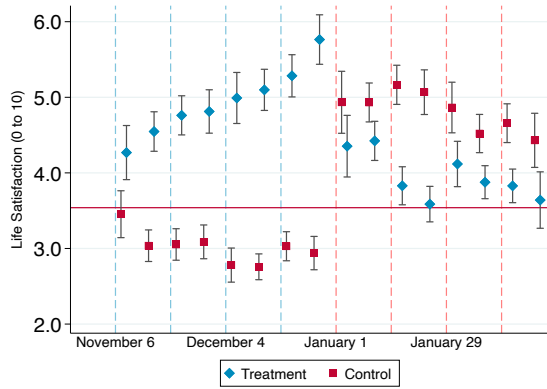




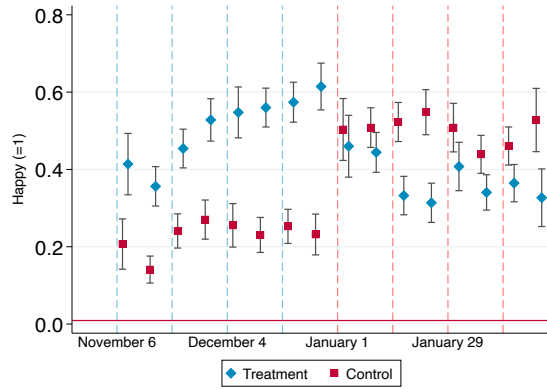
(a) Skipped Meals



(b) Total Bread and Tea Meals



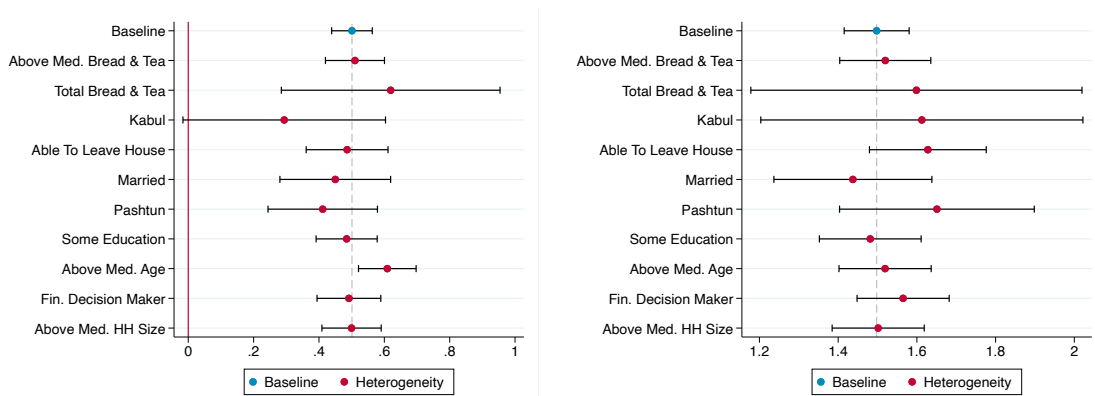
(c) Life Satisfaction



(d) Happy

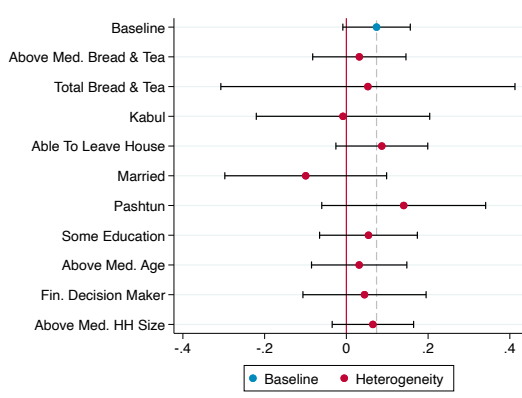
Fig. E2: Effects over Entire Duration of Program

Notes: The red horizontal line corresponds to the mean of the variable at baseline across all individuals in the sample. Red squares (blue diamonds) show the mean value of the variable among Late Group (Early Group) respondents divided in two-week bins since the start of the payments. Blue (red) vertical lines represent the dates in which Early Group (Late Group) participants received their payments. Note that this is the only analysis in which we are using/showing data for the period after the Late Group starts receiving payments (i.e. while the rest of the paper uses only data from survey rounds 1 and 2, this shows in addition data from survey rounds 3 and 4, after the Late Group starts receiving payments). Bars show 95% confidence intervals.



(a) Needs

(b) Wellbeing



(c) Informal Taxation

Fig. E3: Heterogeneity of Treatment Effects

Notes: The blue dots and the gray-shaded vertical lines correspond to the baseline estimate $\hat{\gamma}_1$ from Equation (A1). The red dots correspond to $\hat{\mu}_1 + \hat{\mu}_2 + \hat{\mu}_3$ from Equation (A4). Bars show 95% confidence intervals.

