

# Can Digital Aid Deliver During Humanitarian Crises? \*

Michael Callen  
Miguel Fajardo-Steinhäuser  
Michael G. Findley  
Tarek Ghani

November 11, 2024

## Abstract

Can digital payments help reduce extreme hunger? Humanitarian needs are at their highest since 1945, aid budgets are falling behind, and hunger is concentrating in fragile states where repression and aid diversion present major obstacles. In such contexts, partnering directly with governments is often neither feasible nor desirable, making private digital payment platforms a potentially useful means of delivering assistance. We experimentally evaluated digital payments to extremely poor, female-headed households in Afghanistan, as part of a partnership between community, nonprofit, and private organizations. The payments led to substantial improvements in food security and mental well-being. Despite beneficiaries' limited tech literacy, 99.75% used the payments, and stringent checks revealed no evidence of diversion. Before seeing our results, policymakers and experts are uncertain and skeptical about digital aid, consistent with the lack of prior evidence on digital payments for humanitarian response. Delivery costs are under 7 cents per dollar, which is 10 cents per dollar less than the World Food Programme's global figure for cash-based transfers. These savings can help reduce hunger without additional resources, demonstrating how hybrid partnerships utilizing digital payment platforms can help address grand challenges in difficult contexts.

---

\*Callen: LSE, CERP and CESifo. Email: [m.j.callen@lse.ac.uk](mailto:m.j.callen@lse.ac.uk). Fajardo-Steinhäuser: LSE. Email: [m.fajardo-steinhauser@lse.ac.uk](mailto:m.fajardo-steinhauser@lse.ac.uk). Findley: University of Texas Austin. Email: [mikefindley@utexas.edu](mailto:mikefindley@utexas.edu). Ghani: Washington University in St. Louis. Email: [tghani@wustl.edu](mailto:tghani@wustl.edu). 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, Tim Besley, Robin Burgess, Michael Denly, Arif Husain, Daniel Nielson, Rohini Pande, Lamar Pierce, Soledad Prillaman, Simone Schaner, Jacob Shapiro, Abu Siddique, Garima Siwach, Rebecca Wolfe, Noam Yuchtman and seminar participants at the ACES summer school, the ADB and Yale conference, Brookings, CEPR/TIME/TCD at Trinity College, HBS Tech for All Lab, Innovations for Poverty Action/Northwestern, the Institute for Fiscal Studies, King's College, John's Hopkins School of Advanced International Studies, The University of Maryland, Washington University in St. Louis, Wharton Management in Emerging Markets Conference, and the World Bank shared insightful feedback. Research funding was provided by the UK Foreign, Commonwealth & Development Office, awarded through the J-PAL Crime and Violence Initiative. The authors declare no competing interests. The research protocol was approved by the London School of Economics' Institutional Review Board (study number 89546).

# 1 Introduction

Food insecurity is a systemic global challenge where policy obstacles are multiplying (George et al., 2016). Despite progress over the last half-century, global hunger levels have set new records in each of the last three years and most households experiencing food security crises now live in fragile and conflict-affected states like Afghanistan, Democratic Republic of Congo, and Yemen (FCDO 2023; FSIN 2023; Townsend et al. 2021).<sup>1</sup> In such settings, oppressive state and non-state actors often seek to control resource flows and can restrict humanitarian access to the most vulnerable populations, raising donor concerns that aid will be diverted from intended beneficiaries (Kurtzer, 2019; Cliffe et al., 2023). In such settings, humanitarians face a dilemma: either deliver aid and risk supporting hostile actors and exacerbating conflict (Nunn and Qian, 2014) or suspend operations in the face of urgent needs. Moreover, global aid budgets are failing to keep pace with growing needs, creating pressure to reduce the costs of delivering aid while ensuring it reaches the intended beneficiaries directly.<sup>2</sup>

Can digital payment platforms enable more effective humanitarian crisis response? Humanitarians increasingly transfer physical cash to vulnerable individuals, but that modality is expensive and logistically challenging, subject to diversion by hostile actors that donors do not want to support, and requires in-person contact that is increasingly denied to marginalized populations.<sup>3</sup> Relative to cash, digital payment systems create value by leveraging scalable technology and integrating private firms to facilitate transactions (cf. Dodgson et al., 2015; Wormald et al., 2021; Suri et al., 2023).<sup>4</sup> Delivering aid digitally offers the potential to reduce coordination costs and delays, increase transparency for donors and preserve privacy of beneficiaries, and utilize local supply chains without relying directly on local authorities (Byrd, 2023; De Waal, 2023). While digital payment platforms are widely utilized by development programs for poverty alleviation (Muralidharan, Niehaus, and Sukhtankar, 2016; Suri and Jack, 2016), adoption by humanitarians has been slowed by market and institutional frictions in fragile states (Williamson, 2000; Dorobantu, Kaul, and Zelner, 2017; World Food Programme, 2021b) – particularly given limited evidence on the efficacy of digital aid to assist vul-

---

<sup>1</sup>In 2022, 11 countries - Afghanistan, the Central African Republic, Chad, Democratic Republic of the Congo, Haiti, Nigeria, South Sudan, Sudan, Somalia, Syria, and Yemen - 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 among the roughly 1.350B people analyzed) (Food Security Information Network, 2023).

<sup>2</sup>Humanitarian food security budgets fell 40% from USD 85 per person in 2018 to USD 51 per person in 2021 (Food Security Information Network, 2023).

<sup>3</sup>For example, the share of the World Food Program's (WFP) assistance delivered as cash-based transfers (rather than in-kind aid contributions) rose from 2% in 2010 to 34% in 2021 (McDonough, 2022), a move supported by research evidence (World Bank, 2022).

<sup>4</sup>The United Nations Better than Cash Alliance defines a digital payment as 'the transfer of value from one payment account to another using a digital device or channel. This definition may include payments made with bank transfers, mobile money, QR codes, and payment instruments such as credit, debit, and prepaid cards.' In this study, we examine a specific digital payments platform in Afghanistan, HesabPay, that is available as an app on iOS and Android, and is also accessible to feature phone users via USSD or combined with a QR code card.

nerable, hard-to-reach populations (Aker et al., 2016; Pazarbasioglu et al., 2020; Gen-tilini, 2022).<sup>5</sup>

In this study, we conduct a randomized evaluation to study the impacts of a digital aid program in Afghanistan. Specifically, we compare a treatment group receiving digital aid payments against a comparison group which does not receive aid until later. The program was a hybrid partnership involving community, nonprofit, and private organizations. Locally elected community councils identified extremely poor, tech-illiterate, female-headed households in three cities: Kabul, Herat, and Mazar-i-Sharif. Working with an Afghan digital payments platform and a nationwide Afghan NGO, a U.S.-based nonprofit organization transferred digital value vouchers to beneficiaries' mobile phones that could be spent at participating local private merchants.<sup>6</sup> The intervention transferred 45 USD payments every two weeks (the equivalent of 3.75 months worth of average monthly household income in the sample) for two months.<sup>7</sup> The experiment included 2,409 beneficiary households, randomly assigned to treatment and control groups. Given our study's focus on relieving human suffering, we make use of appropriate methods to measure key outcomes and to avoid false positives, including rigid adherence to a Pre-Analysis Plan (PAP).<sup>8</sup> Crucially, the program leveraged three enabling conditions: widespread access to mobile phones, the digital platform's robust merchant network, and sufficient market availability of food products to meet demand.

Our experiment yields three results on digital aid's efficacy. 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 physical "cash out" option. Second, digital payments improved all prespecified measures of nutritional well-being, with an index of these measures increasing by  $0.5\sigma$  (SE = 0.032;  $p \leq 0.0001$ ) and all prespecified measures of mental and financial health, with an index of these measures improving by  $1.5\sigma$  (SE = 0.042;  $p \leq 0.0001$ ). Last, rigorous tests reveal no evidence of diversion either from beneficiaries or from merchants accepting digital payments. Our conservative estimates of the cost of delivery – including all aspects

---

<sup>5</sup>GiveDirectly (2024) catalogs the vast literature on cash transfers, including important contributions by Haushofer and Shapiro (2016); Bedoya Arguelles et al. (2019); Banerjee et al. (2020); Muralidharan et al. (2021); McIntosh and Zeitlin (2022). While prior cash transfer studies focus mainly on sustained poverty reduction (i.e. SDG 1), we are concerned with basic human survival (i.e. SDG 2); online Appendix A discusses prior cash transfer studies focused on humanitarian goals.

<sup>6</sup>These value vouchers have a denominated local currency value redeemable at participating merchants for any available goods. World Food Programme (2023e) treats value vouchers as an official form of a cash-based transfer, unlike commodity vouchers, which are both tied to specific merchants and goods.

<sup>7</sup>The Afghanistan Cash and Voucher Working Group (CVWG) estimates a Basic Food Basket for a family of seven costs approximately 96 USD per month (Bete, 2022), roughly equivalent to two biweekly direct aid transfers of 4000 AFA and a monthly survey incentive of 350 AFA. This 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.

<sup>8</sup>The Pre-Analysis Plan is registered [here](#).

of recruitment and facilitation – are 6.7 cents per dollar, or less than 40% of the [World Food Programme \(2023b\)](#) global benchmark figure of 17 cents per dollar for cash-based transfers.

Our results exceeded the expectations of experts in this field as elicited prior to the release of the study. To benchmark beliefs, we provided a precise description of the intervention to 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. Asked to predict key outcomes of the study, respondents were skeptical and uncertain. On average, they predicted only 43% of women in our study would be able to use digital payments ( $SD = 25.67$ ); in practice, 99.75% of women purchased goods at least once. Likewise, they predicted, on average, that local authorities would attempt to tax roughly 40% of beneficiaries ( $SD=27.28$ ); in practice, less than 2% of the treatment group reported any diversion attempts, with no statistically significant difference in the control group.