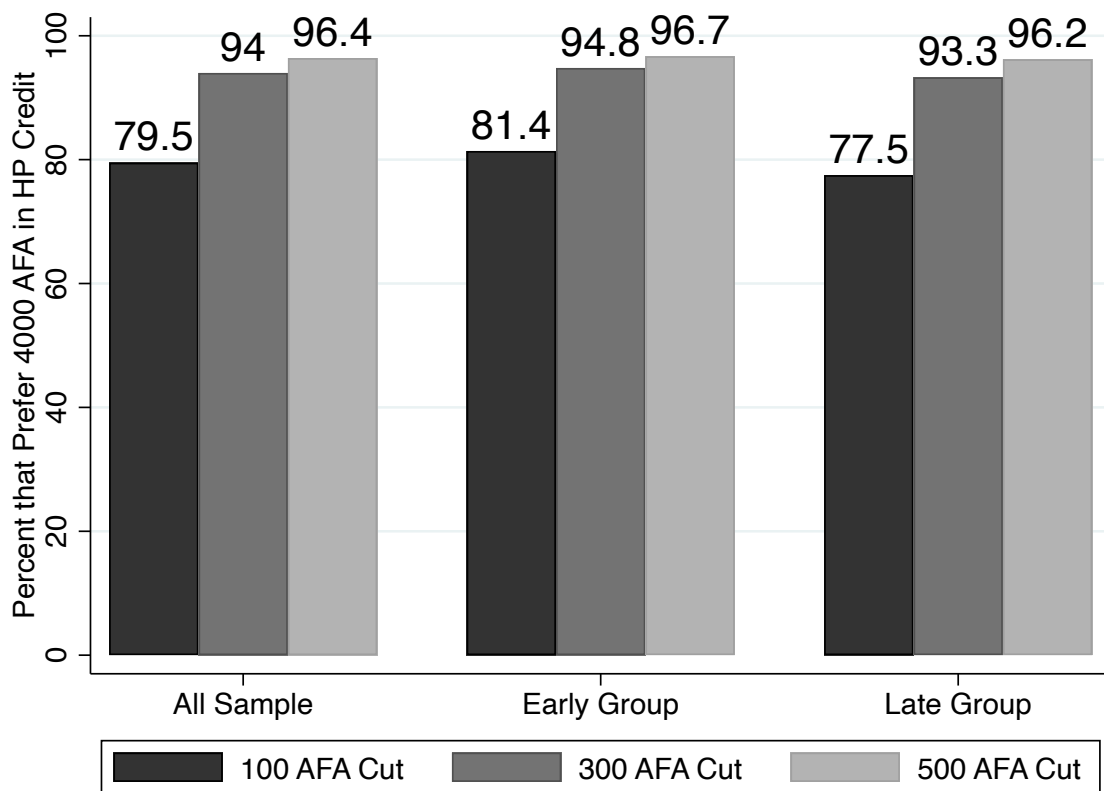
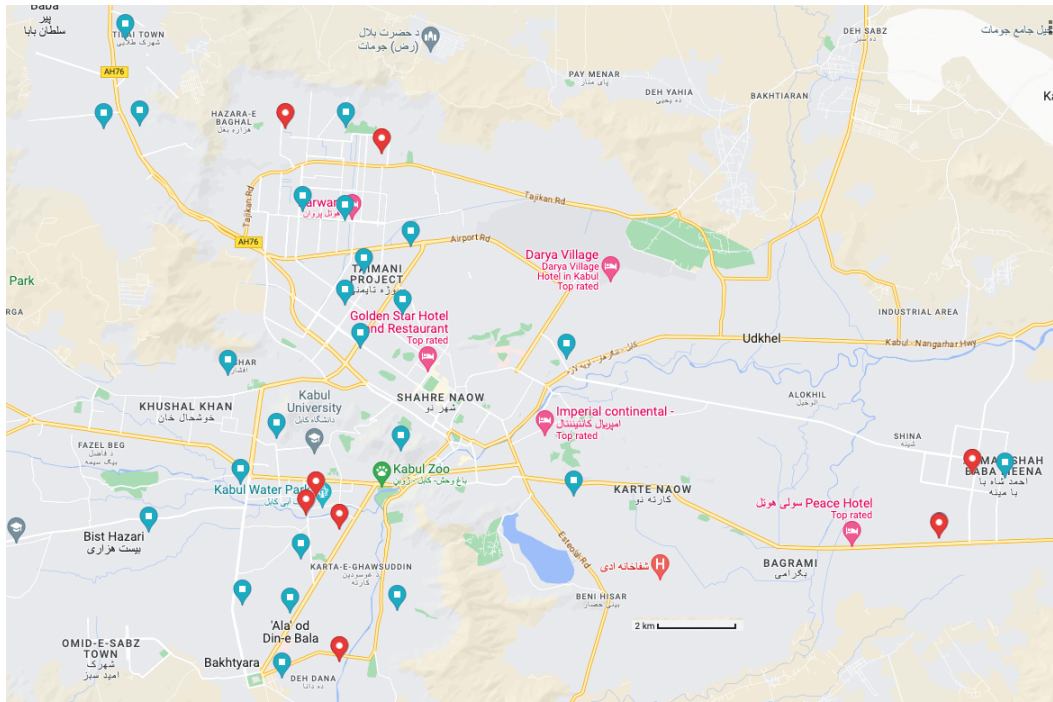


Fig. E4: Hypothetical Question on Preference Between Digital vs. Cash Aid

Notes: During the fourth survey round, we asked participants a hypothetical question to measure their willingness to pay to receive their aid payments in cash rather than digital. The question asked “We are hoping to use what we have learned from these surveys and from your experience with these payments to try to expand the program. While we do not have funding to do so at the moment, we are working to find it. In the future, we are also considering whether to give recipients the option to exchange the voucher for cash, rather than for goods at merchants. If we provide a cash out option, however, the fortnightly payments would be smaller because we have to pay a fee to make physical cash available.” and then proceeded by asking participants “If the fee was X AFN, would you prefer 4000 – X AFN in physical cash or 4000 AFN in HesabPay credit?”, where $X \in \{100, 300, 500\}$. The Figure presents the share of participants who preferred 4000 AFN in HesabPay credit over 3900 AFN (black bar), 3700 AFN (dark gray bar), or 3500 AFN (light gray bar) in cash, for the whole sample (first three bars), or divided by treatment group.



(a) Kabul



(b) Balkh



(c) Herat

Fig. E5: Location of Merchants

Notes: The map shows the location of merchants that participants visited at some point between the start of the program and December 31, 2022. Red pins are those that belong to merchants who participated in the onboarding sessions, while blue pins are those of other merchants. Note that not all merchants could be contacted/located. Coordinates have been jittered.

Table E1: Summary Statistics

| Variable | Mean (1) | SD (2) | N (3) |
|----------------------------------|-------------|-----------|----------|
| <i>Panel A. Outcome Vars</i> | | | |
| 1. Bread-tea breakfast | 6.67 | 0.89 | 2409 |
| 2. Bread-tea lunch | 3.58 | 1.34 | 2409 |
| 3. Bread-tea dinner | 3.51 | 1.49 | 2409 |
| 4. Total bread-tea meals | 13.76 | 2.56 | 2409 |
| 5. Afford Medicine | 0.01 | 0.09 | 2300 |
| 6. Income | 357.92 | 291.75 | 2409 |
| 7. Life satisfaction | 3.53 | 1.17 | 2409 |
| 8. Skipped meals | 2.62 | 1.52 | 2408 |
| 9. Happy | 0.01 | 0.1 | 2408 |
| 10. Employed | 0 | 0.03 | 2405 |
| <i>Panel B. Demographic Vars</i> | | | |
| 1. Married | 0.34 | 0.47 | 2409 |
| 2. Some education | 0.37 | 0.48 | 2404 |
| 3. Fin. decision maker | 0.66 | 0.47 | 2409 |
| 4. Age | 43 | 12.81 | 2409 |
| 5. Number family members | 6.31 | 2.05 | 2409 |
| 6. Has had bank account | 0 | 0.04 | 2409 |
| 7. Able to leave house | 0.65 | 0.48 | 2409 |
| 8. Has transf. airtime | 0 | 0.02 | 2409 |
| 9. Has transf. money | 0 | 0 | 2409 |
| 10. Credit constrained | 1 | 0.05 | 2409 |

Notes: Table shows the value of different variables at baseline, collected during the onboarding sessions, for the whole sample. Column 1 shows the mean of the variable, column 2 the standard deviation, and column 3 the number of respondents of this question at baseline.

Table E2: Attrition

| | Attrited (1) | By Survey Round (2) |
|--------------------------|-------------------|------------------------|
| Treated | -0.001 (0.004) | 0.004 (0.004) |
| Round 2 | | 0.003 (0.003) |
| Treated \times Round 2 | | -0.009** (0.004) |
| Observations | 4,818 | 4,818 |
| R Squared | 0.009 | 0.009 |

Notes: Control for strata fixed effects. Standard errors clustered at individual level. In the first survey round, 29 interviews could not be completed (17 treatment, 12 control). In the second round, 26 interviews could not be completed (10 treatment, 16 control). Households were surveyed once per month, over two months. Each of these months constitutes a survey round.
Levels of significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table E3: Are Digital Payments Diverted? Indices

| | KLK Index Others (1) | KLK Index You (2) | Anderson Index All (3) |
|---|----------------------------|-------------------------|------------------------------|
| Panel A. Baseline | | | |
| Treated | 0.034 (0.026) | 0.043 (0.031) | 0.048 (0.036) |
| Observations | 4,611 | 4,647 | 4,648 |
| Control Mean | 0.000 | 0.000 | 0.000 |
| Panel B. Long-Run | | | |
| β_1 : Treated \times Round 2 | 0.073 (0.052) | 0.025 (0.061) | 0.052 (0.071) |
| β_2 : Round 2 | -0.048* (0.029) | -0.097*** (0.036) | -0.135*** (0.046) |
| β_3 : Treated | -0.074 (0.084) | 0.006 (0.109) | -0.029 (0.125) |
| Observations | 4,611 | 4,647 | 4,648 |
| $(\beta_1 + \beta_2 + \beta_3) / \beta_3$ | 0.670 | -11.440 | 3.920 |
| p-value $\beta_1 + \beta_3 = 0$ | 0.964 | 0.556 | 0.701 |

Notes: In Panels A and B, control for stratification fixed effects, survey round fixed effects, and baseline value of dependent variable, if available. Columns 1-2 create indices following Kling, Liebman & Katz (2007). Column 1 uses the two measures on whether others in their community have been asked to provide informal assistance, while column 2 uses the two measures on whether participants themselves have been asked to provide informal assistance by political actors. Column 3 includes the four measures, but creates the index following Anderson (2008). This analysis was not pre-specified. Includes control for surveyor fixed effects. Standard errors clustered at individual level.

Table E4: Treatment Effects by Survey Round – Secondary Outcomes

| | (1) | (2) | (3) | (4) | (5) |
|---|------------------------|----------------------|---------------------|----------------------|----------------------|
| Panel A. Types of Food | | | | | |
| β_1 : Treated \times Round 2 | 0.367*** (0.060) | 0.240*** (0.049) | 0.014 (0.071) | 0.001 (0.011) | 0.032 (0.023) |
| β_2 : Round 2 | -0.051 (0.040) | -0.124*** (0.030) | 0.165*** (0.051) | -0.008 (0.006) | -0.065*** (0.015) |
| β_3 : Treated | 0.412*** (0.049) | 0.373*** (0.041) | -0.010 (0.056) | 0.012 (0.009) | 0.031 (0.021) |
| Control Mean | 0.698 | 0.518 | 1.394 | 0.021 | 0.074 |
| Observations | 4,763 | 4,763 | 4,763 | 4,763 | 4,763 |
| R Squared | 0.083 | 0.080 | 0.010 | 0.011 | 0.012 |
| p-value $\beta_1 + \beta_3 = 0$ | 0 | 0 | 0.942 | 0.046 | 0 |
| Panel B. Other Economic Outcomes | | | | | |
| β_1 : Treated \times Round 2 | 36.807 (151.022) | 0.050** (0.023) | 0.013 (0.022) | 0.008 (0.019) | |
| β_2 : Round 2 | -165.446** (64.257) | 0.032* (0.017) | 0.030** (0.015) | -0.046*** (0.012) | |
| β_3 : Treated | 110.705 (149.712) | -0.033* (0.018) | 0.010 (0.019) | 0.031** (0.015) | |
| Control Mean | 876.683 | 0.192 | 0.666 | 0.051 | |
| Observations | 4,763 | 4,741 | 4,757 | 3,582 | |
| R Squared | 0.010 | 0.021 | 0.061 | 0.011 | |
| p-value $\beta_1 + \beta_3 = 0$ | 0.048 | 0.361 | 0.202 | 0.002 | |

Notes: This table reports estimated impacts of treatment separately for the first and second survey round. Households were surveyed once per month for two months. Each of these months constitutes a survey round. All specifications control for stratum fixed effects and the baseline value of the dependent variable, if available. Standard errors are clustered at individual level.

Levels of significance: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table E5: Summary Table – Experimenter Demand Effects

| Variable | Experimenter Demand | | | | |
|---------------------------------------|---------------------|--------------------------|-------------------------|-------------------------|---------------------------|
| | Control Mean (1) | Baseline Estimate (2) | Overall Estimate (3) | Control Estimate (4) | Treatment Estimate (5) |
| Panel A. Food Security | | | | | |
| Days skipping meals (past week) | 2.615 | -0.76*** (0.051) | -0.027 (0.057) | -0.099 (0.079) | 0.039 (0.068) |
| Children skipping meals (=1) | 0.873 | -0.117*** (0.012) | 0.016 (0.015) | 0.027 (0.018) | 0.005 (0.023) |
| Regularly eat twice a day | 0.501 | 0.093*** (0.015) | -0.005 (0.013) | -0.006 (0.018) | -0.003 (0.019) |
| Total bread and tea meals (past week) | 13.639 | -1.608*** (0.121) | -0.181 (0.146) | -0.107 (0.199) | -0.268 (0.19) |
| <i>Food Security - KLK Index</i> | 0 | 0.501*** (0.032) | 0.002 (0.036) | -0.002 (0.045) | 0.01 (0.047) |
| Panel B. Informal Taxation | | | | | |
| Inf. tax gov. off. (others) | 0.004 | 0.001 (0.001) | -0.001 (0.002) | -0.002 (0.002) | 0 (0.003) |
| Inf. tax comm. leader (others) | 0.001 | 0.002 (0.002) | -0.002 (0.002) | 0.002 (0.002) | -0.006 (0.005) |
| Inf. tax gov. off. (you) | 0 | 0.002** (0.001) | -0.001 (0.001) | 0 (0) | -0.002 (0.003) |
| Inf. tax comm. leader (you) | 0.003 | 0.002 (0.003) | -0.005* (0.003) | -0.005* (0.003) | -0.005 (0.004) |
| <i>Informal tax. - KLK Index</i> | 0 | 0.074* (0.042) | -0.077 (0.054) | -0.027 (0.036) | -0.124 (0.106) |
| Panel C. Wellbeing | | | | | |
| Better economic situation | 0.048 | 0.335*** (0.011) | 0.003 (0.017) | 0.018 (0.014) | -0.009 (0.025) |
| Satisfied with fin. situation | 0.133 | 0.263*** (0.012) | 0.003 (0.017) | 0.013 (0.018) | -0.005 (0.024) |
| Happy | 0.009 | 0.28*** (0.014) | -0.007 (0.016) | -0.007 (0.018) | -0.004 (0.021) |
| Life satisfaction (std) | 0.011 | 1.682*** (0.058) | -0.007 (0.079) | 0.073 (0.085) | -0.075 (0.103) |
| <i>Economic/Wellbeing - KLK Index</i> | 0 | 1.498*** (0.042) | 0.001 (0.063) | 0.054 (0.055) | -0.04 (0.081) |