Our work builds on two extensive bodies of social science research on cash transfers – one documenting the usefulness of digital transfers to achieve development outcomes ([Bastagli et al., 2019](#); [GiveDirectly, 2024](#)) and the second focused mainly on the benefits of distributing physical cash to vulnerable populations during humanitarian crises ([World Bank, 2022](#)). While the former literature in development contexts often takes the existence of robust digital payment platforms for granted, the latter literature in humanitarian crises rarely engages the private sector as a delivery partner. In Online Appendix A, we systematically review the intersection of these two literatures and find limited evidence regarding digital payments to hard-to-reach populations in fragile states during humanitarian crises – the closest digital aid papers examine comparatively stable settings including refugee camps and natural disaster response (cf. [de Hoop, Morey, and Seidenfeld, 2019](#); [MercyCorps, 2022](#)). Our primary contribution is to provide clear proof-of-concept that digital transfers can cost-effectively address humanitarian needs for highly vulnerable, hard-to-reach groups while avoiding diversion. In exploring the efficacy of digitization for humanitarian response, we echo the emphasis of [George, Merrill, and Schillebeeckx \(2021\)](#) on both technological innovation and “developing business models that infuse innovations with new purpose.” Indeed, the digital payments platform utilized in this study was developed prior to the Taliban’s 2021 return to power with the original intent of serving a growing Afghan middle class, and only afterwards shifted its focus to assist in aid delivery to the most vulnerable citizens.

Our work also contributes to a nascent strategy-related literature on grand challenges affecting large populations in difficult contexts (cf. [Ballesteros, Useem, and Wry, 2017](#); [McGahan and Pongeluppe, 2023](#); [Fangwa et al., 2024](#)). Scholars are increasingly focused on organizational approaches to grand challenges by analyzing the comparative strengths of public and private actors in addressing social issues ([Luo and Kaul,](#)

2019; George et al., 2024). The hybrid arrangement we study in this program – involving community, nonprofit, and private organizations – differs from most previously documented responses to global challenges in that government was not an active partner given donor and humanitarian concerns about engaging the Taliban authorities. This program relied on community councils created by the previous government, which in turn notified local Taliban authorities to ensure the safety of operations. Limiting the Taliban’s opportunities to divert aid was a central motivation for digital aid delivery, and we make considerable methodological efforts to measure diversion rigorously. Given increasing awareness that many of the Sustainable Development Goals will not be met in 2030 absent major innovations in fragile states (Ghani and Gordon 2022; OECD 2022; FCDO 2023), our findings indicate one hopeful pathway – technological and organizational – for the private sector to help address growing global humanitarian needs.

## 2 Research design

We develop, implement, and evaluate an approach to delivering digital transfers to a vulnerable, hard-to-reach population during Afghanistan’s humanitarian crisis. We briefly describe the research design in this section, with the supplementary materials providing more details. Specifically, Online Appendix B describes the context, implementation, estimation strategies, deviations from the PAP and benchmarking of the estimated treatment effects. In addition, Online Appendix C discusses the cost-effectiveness and cost-efficiency estimation and Online Appendix D discusses ethical considerations.

### 2.1 Theory of change

In principle, digital payments should enable the transfer of value to vulnerable beneficiaries who can then purchase food and thus alleviate their household’s food insecurity. However, several enabling conditions must first be met: vulnerable households must have access to mobile phones, enough merchants must accept digital payments to allow convenient and competitive shopping, and the market must have enough goods to meet demand. Even if these conditions are met, other obstacles might still limit the effectiveness of digital aid: illiterate Afghan women may have difficulty utilizing digital payments, local merchants may find digital transactions overly cumbersome, and the Taliban authorities may decide to interfere.

Following an extensive piloting process, this digital aid program was designed with the goal of ensuring a successful implementation despite the challenges posed by the context and population. Beneficiaries were selected through locally-elected Community Development Councils (CDCs), who informed local Taliban authorities of their activities but denied them influence over participant selection. Tech illiteracy was ad-

dressed through a careful onboarding process described below, which also ensured that merchants were well-prepared to meet the needs of this population. Finally, research procedures were put in place to ensure high quality data and valid inference procedures to minimize the risk of false positives.

## 2.2 Partners

The program was a hybrid arrangement designed to minimize interference by Taliban authorities and address common issues facing digital transfers – including technical issues like interoperability and smartphone ownership and social issues like tech literacy and trust. The research team composed of academics designed all research-related components of the project. Locally-elected Community Development Councils (CDCs) identified female-headed households as program beneficiaries through community-level, participatory meetings and facilitated their engagement in the program. Uplift Afghanistan, a U.S.-based nonprofit, received grant funding from a private foundation and managed transfers to beneficiaries. Lastly, HesabPay, a digital payments platform that was compatible across mobile network operators, facilitated transactions through its network of private merchants. This division of labor is consistent with [Luo and Kaul \(2019\)](#), who argue for-profit firms are best-positioned to innovate (as in the case of building a digital payments ecosystem), self-governing collectives have a comparative advantage in private ordering (as in determining which potential beneficiaries to prioritize), and nonprofit organizations are well-suited for fiduciary roles. All partners shared a common interest in testing whether digital payments were a viable and attractive channel for humanitarian aid.

## 2.3 HesabPay's technology and merchant network

HesabPay is a digital payments platform that is interoperable across Afghanistan's mobile networks and transfers value using either a smartphone app or via a feature phone using a combination of transaction initiation via QR code and transaction verification via feature phone USSD.<sup>9</sup> HesabPay uses the Algorand blockchain as a settlement layer, facilitating finalization and recording of digital transactions; HesabPay users are automatically registered for a custodial Algorand wallet (e.g. HesabPay holds their private keys on their behalf) and all transactions are automatically recorded on the Algorand blockchain without disclosing user identities. HesabPay is licensed by Afghanistan's Central Bank as a financial service provider, similar to the regulatory framework for HesabPay's main competition from Afghanistan's mobile money operators.

At the time of the study, HesabPay had an active and growing local merchant acceptance network in Afghanistan's major cities, composed of over one thousand pre-

---

<sup>9</sup>HesabPay was founded in 2016 by entrepreneur Sanzar Kakar. For more details, see the Algorand Foundation case study [here](#).

existing Afghan businesses providing basic foodstuffs and other household items for sale in the same neighborhoods as beneficiaries. In this program, aid payments took the form of a digital value voucher denominated in local currency that could be exchanged for any available goods from any HesabPay enrolled merchant. All beneficiaries conducted a test purchase with a nearby merchant immediately following their onboarding session. While many chose to return to that specific merchant for future purchases, we confirm broader engagement with the acceptance network using beneficiaries' transaction data from the digital payments provider. A1 reports the total number of merchants serving beneficiaries in the transaction data, which is consistently larger than the number from onboarding sessions.<sup>10</sup> A1 maps the locations of these merchants for each of the three cities in which we conducted the intervention, visually demonstrating the decentralized distribution network.

## 2.4 Beneficiary identification and onboarding

Our goal was to identify about 2400 vulnerable women in three Afghan cities (Kabul, Herat and Mazar-i-Sharif) to be part of the intervention. To do so, we worked with the Community Driven Development Organization (CDDO), an Afghan organization that assists CDCs in a wide-array of local activities.<sup>11</sup> Local CDCs identified potential beneficiaries through a 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 deliberation regarding who is most vulnerable. We note that ownership of a feature phone was not a binding constraint for any beneficiaries, reflecting high nationwide levels of mobile phone ownership.<sup>12</sup>

After identification, participants were invited to local onboarding sessions where four activities took place.<sup>13</sup> First, the program was described to them. Second, if they wanted to participate, their informed consent was collected and a baseline survey was

---

<sup>10</sup>Note that this likely underestimates the true number of merchants in any given area, since i) we only contacted numbers with which our beneficiaries transacted more than 10 times, ii) among those numbers with which our participants transacted more than 10 times, we couldn't reach all of them, and iii) there are more merchants in the areas that accept the digital vouchers but our participants did not visit.

<sup>11</sup>The CDCs were established through local elections as part of the National Solidarity Program starting in 2004 (Beath et al., 2016; Beath, Christia, and Enikolopov, 2017b, 2013, 2017a, 2018), 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.

<sup>12</sup>Mobile phone ownership in Afghanistan has grown rapidly over the past two decades, from approximately 25,000 subscribers in 2002 to over 22 million subscribers in 2021 (World Bank, 2023b). 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 (Asia Foundation, 2019).

<sup>13</sup>These took place during September 2022 in each of the three cities, with each session having between 24 and 80 participants.

administered. Third, each woman opened an account with HesabPay, an Afghan commercial digital payments platform. Finally, each woman completed a test purchase with a nearby private merchant using the platform.<sup>14</sup> The CDDO, working together with local CDCs, ensured the safety of staff and beneficiaries by informing local authorities and maintaining independence from external interference in the beneficiary selection process.

## 2.5 Intervention & Randomization

Uplift Afghanistan transferred digital value vouchers to beneficiaries' mobile phones that were redeemable at HesabPay's acceptance network of local merchants for any available goods. Specifically, the intervention transferred 4,000 AFN (approximately \$45 USD) every two weeks for two months to households in 16 urban 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 December 31, 2022 (henceforth, treatment group). 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 (henceforth, control group). The randomization was stratified on two variables: The nahia (neighborhood) in which they registered (our study includes beneficiaries from neighborhoods), and a measure of vulnerability.<sup>15</sup> As can be observed in A2, the treatment groups are balanced on 17 of the 18 outcome and heterogeneity variables we collected at baseline.<sup>16</sup>

While all participants understood they would eventually receive transfers, randomization took place after all onboarding sessions were completed and we only informed participants in both groups that they were going to start receiving their payments a few days before their first payment. Note that this is not a setting where consumption smoothing was feasible given widespread and acute hunger.<sup>17</sup> We identify causal impact by using the late group as a control for the treated early group during the two months in which the early group was receiving its payments and the late group

---

<sup>14</sup>During pilots we conducted prior to the actual intervention, it became clear that participants needed help at first using the digital payments platform, as most of them were illiterate and had no prior experience with mobile money. This is why we decided to use the onboarding sessions to both help participants create their accounts and resolve any remaining questions and also to have participants complete a test purchase with a nearby merchant using the mobile payment platform.

<sup>15</sup>Participants in the same nahia had access to the same merchant acceptance network, motivating this decision. We discuss the stratification procedure in detail in our pre-analysis plan.

<sup>16</sup>As we note on P. 15 of the PAP, with 18 balance variables and a 5% significance level, "we would expect to see p-values of less than 0.05 in 2 of the variables." We thus committed to implement the first randomization draw with less than two variables with a p-value below 0.05. This condition was met on the first random seed we tried, resulting in one unbalanced variable out of 18.