Notes: Control for stratification fixed effects, survey round fixed effects (baseline estimates), and baseline value of dependent variable, if available. Standard errors clustered at individual level in parenthesis. Control mean is the mean in the baseline if available, or across follow up rounds otherwise, for the control group. The KLK Index is created following Katz, Kling, & Liebman (2007), and is the equally-weighted sum of the standardised component variables. The baseline effect is the (pooled) ITT effect of the main treatment (receiving the aid payments). The overall effect is the coefficient on the prime treatment. The control effect is the coefficient on the prime treatment in a regression where the prime treatment and the main treatment are interacted, while the effect on the treated is the sum of the prime treatment and the interaction term between the two treatments from the same regression. Better economic situation is an index that equals 1 if the respondent answered that her economic situation compared to 30 days ago is slightly or much better, and 0 otherwise. Satisfied with financial situation is a dummy that equals 1 if the respondent answered that she agrees a lot or somewhat with the statement that she is highly satisfied with her current financial condition, and 0 otherwise. Happy is a dummy that equals 1 if respondent said that she was very happy or quite happy, and 0 otherwise. Life satisfaction is the score from 1 (dissatisfied) to 10 (satisfied) in terms of how satisfied the respondent is with her life as a whole these days (standardised). Total household income excludes the aid payments.

Table E6: Summary Table – Treatment Effects, Restricted Sample

| | Control Mean (1) | Control SD (2) | Treatment Effect (3) | Standard Error (4) | Naive <i>p</i> -value (5) | Adjusted <i>p</i> -value (6) | N (7) |
|--|------------------------|----------------------|----------------------------|--------------------------|---------------------------------|------------------------------------|----------|
| Panel A. Primary Outcomes | | | | | | | |
| Days skipping meals (past week) | 2.579 | 1.813 | -0.785 | 0.053 | 0 | 0.0002 | 4412 |
| Children skipping meals (=1) | 0.87 | 0.336 | -0.113 | 0.012 | 0 | 0.0002 | 4412 |
| Regularly eat twice a day | 0.493 | 0.5 | 0.1 | 0.015 | 0 | 0.0002 | 4412 |
| Total bread and tea meals (past week) | 13.644 | 3.921 | -1.628 | 0.125 | 0 | 0.0002 | 4412 |
| <i>Food Security - KLK Index</i> | -0.016 | 1.01 | 0.51 | 0.033 | 0 | | 4412 |
| <i>Food Security - Anderson Index</i> | 0.002 | 1.058 | 0.493 | 0.034 | 0 | | 4412 |
| Inf. tax gov. off. (others) | 0.002 | 0.043 | 0.001 | 0.001 | 0.563 | 0.6724 | 4412 |
| Inf. tax comm. leader (others) | 0.002 | 0.043 | 0.003 | 0.002 | 0.118 | 0.125 | 4412 |
| Inf. tax gov. off. (you) | 0.001 | 0.03 | 0.002 | 0.001 | 0.163 | 0.1512 | 4412 |
| Inf. tax comm. leader (you) | 0.005 | 0.071 | 0.001 | 0.002 | 0.612 | 0.6724 | 4412 |
| <i>Informal tax. - KLK Index</i> | -0.007 | 0.923 | 0.075 | 0.041 | 0.068 | | 4412 |
| <i>Informal tax. - Anderson Index</i> | -0.016 | 0.763 | 0.048 | 0.031 | 0.116 | | 4412 |
| Better economic situation | 0.047 | 0.212 | 0.34 | 0.011 | 0 | 0.0002 | 4412 |
| Satisfied with fin. situation | 0.133 | 0.339 | 0.264 | 0.013 | 0 | 0.0002 | 4412 |
| Happy | 0.235 | 0.424 | 0.275 | 0.014 | 0 | 0.0002 | 4412 |
| Life satisfaction | 2.969 | 1.888 | 1.999 | 0.07 | 0 | 0.0002 | 4412 |
| <i>Economic/Wellbeing - KLK Index</i> | 0.001 | 1 | 1.509 | 0.044 | 0 | | 4412 |
| <i>Economic/Wellbeing - Anderson Index</i> | 0.096 | 1.167 | 1.326 | 0.045 | 0 | | 4412 |
| Panel B. Secondary Outcomes | | | | | | | |
| Days eating rice (past week) | 0.672 | 1.05 | 0.615 | 0.036 | 0 | 0.001 | 4412 |
| Days eating beans (past week) | 0.517 | 0.841 | 0.491 | 0.03 | 0 | 0.001 | 4412 |
| Days eating vegetables (past week) | 1.377 | 1.331 | 0 | 0.042 | 0.995 | 0.42 | 4412 |
| Days eating chicken (past week) | 0.018 | 0.138 | 0.014 | 0.006 | 0.021 | 0.022 | 4412 |
| Days eating dairy (past week) | 0.073 | 0.387 | 0.05 | 0.014 | 0 | 0.001 | 4412 |
| Able to buy medicine | 0.078 | 0.269 | 0.031 | 0.01 | 0.003 | 0.005 | 3344 |
| Involved in fin. decisions | 0.672 | 0.47 | 0.019 | 0.016 | 0.23 | 0.152 | 4412 |
| Total household income (past month) | 1093.634 | 1792.27 | 156.559 | 94.73 | 0.099 | 0.071 | 4412 |
| Household's head employed (past month) | 0.284 | 0.451 | -0.008 | 0.014 | 0.579 | 0.278 | 4412 |

Notes: Control for stratification fixed effects, survey round fixed effects, and baseline value of dependent variable, if available. Standard errors clustered at individual level. Primary outcomes show FWER-adjusted *p*-values within each family outcome (following Romano & Wolf, 2005, using 5000 repetitions), while secondary outcomes show FDR-adjusted *p*-values (following Anderson, 2008). The KLK Index is created following Katz, Kling, & Liebman (2007), and is the equally-weighted sum of the standardised component variables. The Anderson Index is created following Anderson (2008), and weights the component variables by the inverse of their variance-covariance matrix. Better economic situation is an index that equals 1 if the respondent answered that her economic situation compared to 30 days ago is slightly or much better, and 0 otherwise. Satisfied with financial situation is a dummy that equals 1 if the respondent answered that she agrees a lot or somewhat with the statement that she is highly satisfied with her current financial condition, and 0 otherwise. Happy is a dummy that equals 1 if respondent said that she was very happy or quite happy, and 0 otherwise. Life satisfaction is the score from 1 (dissatisfied) to 10 (satisfied) in terms of how satisfied the respondent is with her life as a whole these days. Total household income excludes the aid payments. In Round 1, there are 207 women who did not respond to every single question (out of 2381 respondents), while in Round 2 there are 145 (out of 2383 respondents).

Table E7: Number of Merchants by Nahia

| City | Nahia | # Participants (1) | # Onboarding Merchants (2) | # All Merchants (3) |
|-------|-------|-----------------------|-------------------------------|------------------------|
| Kabul | 6 | 82 | 3 | 22 |
| Kabul | 7 | 83 | 2 | 19 |
| Kabul | 8 | 51 | 1 | 8 |
| Kabul | 11 | 137 | 1 | 23 |
| Kabul | 12 | 99 | 1 | 10 |
| Kabul | 13 | 152 | 3 | 25 |
| Kabul | 15 | 85 | 1 | 26 |
| Kabul | 17 | 98 | 1 | 15 |
| Kabul | 22 | 30 | 1 | 8 |
| Herat | 1 | 100 | 3 | 4 |
| Herat | 11 | 100 | 3 | 5 |
| Herat | 13 | 597 | 5 | 7 |
| Balkh | 4 | 200 | 1 | 8 |
| Balkh | 6 | 196 | 1 | 6 |
| Balkh | 7 | 200 | 3 | 4 |
| Balkh | 8 | 199 | 2 | 6 |

Notes: The number of own test merchants reflects the merchants that participants visited during the onboarding session to conduct their test purchase. The number of all merchants includes all numbers that we can identify as merchants in the transaction data.

Table E8: Existence of Necessary Conditions in Fragile Settings

| Country | Food insecurity | | | | CDC | Mobile phone | Mobile money |
|-------------|-----------------|---------------------|----------|-------------|---------|--------------|--------------|
| | Millions | Share of population | Conflict | Freedom | | | |
| DRC | 26.4 | 26% | Yes | Not free | Yes | 45.50 | Medium |
| Ethiopia | 23.6 | 21% | Yes | Not free | Yes | 49.44 | Medium |
| Afghanistan | 19.9 | 46% | Yes | Not free | Yes | 58.26 | Very low |
| Nigeria | 19.5 | 12% | Yes | Partly free | Yes | 99.07 | Medium |
| Yemen | 17.3 | 55% | Yes | Not free | Yes | 60.49 | No data |
| Myanmar | 15.2 | 27% | Yes | Not free | Yes | 95.36 | Medium |
| Syria | 12.1 | 55% | Yes | Not free | No data | 95.20 | No data |
| Sudan | 11.7 | 24% | Yes | Not free | Yes | 80.26 | Very low |
| Ukraine | 8.9 | 25% | Yes | Partly free | No data | 100.0 | No data |
| Pakistan | 8.6 | 43% | Yes | Partly free | Yes | 79.51 | Medium |

Notes: Food insecurity [2], with the share of population being the share of the population *analyzed* by FSIN. Active armed conflict [153]. Level of freedom [29]. Community development council (CDC) presence coded from various sources (See footnotes). Mobile-cellular subscriptions per 100 people [154]. Mobile money prevalence index [155].

Table E9: Actual Values vs. Experts' Predictions

| Variable | Actual Values (1) | Predicted Values (2) | p-value Predicted = Actual (3) |
|--|----------------------|-------------------------|-----------------------------------|
| How Many Bread and Tea Meals in Past Week? | 11.96 | 10.32 (0.45) | < 0.001 |
| Share Reporting Diversion Attempts | 1.99 | 39.85 (2.98) | < 0.001 |
| Share Able to Use Digital Payments | 99.75 | 43.82 (2.75) | < 0.001 |
| Delivery Costs | 6.7 | 10.65 (0.74) | < 0.001 |

Notes: The first column shows the actual values of different elements of the intervention. The second column shows the mean predicted value by our sample of experts (with standard deviations in parentheses). The third column shows the p-value of a test of the mean predicted value being equal to the actual value (i.e. a test of equality of columns 1 and 2). Results are similar when looking at predictions by academics and practitioners separately and are available on request.

Table E10: Experts' Survey Questionnaire

| Statement | Answer Options |
|--|---------------------------------|
| <p>THIS SURVEY CONSISTS OF 5 QUESTIONS THAT SHOULD TAKE LESS THAN 3 MINUTES TO COMPLETE. In conflict settings, distributing humanitarian aid while supporting the dignity and inclusion of vulnerable populations is challenging, in particular where there is a risk of capture by hostile regimes. The widespread adoption of mobile phones suggests one innovative solution: direct aid via digital financial platforms. Digital payments have the potential to empower recipients to meet basic needs using local markets, and to improve transparency while minimizing opportunities for diversions compared to physical cash or commodity distribution. We are piloting a direct aid program to Afghan women using a commercial platform called HesabPay. Local partners have identified 2500 highly vulnerable women in three major cities (Kabul, Herat, and Mazar), each of whom will receive four semi-monthly digital payments of 50 USD. All beneficiaries complete an in-person onboarding that includes identity verification, registration of a digital wallet linked to a unique mobile phone number, and a test transaction using the digital wallet to purchase goods from a registered merchant. Aid payments are unconditional and can be used for purchases at any HesabPay-registered local merchant. To inform our assessment of this program, we are collecting the views of experts like yourself about the likelihood of operational success and anticipated impacts. Your responses will be anonymized. If you opt in, we will contact you again after the study is complete to provide the final research findings as well as the aggregate views of those who completed this survey. Based on the short program description above, please respond to the following questions – please select “Not Applicable” if you are not able to provide a forecast for a given question.</p> | |
| <p>Please check this box to consent to participate in this survey. Only the researchers at the London School of Economics and Political Science will have access to your personal information.</p> | <p>Yes</p> |
| <p>Q1) What percentage of beneficiaries do you expect will be able to use their digital payments to purchase goods directly without resorting to cashing out the aid?</p> | <p>0-100%</p> |
| <p>Q2) How many meals per week do you expect beneficiaries of direct aid to eat only bread and tea? (At baseline, they reported 14 meals of only bread and tea out of 21 meals in the last week.)</p> | <p>0-21 Bread and Tea Meals</p> |
| <p>Q3) What share of beneficiaries do you expect will report efforts to tax or divert their payments?</p> | <p>0-100%</p> |
| <p>Q4) What do you expect the delivery cost to be for our digital direct aid payments - not including costs of beneficiary selection or impact evaluation? (By one recent estimate, the delivery cost of humanitarian aid distributed in physical cash in Afghanistan is approximately 17%.)</p> | <p>0-100 cents</p> |