<sup>17</sup>We pre-specified (on P. 13 of our Pre-Analysis Plan) that our sample would need to be credit constrained in order to interpret differences between our treatment and control groups. Empirically, we asked if women could borrow - only 0.29% of our sample indicated that they could - and we found no evidence of borrowing against future payments in the control group.

was not receiving any payments. The study design and CONSORT flow diagram are depicted in Figure 1. Online Appendix D details the ethical considerations addressed in the design and implementation of the study.<sup>18</sup>

## 2.6 Data collection

Data were collected four ways. First, a baseline survey was completed during the on-boarding 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.<sup>19</sup> Fourth, we attempted to survey all 26 merchants who facilitated initial test transactions with beneficiaries, and succeeded in contacting 19, to confirm how beneficiaries had used their payments and to check whether they were asked to provide favors or extra-legal taxes to local authorities.

We pre-specified all analyses in a PAP, including how the outcome variables would be constructed, and what our primary outcomes variables were going to be. We divided our outcome variables in three families: Basic needs (also called "food security" in the manuscript), well-being and informal taxation (also called "diversion" in the manuscript). In doing so, we tried to provide a level of detail consistent with the standard articulated in [Duflo et al. \(2020\)](#). This standard requires that two research assistants can take the data and the PAP, and with only these two items separately produce identical analysis. For each of these three families of outcomes, we create a summary index following [Kling, Liebman, and Katz \(2007\)](#).

---

<sup>18</sup>We obtained Institutional Review Board approval on 4 May 2022 from the London School of Economics (#89546). 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. 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, the 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. See Online Appendix D for more details, including considerations specific to insecure, humanitarian crisis environments ([Wood, 2006](#); [Campbell, 2019](#); [Puri et al., 2017](#); [Wolfe, 2020](#)).

<sup>19</sup>All surveys were conducted by female enumerators, consistent with local norms. Participants were informed that the survey was completely voluntary and would not affect their aid payments, and they could skip any question they did not want to answer. Participants received a 350 AFA (~4 USD) payment for completing each survey (see Online Appendix D). Overall, we have response rates of about 99% across survey rounds, with no difference in response rates between the treatment and control groups.

## 2.7 Methodology

To estimate treatment effects, we use linear regressions (see Online Appendix 2 for estimating equations). We study the intervention's effect by regressing the pre-specified outcome variables on an indicator for treatment assignment (baseline analysis). We control for the outcome variable's baseline value (if available), strata fixed effects, and survey round fixed effects. Standard errors are clustered at the level of treatment assignment (the participant). Note that due to very high response rates and no compliance issues, we do not conduct treatment-on-the-treated regressions. To avoid false discoveries due to multiple hypothesis testing, we control for the Family Wise Error Rate (FWER) for each family of primary outcomes and control for the False Discovery Rate for secondary outcomes as pre-specified in our PAP.

We conduct several complementary analyses as well, which were also pre-specified. First, we study how the effects change over time by including an interaction term between the treatment assignment indicator and an indicator for the second survey round. We also study dynamic effects graphically by computing the means of given outcome variables at the survey week level for treatment and control groups separately, as shown in Figure 2. Second, we study whether the effects are heterogeneous along several (pre-specified) variables, by interacting each of these variables with the treatment assignment indicator.

We check whether experimenter demand effects are present by randomly informing participants about the goals of the study (see Online Appendix 3 for details). We analyze whether providing information about the goals of the study impacted survey responses in two ways: regressing each outcome on an indicator for the informational assignment indicator, and including an interaction term between an informational assignment indicator and the treatment assignment indicator (to see whether the information affects participants in the treatment and control groups separately).

## 3 Results

We organize our main results around the following primary Research Questions (RQs):

- RQ1: Can extremely poor, tech-illiterate women use digital transfers in Afghanistan?
- RQ2: Do digital transfers improve food security and mental well-being?
- RQ3: Are digital transfers diverted?
- RQ4: How much does it cost to deliver digital aid?

### 3.1 Can extremely poor women use digital transfers?

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 have no schooling and 33.9%

have at most primary education.<sup>20</sup> 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.<sup>21</sup>

Ninety-eight percent 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 effective division of labor between program partners, an emphasis on human-centered design and extensive piloting (Online Appendix 1 describes the piloting process). In particular, guiding beneficiaries through a test transaction during onboarding greatly increased comfort with digital payments.

Beyond the high levels of usage, three additional results support the argument that this tech-illiterate population can use digital transfers. 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 PAP, we also checked whether impacts on food security outcomes vary by pre-intervention need, city, marital status, age, household size, education, and whether the recipient was 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 (A2).

Finally, we asked participants in a hypothetical exercise whether they would prefer to receive their full 4000 AFN payments digitally, or  $4000 - X$  AFN in cash, where  $X$  was either 100 AFN, 300 AFN, 500 AFN. These amounts might reflect the costs required to provide aid as physical cash (see Online Appendix 2 for details). With a conservative fee of 100 AFN (2.5%), 80% of participants preferred digital aid over cash; with a higher fee of 300 AFN (7.5%), the share choosing digital aid over cash increased to ~95% (A3). Collectively, the results indicate that digital payments are a viable option even when people have limited experience using digital technology.

### 3.2 Do digital transfers improve food security and mental well-being?

The payments reduced all four pre-specified measures of food security by meaningful amounts (see Table 1, Panel A for estimates and control means). Beneficiaries, on average, skipped meals in 0.76 fewer days per week ( $SE = 0.051, p \leq 0.0001$ ), children were 11.7 percentage points less likely to have skipped meals over the past week ( $SE = 0.012, p \leq 0.0001$ ), the share of households where everyone was able to eat regularly during the prior week increased by 9.3 percentage points ( $SE = 0.015, p \leq 0.0001$ ), and

---

<sup>20</sup>Using 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.

<sup>21</sup>Three 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.

beneficiaries reduced meals of only bread and tea by 1.608 (SE = 0.121,  $p \leq 0.0001$ ). An index of these four measures – constructed as the average of the standardized measures following [Kling, Liebman, and Katz \(2007\)](#) – improved by 0.5 SDs (SE = 0.032,  $p \leq 0.0001$ ).

Under secondary outcomes, we find evidence of a more diverse diet (see Table 1, Panel B for estimates and control means). With a recall period of one week prior to the survey, beneficiaries ate rice on 0.6 days (SE = 0.035,  $p \leq 0.0001$ ), beans on 0.49 days (SE = 0.029,  $p \leq 0.0001$ ), chicken on 0.01 days (SE = 0.006,  $p = 0.035$ ), and dairy on 0.05 days (SE = 0.013,  $p \leq 0.0001$ ) more than the control group. Participants also report an increased ability to purchase medicine when needed. Consistent with pre-registered expectations in our PAP, we do not find any increases in outside income, employment, or agency over financial decisions.<sup>22</sup>

Primary food security outcomes kept improving for the two months during which beneficiaries were paid (Table 2), and more modest improvements remain for at least two months after payments conclude (Figure 2). Regarding dietary diversity, estimates reported in A3 indicate that consumption of beans and rice is increasing over time.

To assess if these estimated impacts are in line with what we should expect given the size of the payments, we collected data from a subsample of 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 ([Bete, 2022](#)). While not eliminated, skipped meals declined after each payment (Figure 2, Panel A). Online Appendix 5 provides further discussion of the magnitudes of needs results.

Turning to mental well-being, since the Taliban takeover in 2021, Afghanistan consistently ranks as the country with the lowest levels of happiness ([Gallup, 2022a,c](#)). Our participants report extremely low levels of happiness and financial health. Treatment improved all four prespecified measures of mental and financial health, with an index of these measures improving by 1.5 SDs (Table 1, Panel B), although from a very low base and these impacts disappear as soon as payment stops (Figure 2). 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 increased measures of mental well-being: 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,  $p \leq$

---

<sup>22</sup>Regarding income and employment, we write in our PAP: “due to the existing restrictions on women’s liberties in Afghanistan, we believe it is unlikely that [these] will change.”

0.0001), relative to the control group.<sup>23</sup> 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 (Table 2, Panel B).<sup>24</sup>

Our results are robust to common concerns in Randomized Control Trials (RCTs). First, we find no evidence of treatment beneficiaries providing part of their payments to control beneficiaries or of control beneficiaries borrowing against payments they would receive in the last two months of the program (see Online Appendix 3). Second, given that participants had 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 those estimated on the complete sample (see A5).

Finally, 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. Specifically, in the second round of follow-up surveys, 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 find no evidence that the prime influenced responses (see Online Appendix 3 and A4).

### 3.3 Are 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 sponsor international terror. Reports of aid diversion across multiple fragile countries have emerged in the past year (O'Donnell, 2023; UN News, 2023; Politico, 2023) and have been documented extensively for decades (Barnett, 2013) with scholarly evidence suggesting that diversion

---

<sup>23</sup>While these results may seem unrealistically large, this is attributable to low base levels, as is seen in other surveys in contemporary Afghanistan (Gallup, 2022a,c).

<sup>24</sup>Our food security results complement a broader literature demonstrating mostly positive impacts, whereas our mental well-being results contribute newer insights to a more nascent literature. As discussed in Online Appendix 2, our review of 23 cash-based programs in humanitarian contexts found that 21 included food security outcomes, with 16 of those 21 finding any positive effect. Less attention has focused on the mental impacts of cash-based programs, with only seven of the 23 studies considering a relevant outcome and six finding any positive effect.

prolongs conflict (Findley, 2018; Nunn and Qian, 2014). 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 they directly or indirectly supported the Taliban (Binesh, 2023). Reports of Taliban aid requirements and infiltration of UN assistance (O'Donnell, 2023; CBS News, 2023) are deepening these concerns, adding further pressure to cut assistance. Therefore, the humanitarian mandate to address hunger depends partly on avoiding diversion.

There are at least four ways that hostile regimes might capture aid. First, they might influence who is eligible, including by creating fictitious “ghost” beneficiaries. Second, they can ask recipients to hand over aid transfers. Third, they can capture aid while in transit. Fourth, they can ask merchants or payments platforms who are serving aid beneficiaries to pay bribes or additional tax. We organize the presentation of results around these four potential diversion strategies and discuss whether digital delivery might constrain them. Such diversion strategies are widely documented both in Afghanistan (Special Inspector General for Afghanistan Reconstruction, 2023b; Sopko, 2023) and in other countries (see Online Appendix 1 for a discussion of diversion of humanitarian aid in the literature).

**Influencing eligibility:** In our study, beneficiaries were identified by local elected community councils in consultation with communities using a “well-being analysis” described in Online Appendix B. The prevalence of mobile phones enables phone surveys as a means of checking that beneficiaries are indeed vulnerable, even at large-scale. Our surveys confirm that beneficiaries were quite vulnerable (see A6). The digital payments platform also requires Know Your Customer identity verification, which can automatically compare beneficiary names against sanctions lists to further guard against capture by hostile actors. Evidence from a diverse set of global contexts (Guggenheim and Petrie, 2022; Samii, 2023) and the pre-2021 Afghanistan NSP program specifically (Beath, Christia, and Enikolopov, 2018; Burde et al., 2023) also indicates that most diversion occurs at national and regional levels and primarily through bureaucratic and partisan channels. At the local level, in contrast, CDCs develop transparency and monitoring mechanisms to prevent diversion, making it more challenging for local authorities to interfere in influencing eligibility.

**Taxing beneficiaries:** To check for taxation, 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 (Reinikka and Svensson, 2006), and then whether they themselves had been asked. Results are presented in the first four columns of Table 3, Panel A, which show that the treatment group does not report informal payments to authorities in larger proportion than the control group regardless of how the question was framed (all coefficients are insignificant at traditional levels, and precisely estimated). When looking at the KLK Index combining the four individual questions (column 5), the coefficient is marginally significant at the 10% level. However, this is driven by a single individual who answered yes to three of the four questions, which due to the way the index is constructed, receives an extremely high index value of 34 standard deviations. Removing this observation leads the results using the KLK Index to lose statistical significance at conventional levels (column 4 in A7).

We find very low levels of informal taxation: Overall, only 27 beneficiaries in the treatment group answered yes to any of these four questions since payments started, compared to 21 beneficiaries in the control group. When using an indicator for whether the respondent answered yes to any of the four informal taxation questions the difference is statistically insignificant (Table 3, column 6).<sup>25</sup>

Estimates in Table 3, Panel B, indicate that the results are also not increasing over time, when beneficiaries could have become more visible to local authorities. Based on qualitative debriefs with a subsample of our respondents, one potential reason for limited diversion is that participants are too poor for local authorities to ask them for payments. Moreover, it is politically and logically costly to tax vulnerable beneficiaries after they have received their payments.

Other surveys in Afghanistan indicate that significant shares of Afghans are comfortable reporting corrupt behaviors, at least before the Taliban takeover (Asia Foundation, 2019). Even so, given the authoritarian context and the fact that these are questions that may be sensitive for participants (Blair, Coppock, and Moor, 2020), there could be a lack of positive responses due to fear of reporting inappropriate behavior by local authorities. We therefore conducted a list experiment, an established method to measure the presence of sensitive behaviors, with beneficiaries to gauge the extent of informal taxation (see Online Appendix 2 for details). 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, which includes the informal taxation statement, do not report experiencing more items on the list than those receiving the shorter list. The effects are precisely estimated and small in magnitude, supporting the argument that diversion was minimal.

We also have access to the beneficiaries' transaction data from the digital payments provider, which coupled with the fact that they could not cash out their pay-

---

<sup>25</sup>Note that this categorical variable for whether the respondent answered yes to any of the four diversion questions was not pre-specified in our PAP. In Online Appendix 4 we describe all deviations from the PAP.

ments and that they spent almost all the money they received, provides a clear picture of how beneficiaries spent their funds. Beneficiaries spent 74.2% of their funds at the merchants they visited during the beneficiaries' onboarding sessions, 20.9% at other registered merchants, 3.5% at individually-registered accounts, and 0.1% on airtime purchases. Combined with the fact that almost no respondents indicate providing assistance to local authorities, this also supports the argument that there was little diversion. Moreover, only 6.7% of beneficiaries reported that someone else decided how to use the aid payments – all of whom were other household members.

Importantly, once beneficiaries are onboarded, distribution does not require travel outside the home as it might with food or cash distributions. The distribution of digital aid is therefore less visible than cash or in-kind distribution. Finally, the distribution of digital payments is instantaneous, with recipients able to access their funds immediately. Figure 3, 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. A similar pattern holds every time participants receive their funds. Panel B displays how participants spent their money on a daily basis.

**Capturing aid in transit:** Relative to in-kind or physical cash, which rely on intermediaries that staff convoys or physical distribution points, digital payments may reduce the opportunities for in-transit theft by transferring aid directly to beneficiaries. Also, with automatically generated 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. The digital payments platform in this study utilizes the Algorand blockchain, so all payments and purchases were automatically recorded on an immutable ledger that facilitates external auditing. Finally, phone surveys also provide a scalable means of confirming the delivery of benefits, as also demonstrated in [Muralidharan et al. \(2021\)](#).

**Taxing merchants or intermediaries:** The government has the legal authority to tax both merchants and digital payment platforms. To check for merchant taxation, we surveyed 26 merchants who assisted beneficiaries with an initial test transaction to ask about extra-legal taxes (merchants do need to pay ordinary sales tax on any transactions). None of the merchants reported paying extra-legal taxes, being asked to provide favors, or being aware of other merchants paying extra-legal taxes or providing favors.

In our study the digital payments provider reported no attempts at extra-legal taxation. This could plausibly change for a larger initiative. In this scenario, a single tax would be levied on a single provider. Economic theory suggests a single, negotiated tax

is preferable to numerous, less-organized attempts at interception, potentially reducing overall diversion ([Shleifer and Vishny, 1993](#)).

### 3.4 How much does it cost to deliver digital aid?

The cost-efficiency of digital aid, or the program costs required to serve a given number of beneficiaries, outperforms comparable estimates. The total cost per beneficiary (CPB) of delivering the entire \$180 (across four disbursements) to our beneficiaries is \$2.44 USD excluding recruitment costs or \$12 USD including recruitment costs. Correspondingly, the cost to deliver a single dollar including the transferred dollar, termed the total cost transfer ratio (TCTR), is either \$1.014 (\$182.44/\$180.00) or \$1.067 (\$192.00/\$180.00). Excluding the transferred dollar, the cost-transfer ratios (CTRs) are either 1.4 cents when recruitment costs are excluded or 6.7 cents when recruitment costs are included. For comparison, the World Food Programme's (WFP) global CTR for providing cash-based humanitarian aid is 17 cents per dollar ([World Food Programme, 2023b](#)). To contextualize this difference, we estimate that if WFP had delivered all \$357M of its 2022 cash-based assistance in Afghanistan digitally, the savings would be sufficient to support an additional 77,000 households or 538,075 individuals for the four-month lean season.

We also computed the costs required to produce a given level of program impact, a measure of cost-effectiveness. Online Appendix [C](#) details these cost-efficiency, scale, and cost-effectiveness estimates and reports ranges of estimates based on more or less conservative assumptions. The cost-efficiency estimates compare favorably to other cash-based transfer programs in humanitarian contexts. Although cost analysis is rarely performed for humanitarian cash programs ([Gentilini, 2016](#); [World Bank, 2022](#)), we identified three organization-specific global estimates and 12 individual studies with estimates. When cost categories are comparable, all of these organizations and studies have higher costs of delivery than our digital aid intervention.

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 physical cash 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. When payments are delivered digitally 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. This is particularly im-

portant when considering women in food-insecure settings as existing research documents substantial benefits to channeling resources to women directly (Field et al., 2021; Riley, 2022). Beneficiaries in our study used their digital transfers across the merchant acceptance network, increasing convenience and privacy while also ensuring competitive pressure on any single merchant who might otherwise seek to engage in price gouging.

## 4 Experts' beliefs and methodological safeguards

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. This may lead to a reasonable reluctance to adopt new technologies, especially if they are untested or perceived as risky. To better understand these issues, we surveyed experts, asking them to predict the treatment impacts we would later estimate. To help address the same issues, we took several methodological steps to safeguard against false positives.

We measure the expert views by surveying 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.<sup>26</sup> The survey provided a precise description of the intervention and asked respondents to predict the corresponding treatment impacts for several key outcomes we report here; see [A8](#) for the full text of the survey prompt.

Table [4](#) compares the actual values of the intervention's components (column 1) to the respondents' predicted values (column 2), and shows the *p*-value of a test of equality (column 3). Across all four components, differences between expert predictions and the values from our data are statistically significant. These respondents predicted, on average, that only 43% of the women in our study would be able to use digital payments (SD = 25.67pp). In practice, as described above, 99.75% of women purchased goods at least once. They also predicted, on average, that local authorities would attempt to tax roughly 40% of beneficiaries (SD = 27.28). In practice, less than 2% of the treatment group reported any diversion attempts, with no statistically significant difference in the control group. Moreover, respondents had little agreement in their views, reflected in the large variance across predictions. A combination of skepticism and uncertainty about potential efficacy might partly explain an unwillingness to adopt innovations, especially when stakes are high. Experts did a better job of predicting cost-efficiency, guessing on average that delivery costs would be 10.65 cents on the

<sup>26</sup>The 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.

dollar, while the actual delivery costs were 6.7 cents on the dollar, indicating a more common belief that digital aid delivery is relatively cheap.

Because of the potential stakes, we implemented three methodological safeguards against false positives (whereby the null hypothesis that an innovation has no impact is incorrectly rejected in favor of the alternative that it is effective). First, we conducted an RCT, which provides 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 (Wolfe, 2020). Second, we strictly followed our PAP, which, while limiting our capacity to learn and adapt, helps ensure that  $p$ -values are correctly calculated, providing additional protection against false positives (Olken, 2015; Duflo et al., 2020). We also adjust our estimates to take into account the multiple hypotheses we are testing following the PAP. Third, as we rely on self-reported measures and not on anthropometrics or biometric markers due to surveying limitations, survey bias and strategic responses by participants are a source of concern. In our case, we followed the literature on testing and avoiding survey response bias (De Quidt, Haushofer, and Roth, 2018) and took the most conservative approach we could think of to test for survey bias (see Online Appendix 3).