Table E11: Cost-efficiency calculations

| Panel A: Raw costs | |
|--|--------------|
| <i>Costs by category</i> | |
| –Distribution: Personnel | \$4054.18 |
| –Distribution: Facilities | \$914.36 |
| –Distribution: Technology | \$910.62 |
| –Digital aid transfers | \$433,620.00 |
| <i>Total costs without onboarding</i> | \$439,499.16 |
| –Onboarding costs | \$23,029 |
| <i>Total costs with onboarding</i> | \$462,528.16 |
| <i>Total costs treatment group (N = 1,208)</i> | \$231,936 |
| <i>Total costs control group (N = 1,201)</i> | \$230,592 |
| Panel B: Costs per beneficiary (CPB) | |
| <i>Including the digital aid transfer</i> | |
| –With onboarding | \$192.00 |
| –Without onboarding | \$182.44 |
| <i>Excluding the digital aid transfer</i> | |
| –With onboarding | \$12 |
| –Without onboarding | \$2.44 |
| Panel C: Transfer ratios | |
| <i>Total cost-transfer ratios (TCTR)</i> | |
| –With onboarding | 1.067 |
| –Without onboarding | 1.014 |
| <i>Cost-transfer ratios (CTR)</i> | |
| –Cost-transfer ratio with onboarding | 6.7 cents |
| –Cost-transfer ratio without onboarding | 1.4 cents |

Note: Based on 2,409 beneficiary households each receiving 180 USD for a total of 433,620 USD distributed.

Table E12: : Cost-effectiveness ratios: days with skipped meals and corresponding calorie intake

| CERs for Reduction in Days with Skipped Meals | | | | |
|--|------------------|--------------------------|-------------------|------------|
| <i>Group</i> | <i>Mean/Week</i> | <i>Skipped Days</i> | <i>% Skipped</i> | <i>CER</i> |
| <i>Panel A: The entire eight-week program (Weeks 1–8)</i> | | | | |
| Late Group | 2.569 | 24,682.952 | 36.700% | |
| Early Group | 1.809 | 17,482.176 | 25.843% | |
| Effect | -0.760 | -7,344.640 | -10.857% | \$31.58 |
| <i>Panel B: The second half of the program (Weeks 5–8)</i> | | | | |
| Late Group | 2.448 | 11,760.670 | 34.973% | |
| Early Group | 1.401 | 6767.699 | 20.009% | |
| Effect | -1.048 | -5,061.520 | -14.964% | \$22.91 |
| <i>Panel C: The final week of the program (Week 8)</i> | | | | |
| Late Group | 2.508 | 3,012.108 | 35.823% | |
| Early Group | 1.088 | 1,314.304 | 15.543% | |
| Effect | -1.42 | -1,715.36 | -20.286% | \$16.90 |
| CERs for Reduction in Foregone Calories | | | | |
| <i>Group</i> | <i>Mean/Week</i> | <i>Foregone calories</i> | <i>% Foregone</i> | <i>CER</i> |
| <i>Panel D: The full eight-week program (Weeks 1–8)</i> | | | | |
| Late Group | 34,041.82 | 327,073,797 | 36.700% | |
| Early Group | 23,971.06 | 231,656,314 | 25.843% | |
| Effect | -10,070.76 | -97,323,825 | -10.857% | \$2.38 |
| <i>Panel E: The second half of the program (Weeks 5–8)</i> | | | | |
| Late Group | 32,439.77 | 155,840,670 | 34.973% | |
| Early Group | 18,559.35 | 89,678,782 | 20.009% | |
| Effect | -13,880.42 | -67,070,202 | -14.964% | \$1.73 |
| <i>Panel F: The final week of the program (Week 8)</i> | | | | |
| Late Group | 33,233.51 | 39,913,443 | 35.829% | |
| Early Group | 14,417.088 | 17,415,842 | 15.543% | |
| Effect | -18,816.42 | -22,730,235 | -20.286% | \$1.28 |

Notes: The cost-effectiveness ratio is computed as the Early Group's costs divided by the reduction in days with skipped meals for the beneficiaries (treatment group costs / (# beneficiaries * treatment effect * # of weeks)) and is interpreted as the number of dollars required for a reduction of a day with skipped meals. For aggregate estimates, there are 67,648 and 67,256 possible days for the Early Group and Late Group respectively (1208 or 1201 participants * 56 days). The cost-effectiveness ratio for calorie intake is computed similarly and is interpreted as the number of dollars required for a reduction of 1,000 foregone calories. Results for calorie intake are based on the Afghanistan CVWG estimate of 2,100 calories/person/day, which translates to 13,251 calories/household/day and 92,757 calories/household/week. Across the eight-week program, there are 896,403,648 and 891,209,256 possible calories for the Early Group and Late Group respectively (1208 or 1201 participants * 92,757 calories per household week * 8 weeks).

Table E13: : Cash-based transfer programs with rigorous evaluation in crisis contexts