## 5 Discussion

We evaluate a digital aid program and show that extremely poor, tech-illiterate women can receive humanitarian aid through digital payment systems to meet basic food security and mental well-being outcomes. We also find no evidence of diversion. Accounting for all aspects of facilitation, the estimated 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. Our estimates are also better than the predictions of policy and research experts. Why did digital aid outperform these expectations?

The hybrid partnership that delivered the program was carefully designed to address common issues working with poor, tech-illiterate populations in a politically fragile context. Vulnerable households were identified by locally-elected CDCs, which ensured the safety of staff and beneficiaries by informing local authorities and maintained independence from interference in the beneficiary selection process. A U.S.-based nonprofit provided fiduciary oversight. The digital payments platform was compatible across mobile operators and did not require a smartphone, ensuring beneficiaries could use existing phones to access transfers. In addition, user-centered design principles helped minimize potential confusion: most importantly, beneficiaries were trained to use the payments platform by completing a test transaction with a merchant.

We expect that whether the findings obtained from this context and population generalize to other settings will depend most critically on three key enabling conditions. First, vulnerable households must have access to phones. Second, there must be enough merchants who accept digital payments to allow convenient and competitive shopping. And third, markets must have enough goods to meet demand. Transitioning from in-kind to cash and digital support can increase competition for goods and so may increase market prices and thereby displace hunger to other vulnerable groups. Studies of famine highlight that in extreme cases of sudden and widespread deprivation – as in Gaza at the time of writing ([Chotiner, 2024](#)) – increasing food supply is critical, but that response to most famines historically required complementary strategies to redistribute resources ([Dreze and Sen, 1990](#); [Sen, 1999](#)).

Our study is motivated by the challenges facing populations who are isolated by the state in complex crisis situations like Afghanistan, where basic human survival – not sustained poverty reduction – is the immediate policy goal. Evidence on whether digital humanitarian aid can help achieve this objective is limited. Most cash-based programs are carried out in more stable contexts, contributing to a rich evidence base on outcomes like poverty reduction, education, health, financial inclusion, and social protection ([Bastagli et al., 2019](#); [GiveDirectly, 2024](#)).

Based largely on this research, humanitarian cash-based programming is growing and digital delivery is increasingly popular ([Urquhart et al., 2023](#)), but research during humanitarian crises with populations facing extreme cultural and political constraints is still nascent. In Online Appendix 1, we review the 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. This knowledge gap is consequential as these contexts and groups are increasingly affected by acute hunger ([Foreign, Commonwealth and Development Office, 2023](#)). As our expert survey demonstrates, our results are not readily inferred from prior studies: illiterate Afghan women could have been unable to utilize digital payments, local merchants could have refused to accept digital transactions, and the Taliban authorities could have interfered. Our study thus provides an “existence result” on the feasibility of digital aid in humanitarian response, with insight into the organizational and technological factors that contributed to success.

The exigencies of this context and the fact that we work with a hard-to-reach group, however, necessitated three design choices that affect our generalizability. First, our experimental data cover only two months before the early group is treated. This time frame, however, is consistent with our learning objective: can digital payments cost-effectively reduce acute hunger. As such, we pre-specified this as the relevant time frame, and relatedly that we did not expect changes in outcomes like income or employment. Such impacts were unlikely – and did not realize – given the large humanitarian needs, prohibitions on women’s employment, and relatively modest transfer

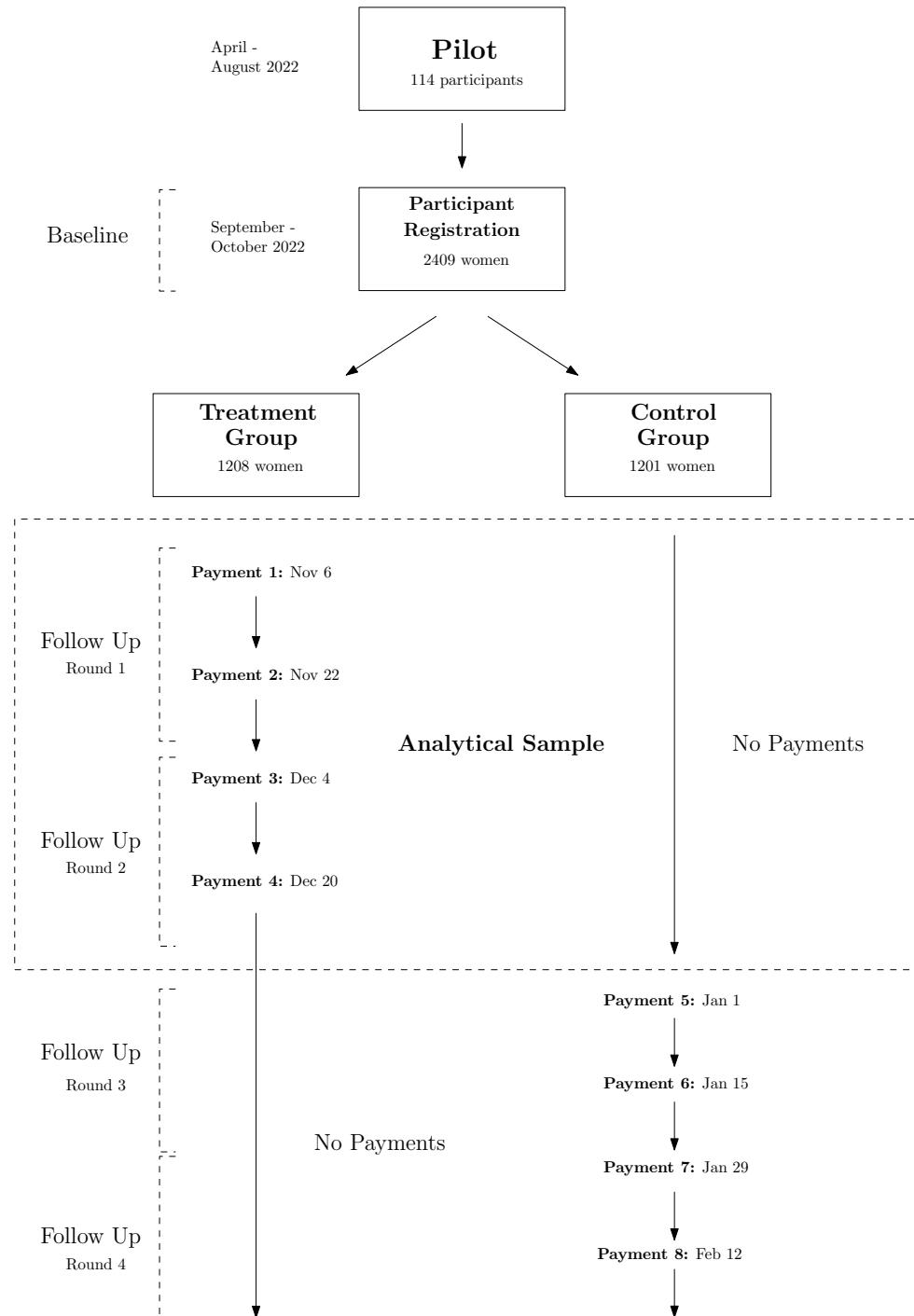
size. Second, working with our sample required approval of the *de facto* authorities. Our partners could not justify seeking Taliban permission to onboard women to the digital payments app without eventually providing humanitarian assistance. We thus opted for a short, staggered program where every participant received some assistance during the lean season. While the control group knew they would receive transfers, they were unaware when these would start and we find no evidence of borrowing or other anticipatory changes in behavior. Our late group, thus, provides a valid counterfactual. Third, because we could not visit our participants, and to avoid survey fatigue over the phone, we focused on a small 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.

Nonetheless, this exercise provides proof-of-concept that digital aid represents a potential cost-effective complement to existing modalities. We see five important directions for future research to help humanitarians select strategies to address hunger. First, relative to physical goods and cash, digital payments enable humanitarians to more easily vary when aid is delivered, its frequency, and the amount. Research can help identify how to optimally take advantage of this added flexibility. Second, more evidence is needed on the costs and benefits of digital relative to cash delivery. Some prior work addresses this question ([Aker et al., 2016](#)), but we think more of this work in fragile contexts controlled by hostile actors and where the vulnerable are hard-to-reach will be relevant given the rise in hunger among such groups. This is best pursued in experiments which randomize both digital delivery and cash delivery against a control group. Third, humanitarians need further evidence on how key enabling conditions – such as mobile phone penetration and the availability of merchants who accept digital payments – affect these relative costs and benefits, and how best to organize future hybrid partnerships based on differences in enabling conditions. Fourth, additional focus is needed on approaches to targeting, onboarding and monitoring beneficiaries in hostile environments that minimize risks to participants ([Aiken et al., 2022](#); [Jeong and Trako, 2022](#)). In particular, the one point of physical contact in our study between humanitarians and beneficiaries was during onboarding. Further innovations might eliminate the need for contact between humanitarians and those in need altogether and apply automated screening of data from digital transactions and phone surveys to monitor user experience and identify potential fraudulent patterns. And fifth, implementing digital payments at scale may affect diversion strategies, market prices, and the broader use of digital payments technologies for other financial transactions. It may also spur the transformation of humanitarian organizations into leaner organizational models to best complement the capacities of private and community organizations. Future work with large-scale experiments should explore such changes.