| Country/ Study | Target Beneficiaries | Cash Type | Delivery Mechanism | Nutrition Effect | Well-Being Effect | Diversion Discussed | Usage (Mobile) | Decent- ralized | Trans- parent | Cost (TCTR) |
|--------------------------------|-------------------------|--------------|-----------------------|---------------------|----------------------|------------------------|-------------------|--------------------|------------------|----------------|
| 1. Ecuador [156] | Refugees | C/V/F | ATM/physical | + | | | | ✓ | | 1.075 |
| 2. Philippines [102] | Affected HHs | C | Mobile money | ∅ | | | ✓ | | | 1.020 |
| 3. Philippines [157] and [158] | Poor HHs | C | Physical | + | | | | | | |
| 4. Lebanon [159] and [160] | Refugees | C | ATM | + | + | | | ✓ | | |
| 5. Lebanon [43] | Refugees | C | ATM | + | ∅ | ✓ | | ✓ | | |
| 6. Lebanon [144] | Refugees | C | Not stated | | | | | | | |
| 7. Sri Lanka [161] | Affected HHs | C/F | Physical | + | | | | | | ✓ |
| 8. Bangladesh [97] | At-risk HHs | C | Mobile money | + | + | | ✓ | ✓ | ✓ | |
| 9. Uganda [41] | Poor women | C | Physical | + | | ✓ | | | | ✓ |
| 10. Niger [63] | Poor women | C | Mobile/physical | + | | | ✓ | | ✓ | 1.059 |
| 11. Niger [162] | Poor HHs | C/F | ATM | + | | | | | | 1.053 |
| 12. Niger [146] and [163] | Poor women | C | Physical | + | + | | | | | 1.252 |
| 13. Niger [164] | Poor children | C/F | Physical | + | | | | | | |
| 14. Iraq [129] | Poor HHs | C | Physical | + | | | | | | |
| 15. D.R.C. [42] | Displaced | C/V | Bank/physical | ∅ | | ✓ | | | | 1.087 |
| 16. D.R.C. [165] | Displaced | C/V | Multiple | ∅ | | | | | | |
| 17. D.R.C. [101] | Displaced | V | Physical | ∅ | + | | | | | 1.437 |
| 18. C.A.R. [99] | Poor HHs | C | Physical | ∅ | + | | | | | |
| 19. Yemen [166] and [167] | Poor HHs | C/F | Physical | + | | | | | | 1.083 |
| 20. Yemen [168] | Poor women | C | Physical | + | | | | | | |
| 21. Afghanistan [98] | Poor women | C | Physical | + | + | | | | | 2.017 |
| 22. Afghanistan [151] | At-risk youth | C | Mobile money | | | | ✓ | ✓ | ✓ | ✓ |
| 23. South Sudan [100] | Poor HHs | C | Physical | + | | | | | | |

Notes: Abbreviations: HHs=households; C=Cash; V=Voucher; F=Food. Symbols apply to cash/voucher findings/discussion: +=analytically examined and positive effect; -=analytically examined and negative effect; ∅=analytically examined and null effect; ✓=discussed and/or a possible advantage, but either not analyzed or, as with the cost metric, insufficient information to derive the TCTR; Organization: countries are ordered by increasing level of fragility based on OECD [143]. Ecuador and Philippines are not ranked or discussed; Lebanon and Sri Lanka are not ranked, but discussed as exhibiting early-warning signs.

References: 1. Hidrobo et al. [156]. 2. MercyCorps [102]. 3. Kandpal et al. [157] and Crost et al. [158]. 4. Salti et al. [159] and Moussa et al. [160]. 5. Lehmann and Masterson [43]. 6. de Hoop et al. [144]. 7. Sandström and Tchatchua [161]. 8. Pople et al. [97]. 9. Blattman et al. [41]. 10. Aker et al. [63]. 11. Hoddinott et al. [162]. 12. Bossuroy et al. [146]. 13. Langendorf et al. [164]. 14. Kurtz et al. [129]. 15. Aker [42]. 16. Bonilla et al. [165]. 17. Quattrochi et al. [101]. 18. Alik-Lagrange et al. [99]. 19. Schwab [166]. 20. Kurdi [168]. 21. Bedoya et al. [98]. 22. Lyall et al. [151]. 23. Chowdhury et al. [100].

Table E14: Related nutrition, psychological well-being, and diversion results

| Program | Relevant study context & notes | Selected results |
|---|--|---|
| 1. 2011 WFP food security & nutrition for refugee integration (Ecuador, [156]). | RCT evaluation of 240 USD cash / voucher / food / control to 2,357 HHs to address non-acute, persistent displacement (by government request). | Log calorie intake per capita increased by 12% (cash), 18% (vouchers) 21% (food); food consumption scores (FCS) improved by 11% (cash), 16% (vouchers), and 12% (food). |
| 2. 2014 Mercy Corps TabangKO for typhoon recovery (Philippines, [102]). | RCT evaluation of 89 USD lump-sum / varying disbursement cash to 25,480 HHs to assist immediate recovery & build resilience (with government targeting assistance). | None of the treatment arms had differential impacts on nutrition/food outcomes (coping strategies index, FCS dietary diversity). |
| 3. 2009–11 Pantawid Pamilya poverty reduction program (Philippines, [157] and [158]). | RCT evaluation of 176–330 USD cash / control for longer-term poverty reduction in 65 non-acute, high-need villages in 714 study HHs (government implemented). | Dairy and egg consumption increased by 6.9pp and 8.2pp, and broader increases in height-for-age, increased antenatal & prenatal care, decreased severe stunting, & less violence. |
| 4. 2017–18 WFP/UNHCR multi-purpose cash assistance (Lebanon, [159] and [160]). | RD assessment of 2100 USD cash (shorter/longer) / control to 56,000 refugee HHs (6,287 in study) to meet survival needs (contributing to government crisis response). | Dairy consumption +0.6 days (control: 2.52), borrow food -0.27 days (control: 1.74), eating elsewhere +0.1 days (control: 0.17); happiness increased, stress decreased. |
| 5. 2013–14 UNHCR/IRC Winter Cash Assistance Program (Lebanon, [43]). | RD assessment of 575 USD cash / control to 87,700 Syrian refugee HHs (1,358 in study) to provide winter shelter (with government targeting assistance). | Days w/: meals skipped -0.65 (control: 3.25), less-preferred meals -0.6 days (control: 4.7), restricted consumption -0.3 (control: 2.6), & reduced portions -0.4 (control: 3.2). |
| 6. 2016–17 No Lost Generation education program (Lebanon, [144]). | RD assessment of 441 USD cash / control to 1440 Syrian refugee HHs to offset commuting / foregone income (part of government education initiative). | No nutrition, psychological well-being, or diversion outcomes reported. |
| 7. 2005–6 WFP cash transfers for tsunami relief (Sri Lanka, [161]). | RCT evaluation of cash / food (amount unclear) to 3,200 HHs (1,360 in study) to address medium-term recovery from tsunami a year earlier (government role unclear). | Estimates not reported: food spending increased; dietary diversity improved; & consumption of staples decreased. Greatest impact in poorest areas, with neediest HHs. |
| 8. 2020 WFP anticipatory cash for flood preparation & relief (Bangladesh, [97]). | Natural experiment of 53 USD cash / control to 23,434 HHs (8,954 in study) ahead of imminent flooding (with government forecasting/targeting assistance). | Percentage of HHs reporting: days w/out meals -10pp (control: 28%) and children eating 3 meals +3pp (control: 80%); Increased life satisfaction +12.5% (control: 2.8) on a 10-pt scale. |
| 9. 2009–11 AVSI Women’s Income Generating Support (Uganda, [41]). | RCT evaluation of 150 USD cash, training, & supervision / control to 1,800 women to promote microenterprise in post-conflict region (supporting government programming). | Number of times in past week going hungry -0.1 (control: 0.2), usual number of meals per day +0.06 (control: 1.76), and diversion less than 1% of grants. |
| 10. 2009–10 Concern Worldwide social protection program (Niger, [63]). | RCT evaluation of 225 USD cash (mobile vs. physical vs. physical+phone) to 10,000 women (1152 in study survey) for drought response (government role unclear). | Mobile increased dietary diversity +0.28–0.51 (physical: 3.17; physical+phone: 2.94) on 12-pt scale and # of meals by under-5 children in past day +0.33 (physical: 3.17). |
| 11. 2011 WFP cash/food aid in Zinder region (Niger, [162]). | RCT evaluation of 50 USD cash-for-work / food-for-work to 2,209 poor HHs to address lean-season drought among highly poor (by government request). | Dietary diversity reduced -0.56 (food mean: unclear) on 25-pt scale, food consumption reduced -3.9 points (food: 41.5) on 112-pt scale, & increased grain purchases. |
| 12. 2016–19 Supplement to government safety nets program (Niger, [146]). | RCT evaluation of 127 USD cash / psychosocial / joint (w/core support) to 4712 poor women (1191 in cash) in HHs in government safety nets program. | Cash increased food security +0.20sd (joint: +0.25sd), dietary diversity +3.69 (joint +6.11), days eating foods from 8 groups, & life satisfaction +0.20sd (joint +0.45sd). |

| | | |
|---|---|---|
| 13. 2011 Cash for prevention of young child malnutrition (Niger, [164]). | RCT evaluation of 260 USD cash / cash+various food supplements to 5,395 HHs with at-risk 6–23 month old children (in collaboration w/government ministries). | Cash+food reduced moderate acute malnutrition by double (Hazard ratios: 2.07–2.42). Cash+super cereal reduced severe acute malnutrition by triple (Hazard ratio: 3.13). |
| 14. 2019–20 multi-purpose cash assistance for economic recovery (Iraq, [129]). | RCT evaluation of 1200 USD cash (lump-sum, multiple disbursements) / control to 827 poor HHs in protracted crisis (supporting the government safety nets program). | Food consumption +4.515 (control: appx. 44) on 112-pt scale, reliance on coping strategies -12.18 (control: 25.01), and insecurity -0.18 (control: unclear) on 1–5 scale. |
| 15. 2011 Concern Worldwide social protection program (D.R.C., [42]). | RCT evaluation of 130 USD cash / vouchers to 474 displaced HHs to increase access to food & non-food items amid ongoing violent conflict (government role unclear). | No improvements to: diet diversity (3.05 on 12-pt scale; voucher: 3.07), meals / day (1.35; voucher: 1.38), or months of adequate food (1.25; voucher: 1.26). |
| 16. 2013–15 Multi-purpose cash for displaced communities in crisis (D.R.C., [165]). | 2 RCT evaluations: 120 USD cash (lump sum vs. 3 disbursements) to 196 HHs & gender of HH beneficiary (male, female, choice) for 157 HHs (government role unclear). | No differential impacts of number of transfers or gender of HH beneficiary on food security, children’s health, or other key program indicators. |
| 17. 2017–18 RRPM voucher assistance for displaced (D.R.C., [101]). | RCT evaluation of 55–90 USD vouchers / control to 488 displaced HHs (488 control) to improve health & well being (government role unclear). | Psych. well-being +0.32sd at 6-weeks, +0.18 at 1-year; life satisfaction +0.59 (control: 3.29) on 10-pt scale, & coping/meal skipping -0.07 (control: 1.79) on 11-question index. |
| 18. 2016–19 WB Londo cash-for-public works program (C.A.R., [99]). | Natural experiment of 120 USD cash-for-public-works / control to 3,470 poor study HHs to improve livelihoods amid ongoing violent conflict (government role unclear). | No change in days with skipped meals (0.30; control: 0.31), happiness increases +0.14sd (control: -0.10), and satisfaction with security increases 0.08sd (control: -0.07). |
| 19. 2011–12 WFP cash & food transfer program (Yemen, [167] and [166]). | RCT evaluation of 147 USD cash / food to 1,983 poor HHs facing lean-season food insecurity (in collaboration with government). | Food consumption +2.27 (food: 48.82) on 112-pt scale, days/week reducing meal frequency +0.37 (food: 0.63), and days/week adults reduced intake +0.47 (food: 0.3). |
| 20. 2015–17 Yemen SWF / WB Cash-for-Nutrition program (Yemen, Kurdi [168]). | RCT evaluation of 1122 USD cash / control to 1,001 women in high-need HHs facing food insecurity amid violent conflict (in collaboration with government). | Child dietary diversity +0.616 (control: 1.88) of 7 food groups; likelihood of coping through selling gold -0.104 (control: 0.244) among the highest tercile. |
| 21. 2015–18 Targeting the Ultra Poor program (Afghanistan, [98]). | RCT evaluation of 180 USD cash, livestock assets, training, & coaching / control to 491 ultra-poor female-headed HHs (implemented by government-owned organization). | HHs in which adults skip/reduce meals -23pp (control: 0.44) & children skip/reduce -20pp (control: 0.59); life satisfaction +0.44 (control: 0.00) on 10pt-scale. |
| 22. 2016 Mercy Corps INVEST vocation/skills program (Afghanistan, [151]). | RCT evaluation of 75 USD cash / training / combined to 1,841 youth (756 control) at risk of insurgent recruitment (supported by local government). | No nutrition, psychological well-being, or diversion outcomes reported. |
| 23. 2013–14 Cash / Targeting the Ultra Poor (South Sudan, [100]). | RCT evaluation of 350–410 cash / TUP bundle / control to 125 poor HHs to increase productivity amid ongoing violent conflict (government role unclear). | Null: % skip day of meals / week (0.06; control: 0.41), go to bed hungry (0.03; control: 0.40), no food in home (-0.01; control: 0.45), & limiting portions (-0.04; control: 0.48). |

References: 1. Hidrobo et al. [156]. 2. MercyCorps [102]. 3. Kandpal et al. [157] and Crost et al. [158]. 4. Salti et al. [159] and Moussa et al. [160]. 5. Lehmann and Masterson [43]. 6. de Hoop et al. [144]. 7. Sandström and Tchatchua [161]. 8. Pople et al. [97]. 9. Blattman et al. [41]. 10. Aker et al. [63]. 11. Hoddinott et al. [162]. 12. Bossuroy et al. [146]. 13. Langendorf et al. [164]. 14. Kurtz et al. [129]. 15. Aker [42]. 16. Bonilla et al. [165]. 17. Quattrochi et al. [101]. 18. Alik-Lagrange et al. [99]. 19. Schwab [166]. 20. Kurdi [168]. 21. Bedoya et al. [98]. 22. Lyall et al. [151]. 23. Chowdhury et al. [100].