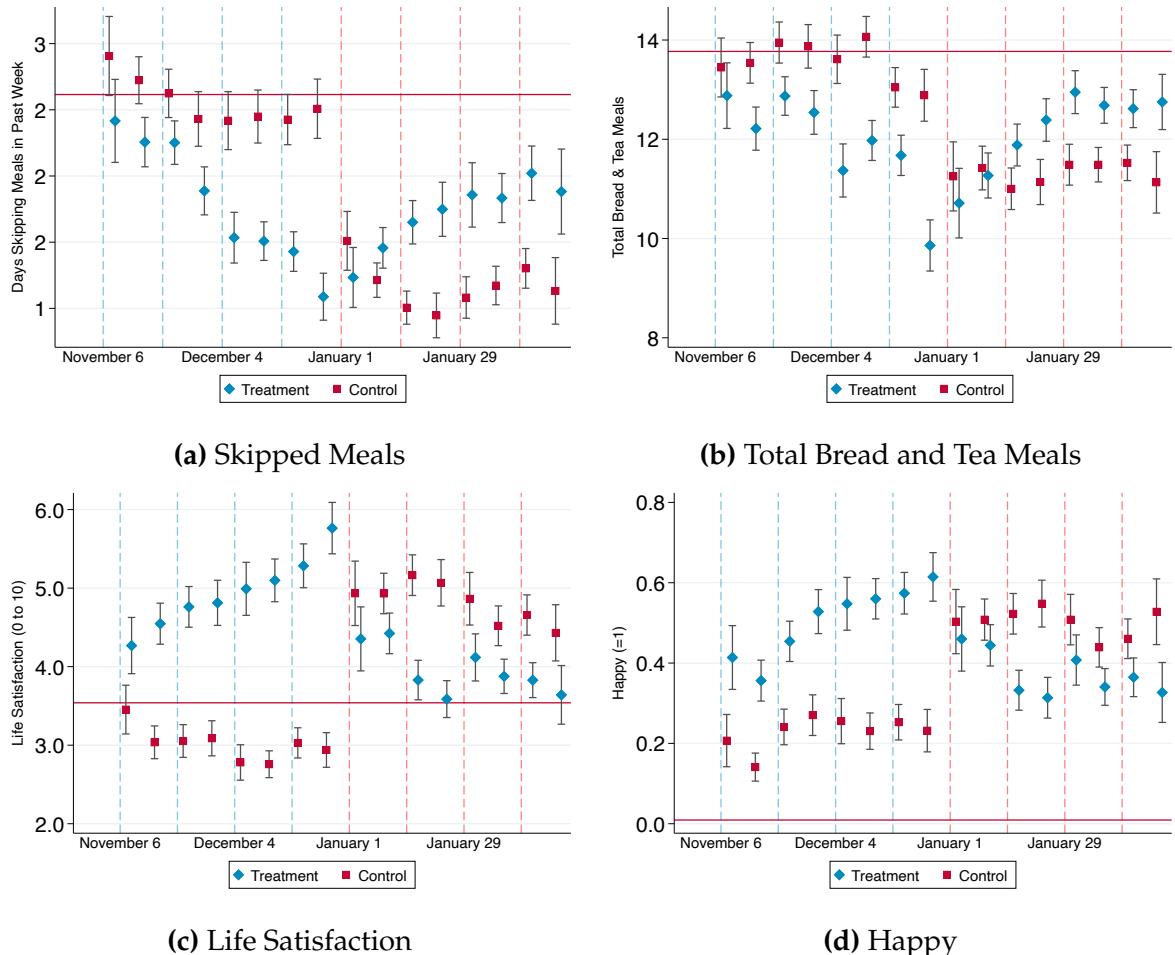
Encouragingly, aid agencies are increasing their use of digital payments. The WFP is using insights from our study to scale digital payments to over 100,000 house-

holds in Afghanistan in partnership with HesabPay, which has been able to maintain its license to operate despite the dynamic strategic and political conditions in the country. As this program scales, it will need to grapple with the complex interdependencies between humanitarian agencies, digital payment platforms, and governments in fragile states. While local authorities could still choose to block digital aid provision, doing so might impede future humanitarian commitments and draw increased public attention to the issues of aid diversion and vulnerable populations. The centralized management of digital payments relative to cash or food distribution also reduces the need to engage with local actors who might complicate aid delivery and increases humanitarian bargaining power. 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 payment platforms), as shown in [A9](#), where Afghanistan ranks as a particularly challenging case, suggests that this approach could increase food security in other fragile states with high levels of hunger.

## Tables and figures

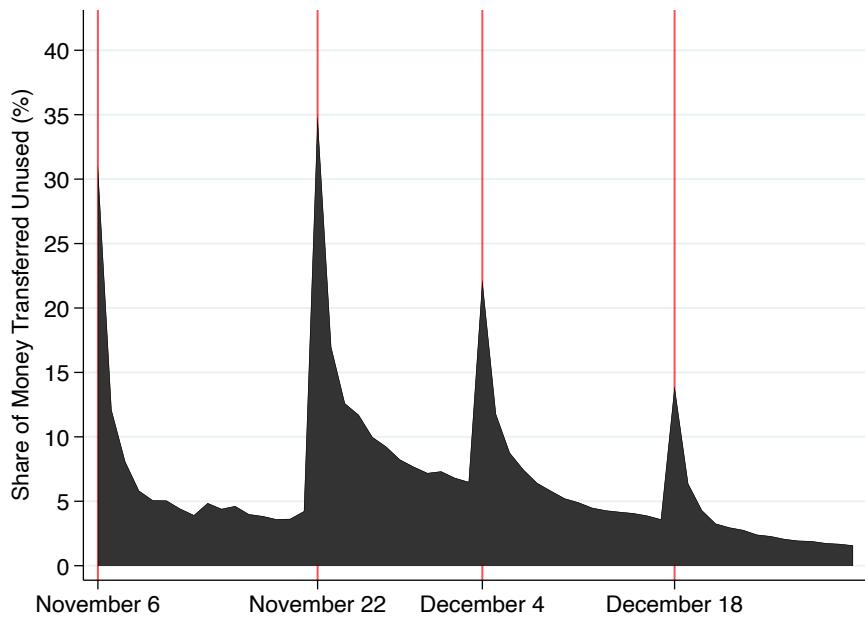


**Figure 1:** Intervention Timeline and Design

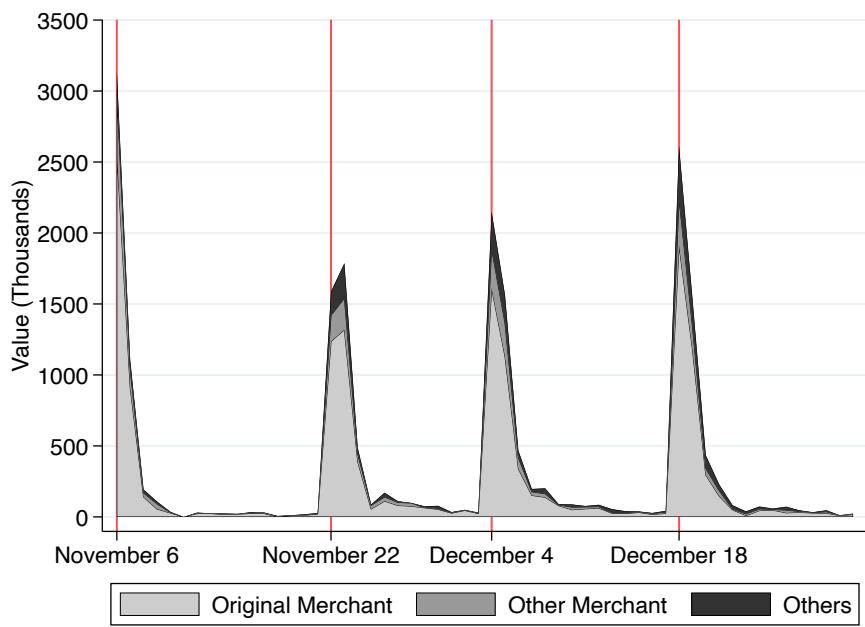


**Figure 2: Effects over Entire Duration of Program**

**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 control group (treatment group) respondents divided in two-week bins since the start of the payments. Blue (red) vertical lines represent the dates in which treatment group (control group) participants received their payments. Note that this is the only analysis in which we are using/showing data for the period after the control 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 control group starts receiving payments). Bars show 95% confidence intervals.



(a) Share of Money Unused



(b) Transactions Flow

**Figure 3: Account Usage Over Time**

**Notes:** Panel A shows the cumulative share of money that participants in the treated group 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 belongs to either a different test merchant or another merchant that did not participate in the onboarding sessions.

**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)
<b>Panel A. Primary Outcomes</b>							
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
<b>Panel B. Secondary Outcomes</b>							
Days eating rice (past week)	0.698	1.058	0.597	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	0.934	0.453	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	0.001	0.002	4763
Able to buy medicine	0.051	0.221	0.034	0.01	0.001	0.002	3582
Involved in fin. decisions	0.666	0.472	0.017	0.015	0.272	0.185	4757
Total household income (past month)	876.683	1581.531	128.482	90.93	0.158	0.118	4763
Household's head employed (past month)	0.192	0.394	-0.008	0.014	0.578	0.311	4741

*Notes:* Stratification fixed effects, survey round fixed effects, and baseline values of dependent variables, if available, are included. Standard errors are clustered at the 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 – Food Security & Wellbeing Outcomes**

	(1)	(2)	(3)	(4)	(5)
<b>Panel A. Food Security</b>	Days Skipping Meals	Children Skipping Meals (=1)	Regularly Eat Twice a Day (=1)	Total Bread & Tea Meals	KLK Index
$\beta_1$ : Treated $\times$ Round 2	-0.580*** (0.094)	-0.096*** (0.021)	-0.005 (0.028)	-0.948*** (0.219)	0.309*** (0.056)
$\beta_2$ : Round 2	-0.195*** (0.072)	0.025* (0.013)	0.072*** (0.020)	-0.322** (0.151)	0.096** (0.038)
$\beta_3$ : Treated	-0.470*** (0.072)	-0.069*** (0.015)	0.096*** (0.020)	-1.134*** (0.162)	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
<b>Panel B. Wellbeing</b>	Better Economic Situation	Satisfied with Financial Sit.	Happy	Life Satisfaction	KLK Index
$\beta_1$ : Treated $\times$ Round 2	0.219*** (0.021)	0.200*** (0.024)	0.102*** (0.026)	0.842*** (0.127)	0.837*** (0.078)
$\beta_2$ : Round 2	-0.012 (0.009)	-0.036*** (0.014)	0.030* (0.017)	-0.223*** (0.075)	-0.073* (0.040)
$\beta_3$ : Treated	0.226*** (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

*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.

**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)
<b>Panel A. Baseline</b>						
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	0.002	0.002	0.001	0.007	0.012	0.009
<b>Panel B. Long-Run</b>						
$\beta_1$ : Treated $\times$ Round 2	-0.000 (0.003)	0.006* (0.003)	0.002 (0.003)	-0.001 (0.005)	0.091 (0.084)	0.004 (0.006)
$\beta_2$ : Round 2	-0.002 (0.002)	-0.003 (0.002)	-0.003* (0.002)	-0.008** (0.003)	-0.137*** (0.045)	-0.010** (0.004)
$\beta_3$ : Treated	0.001 (0.005)	-0.007 (0.005)	-0.001 (0.005)	0.003 (0.009)	-0.062 (0.140)	-0.004 (0.011)
$(\beta_1 + \beta_2 + \beta_3) / \beta_3$	-0.770	0.500	1.880	-1.680	1.740	2.550
Observations	4,558	4,596	4,631	4,626	4,648	4,509
<b>Panel C. List Experiment</b>						
	All Sample	Late Sample	Early Sample			
Treated List	0.011 (0.023)	0.019 (0.032)	0.008 (0.032)			
Observations	2,358	1,171	1,187			
Mean Items Mentioned	0.746	0.691	0.800			

*Notes:* In Panels A and B, stratification fixed effects, survey round fixed effects, and baseline values of dependent variables, if available, are included. 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". This includes a control for surveyor fixed effects. Standard errors are clustered at the individual level.

**Table 4:** Actual Values vs. Experts' Predictions

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

## References

Active Learning Network for Accountability and Performance in Humanitarian Action. 2023. “ALNAP: Strengthening Humanitarian Action through Evaluation and Learning.” .

Agur, Itai, Soledad Martinez Peria, and Celine Rochon. 2020. “Digital Financial Services and the Pandemic: Opportunities and Risks for Emerging and Developing Economies.”

Aiken, Emily, Suzanne Bellue, Dean Karlan, Chris Udry, and Joshua E Blumenstock. 2022. “Machine learning and phone data can improve targeting of humanitarian aid.” *Nature* 603 (7903):864–870.

Aker, Jenny C. 2017. “Comparing cash and voucher transfers in a humanitarian context: Evidence from the Democratic Republic of Congo.” *World Bank Economic Review* 31:44–70.

Aker, Jenny C, Rachid Boumnijel, Amanda McClelland, and Niall Tierney. 2016. “Payment Mechanisms and Antipoverty Programs: Evidence from a Mobile Money Cash Transfer Experiment in Niger.” *Economic Development and Cultural Change* .

Alik-Lagrange, Arthur, Niklas Buehren, Markus Goldstein, and Johannes Hoogeveen. 2019. “The Welfare Impacts of Public Works in Fragile Conflict and Violent Economies: The Londö Public Works in the Central African Republic \*.” URL <http://hdr.undp.org/en/composite/HDI>.

American Political Science Association. 2020. “Principles and Guidance for Human Subjects Research.”

Anderson, Mary B. 1999. *Do no harm: how aid can support peace—or war*. Lynne Rienner Publishers.

Anderson, Michael L. 2008. “Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects.” *Journal of the American statistical Association* 103 (484):1481–1495.

Andersson-Manjang, S.K. 2021. “The Mobile Money Prevalence Index (MMPI) A Country-Level Indicator for Assessing the Adoption, Activity and Accessibility of Mobile Money.” GSMA URL [www.gsma.com/mobilemoney](http://www.gsma.com/mobilemoney).

Angelucci, Manuela. 2008. “Love on the Rocks: Domestic Violence and Alcohol Abuse in Rural Mexico.” *BE Journal of Economic Analysis & Policy* 8.

Asia Foundation. 2019. “A Survey of the Afghan People.” *Asia Foundation Report* .

Attah, Ramlatu, Valentina Barca, Andrew Kardan, Ian MacAuslan, Fred Merttens, and Luca Pellerano. 2016. “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.

Baele, Stéphane J. 2013. "The ethics of New Development Economics: is the Experimental Approach to Development Economics morally wrong?" *The Journal of Philosophical Economics* 7.

Bailey, Sarah and Paul Harvey. 2015. "State of evidence on humanitarian cash transfers Background Note for the High Level Panel on Humanitarian Cash Transfers."

Ballesteros, Luis, Michael Useem, and Tyler Wry. 2017. "Masters of disasters? An empirical analysis of how societies benefit from corporate disaster aid." *Academy of management journal* 60 (5):1682–1708.

Banerjee, Abhijit, Michael Faye, Alan Krueger, Paul Niehaus, and Tavneet Suri. 2020. "Effects of a Universal Basic Income during the pandemic." *Innovations for Poverty Action Working Paper* .

Baranov, Victoria, Lisa Cameron, Diana Contreras Suarez, and Claire Thibout. 2021. "Theoretical Underpinnings and Meta-analysis of the Effects of Cash Transfers on Intimate Partner Violence in Low- and Middle-Income Countries."

Barnett, Michael. 2013. *Empire of humanity: A history of humanitarianism*. Cornell University Press.

Bastagli, Francesca, Jessica Hagen-Zanker, Luke Harman, Valentina Barca, Georgina Sturge, and Tanja Schmidt. 2019. "The Impact of Cash Transfers: A review of the evidence from low- and middle-income countries."

Beath, Andrew, Fotini Christia, Georgy Egorov, and Ruben Enikolopov. 2016. "Electoral rules and political selection: Theory and evidence from a field experiment in Afghanistan." *The Review of Economic Studies* 83 (3):932–968.

Beath, Andrew, Fotini Christia, and Ruben Enikolopov. 2013. "Empowering women through development aid: Evidence from a field experiment in Afghanistan." *American Political Science Review* 107 (3):540–557.

———. 2017a. "Can development programs counter insurgencies?: Evidence from a field experiment in Afghanistan." .

———. 2017b. "Direct democracy and resource allocation: Experimental evidence from Afghanistan." *Journal of Development Economics* 124:199–213.

———. 2018. "Do elected councils improve governance? Experimental evidence on local institutions in Afghanistan." .

Bedoya, Guadalupe, Aidan Coville, Johannes Haushofer, Mohammad Isaqzadeh, Jeremy P Shapiro, and Woodrow Wilson School. 2019. "No Household

Left Behind: Afghanistan Targeting the Ultra Poor Impact Evaluation." *National Bureau of Economic Research* URL <http://www.worldbank.org/en/topic/fragilityconflictviolence/overview>.

Bedoya Arguelles, Guadalupe, Aidan Coville, Johannes Haushofer, Mohammad Isaqzadeh, and Jeremy Shapiro. 2019. "No Household Left Behind: Afghanistan Targeting the Ultra Poor Impact Evaluation." *NBER Working Paper* (w25981).

Bete, George. 2022. "Afghanistan Cash & Voucher Working Group (CVWG) Minimum Expenditure Basket (MEB) and Setting the Transfer Value: Guidance Document."

Binesh, Banafsha. 2023. "Sopko Speaks of US Legislation Restricting Aid to Afghanistan." *Tolo News* .

Blair, Graeme, Alexander Coppock, and Margaret Moor. 2020. "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.

Blattman, Christopher, Eric P. Green, Julian Jamison, M. Christian Lehmann, and Jeanne Annan. 2016. "The returns to microenterprise support among the ultrapoor: A field experiment in postwar Uganda." *American Economic Journal: Applied Economics* 8:35–64.

Blofield, Merike, Felicia M. Knaul, Renzo Calderón-Anyosa, Amber Peterman, Julianne Martinez Franzoni, Megan O'Donnell, and Flavia Bustreo. 2022. "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."

Blumenstock, Joshua, Michael Callen, and Tarek Ghani. 2018. "Why do defaults affect behavior? Experimental evidence from Afghanistan." *American Economic Review* 108 (10):2868–2901.

Blumenstock, Joshua, Michael Callen, Tarek Ghani, and Robert Gonzalez. Forthcoming. "Violence and financial decisions: Evidence from mobile money in Afghanistan." *Review of Economics and Statistics* :1–45.

Blumenstock, Joshua E, Michael Callen, Tarek Ghani, and Lucas Koepke. 2015. "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*. 1–10.

Blumenstock, Joshua Evan, Michael Callen, Anastasiia Faikina, Stefano Fiorin, and Tarek Ghani. 2023. "Strengthening Fragile States: Evidence from Mobile Salary Payments in Afghanistan." Available at SSRN 4473387 .

Bobonis, Gustavo J. 2011. "The Impact of Conditional Cash Transfers on Marriage and Divorce." *Economic Development and Cultural Change* :281–312.

Bobonis, Gustavo J., Melissa González-Brenes, and Roberto Castro. 2013. "Public transfers and domestic violence: The roles of private information and spousal control." *American Economic Journal: Economic Policy* 5:179–205.

Bonilla, Juan, Kaitlin Carson, Gilbert Kiggundu, Mitchell Morey, Hannah Ring, Eleonora Nillesen, Gabriele Erba, and Steven Michel. 2017. "Humanitarian Cash Transfers in the Democratic Republic of the Congo: Evidence from UNICEF's ARCC II Programme." URL [www.air.org](http://www.air.org).

Bossuroy, Thomas, Markus Goldstein, Bassirou Karimou, Dean Karlan, Harounan Kazianga, William Parienté, Patrick Premand, Catherine C. Thomas, Christopher Udry, Julia Vaillant, and Kelsey A. Wright. 2022. "Tackling psychosocial and capital constraints to alleviate poverty." *Nature* 605:291–297.

Boudreau, Laura E, Sylvain Chassang, Ada Gonzalez-Torres, and Rachel M Heath. 2023. "Monitoring Harassment in Organizations." *National Bureau of Economic Research Working Paper* .

Broadcasting Board of Governors. 2015. "Media Use in Afghanistan." *Broadcasting Board of Governors Gallup Report* .

Bryer, David and Edmund Cairns. 1997. "For Better? For Worse? Humanitarian Aid in Conflict." *Development in Practice* :363–374.

Buller, Ana Maria, Amber Peterman, Meghna Ranganathan, Alexandra Bleile, Melissa Hidrobo, and Lori Heise. 2018. "A mixed-method review of cash transfers and intimate partner violence in lowand middle-income countries." *World Bank Research Observer* 33:218–258.

Burde, Dana, Joel Middleton, Roxanne Rahnama, and Cyrus Samii. 2023. "Can Communities Take Charge? A Randomized Controlled Trial on Sustaining Schools in Afghanistan."

Burton, Jo. 2020. "'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.

Byrd, William. 2023. "Afghanistan's Crisis Requires a Coherent, Coordinated International Response." *United States Institute of Peace* .

Campbell, Susanna. 2019. "Ethics of Research in Conflict Environments." *Journal of Global Security Studies* 2.

Carril, Alvaro. 2017. "Dealing with misfits in random treatment assignment." *The Stata Journal* 17 (3):652–667.

CBS News. 2023. "U.S. taxpayers helping fund Afghanistan's Taliban? Aid workers say they're forced "to serve the Taliban first"." *CBS News* .

Chotiner, Isaac. 2024. "Gaza Is Starving." *The New Yorker* 1.

Chowdhury, Reajul, Elliott Collins, Ethan Ligon, and Munshi Sulaiman. 2017. "Valuing Assets Provided to Low-Income Households in South Sudan."

Clarke, Paul Knox and James Darcy. 2014. "Insufficient evidence: The quality and use of evidence in humanitarian action." URL [www.alnap.org/pool/files/discussion-starter-evidence-alnap-2014.pdf](http://www.alnap.org/pool/files/discussion-starter-evidence-alnap-2014.pdf).

Cliffe, Sarah, Renata Dwan, Betty Wainaina, and Leah Zamore. 2023. "Aid strategies in 'politically estranged' settings." *Chatham House Research Paper*.

Crost, Benjamin, Joseph H. Felter, and Patrick B. Johnston. 2016. "Conditional cash transfers, civil conflict and insurgent influence: Experimental evidence from the Philippines." *Journal of Development Economics* 118:171–182.

Davies, Shawn, Therése Pettersson, and Magnus Öberg. 2023. "Organized violence 1989–2022, and the return of conflict between states." *Journal of Peace Research* URL <http://journals.sagepub.com/doi/10.1177/00223433231185169>.

de Hoop, Jacobus, Mitchell Morey, and David Seidenfeld. 2019. "No Lost Generation: Supporting the School Participation of Displaced Syrian Children in Lebanon." *Journal of Development Studies* 55:107–127.

De Quidt, Jonathan, Johannes Haushofer, and Christopher Roth. 2018. "Measuring and bounding experimenter demand." *American Economic Review* 108 (11):3266–3302.

De Waal, Alex. 2023. "The crisis in Sudan calls for a new model of humanitarian aid." *Financial Times*.

De Waal, Alexander. 1997. *Famine crimes: politics & the disaster relief industry in Africa*. Indiana University Press.

Department for International Development. 2017. "The effects of DFID's cash transfer programmes on poverty and vulnerability: An impact review." URL [www.nationalarchives.gov.uk/doc/open-government-licence/version/3](http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3).

Dodgson, Mark, David Gann, Irving Wladawsky-Berger, Naveed Sultan, and Gerard George. 2015. "Managing digital money." *Academy of management journal* 58 (2):325–333.

Doocy, Shannon and Hannah Tappis. 2016. "Systematic Review 28 Cash-based approaches in humanitarian emergencies A systematic review."

Dorobantu, Sinziana, Aseem Kaul, and Bennet Zelner. 2017. "Nonmarket strategy research through the lens of new institutional economics: An integrative review and future directions." *Strategic Management Journal* 38 (1):114–140.

Dreze, Jean and Amartya Sen. 1990. *Hunger and public action*. Clarendon Press.

Duflo, Esther, Abhijit Banerjee, Amy Finkelstein, Lawrence Katz, Benjamin Olken, and Anja Sautmann. 2020. "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* .

Ecker, Olivier, Jean-François Maystadt, and Zhe Guo. 2019. "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* .

Fangwa, Anicet A, Caroline Flammer, Marieke Huysentruyt, and Bertrand V Quélin. 2024. "The governance of nonprofits and their social impact: Evidence from a randomized program in healthcare in the Democratic Republic of Congo." *Management Science* 70 (5):2732–2755.

Field, Erica, Rohini Pande, Natalia Rigol, Simone Schaner, and Charity Troyer Moore. 2021. "On her own account: How strengthening womenâ€™s financial control impacts labor supply and gender norms." *American Economic Review* 111 (7):2342–2375.

Findley, Michael, Joseph K Young, Daniel Strandow, and Olgahan Cat. 2023. "Aiding War: Foreign Aid and the Intensity of Violent Armed Conflict." *International Studies Quarterly* 67 (3).

Findley, Michael G. 2018. "Does foreign aid build peace?" *Annual Review of Political Science* 21:359–384.

Food Security Information Network. 2023. "2023 Global Report on Food Crises." .

Foreign and Commonwealth Development Office (FCDO). 2023. "International development in a contested world: ending extreme poverty and tackling climate change." .

Foreign, Commonwealth and Development Office. 2023. "International development in a contested world: ending extreme poverty and tackling climate change: A White Paper on International Development." URL <https://assets.publishing.service.gov.uk/media/6560874b0c7ec8000d95bdcf/international-development-in-a-contested-world-ending-extreme-poverty-and-tackling-climate-change.pdf>.

Freedom House. 2023. "Freedom in the World 2023: Marking 50 Years in the Struggle for Democracy." *Freedom House Report* URL [www.freedomhouse.org](http://www.freedomhouse.org).

Gallup. 2022a. "Afghans Lose Hope Under the Taliban." .

———. 2022b. "Suffering Reaches Unprecedented Levels in Afghanistan." .

———. 2022c. "Taliban Plunge Afghans Into Economic Ruin." .

———. 2023. "Digital Freedom Out of Reach for Most Afghan Women." .

Gentilini, Ugo. 2016. "Revisiting the "cash versus food" debate: new evidence for an old puzzle?" *The World Bank Research Observer* 31 (1):135–167.

———. 2022. “Cash Transfers in Pandemic Times: Evidence, Practices, and Implications from the Largest Scale Up in History.” *World Bank Report* .

George, Gerard, Thomas J. Fewer, Sergio Lazzarini, Anita M. McGahan, and Phanish Puranam. 2024. “Partnering for Grand Challenges: A Review of Organizational Design Considerations in Publicâ€“Private Collaborations.” *Journal of Management* 50 (1):10–40. URL <https://doi.org/10.1177/01492063221148992>.

George, Gerard, Jennifer Howard-Grenville, Aparna Joshi, and Laszlo Tihanyi. 2016. “Understanding and tackling societal grand challenges through management research.” *Academy of management journal* 59 (6):1880–1895.

George, Gerard, Ryan K Merrill, and Simon JD Schillebeeckx. 2021. “Digital sustainability and entrepreneurship: How digital innovations are helping tackle climate change and sustainable development.” *Entrepreneurship theory and practice* 45 (5):999–1027.

Ghani, Tarek and Grant Gordon. 2022. “Predictable Disasters.” *Breakthrough: The Promise of Frontier Technologies for Sustainable Development* :103.

GiveDirectly. 2024. “Cash Evidence Explorer.” *Vox* .

Goodhand, Jonathan. 2002. “Aiding violence or building peace? The role of international aid in Afghanistan.” *Third World Quarterly* :837–859.

Gordon, Laura. 2015. “Risk and humanitarian cash transfer programming: Background Note for the High Level Panel on Humanitarian Cash Transfers.”

Grossman, Guy, Macartan Humphreys, and Gabriella Sacramone-Lutz. 2014. “I wld like u WMP to extend electricity 2 our village: On information technology and interest articulation.”

Guggenheim, Scott and Charles Petrie. 2022. “Alternative Aid Modalities: Community Development.”

Hagen-Zanker, Jessica, Martina Ulrichs, Rebecca Holmes, and Zina Nimeh. 2017. “Cash transfers for refugees The economic and social effects of a programme in Jordan.” URL [www.odi.org/twitter](http://www.odi.org/twitter).

Haushofer, Johannes and Jeremy Shapiro. 2016. “The short-term impact of unconditional cash transfers to the poor: experimental evidence from Kenya.” *The Quarterly Journal of Economics* 131 (4):1973–2042.

Hidrobo, Melissa and Lia Fernald. 2013. “Cash transfers and domestic violence.” *Journal of Health Economics* 32:304–319.

Hidrobo, Melissa, John Hoddinott, Amber Peterman, Amy Margolies, and Vanessa Moreira. 2014. “Cash, food, or vouchers? Evidence from a randomized experiment in northern Ecuador.” *Journal of Development Economics* 107:144–156.

Hidrobo, Melissa, Amber Peterman, and Lori Heise. 2016. "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.

Hoddinott, John, Susanna Sandström, and Joanna Upton. 2018. "The impact of cash and food transfers: Evidence from a randomized intervention in Niger." *American Journal of Agricultural Economics* 100:1032–1049.

Humphreys, Macartan. 2015. "Reflections on the Ethics of Social Experimentation." *Journal of Globalization and Development* 6:87–112.

International Rescue Committee. 2016. "Cost Efficiency Analysis: Unconditional Cash Transfer Programs."

———. 2018. "Seven steps to scaling cash relief Driving outcomes and efficiency International Rescue Committee |."

Jeong, Dahyeon and Iva Trako. 2022. "Cash and In-Kind Transfers in Humanitarian Settings A Review of Evidence and Knowledge Gaps." *World Food Programme and World Bank* .

Kandpal, Eeshani, Harold Alderman, Jed Friedman, Deon Filmer, Junko Onishi, and Jorge Avalos. 2016. "A conditional cash transfer program in the philippines reduces severe stunting." *Journal of Nutrition* 146:1793–1800.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz. 2007. "Experimental analysis of neighborhood effects." *Econometrica* 75 (1):83–119.

Kurdi, Sikandra. 2021. "The nutritional benefits of cash transfers in humanitarian crises: evidence from Yemen." *World Development* 148.

Kurtz, Jon, Vaidehi Krishnan, Virginia Leape, Vladimir Jovanovic, Marlaina Ross, and Reimar Macaranas. 2021. "A Million-Dinar Question: Can Cash Transfers Drive Economic Recovery in Conflict-Driven Crises?"

Kurtzer, Jacob D. 2019. "Denial, Delay, Diversion: Tackling Access Challenges in an Evolving Humanitarian Landscape." *Center for Strategic and International Studies URL* [www.csis.org](http://www.csis.org).

Langendorf, Céline, Thomas Roederer, Saskia de Pee, Denise Brown, Stéphane Doyon, Abdoul Aziz Mamaty, Lynda W.M. Touré, Mahamane L. Manzo, and Rebecca F. Grais. 2014. "Preventing Acute Malnutrition among Young Children in Crises: A Prospective Intervention Study in Niger." *PLoS Medicine* 11.

Lee, David S. 2009. "Training, wages, and sample selection: Estimating sharp bounds on treatment effects." *The Review of Economic Studies* 76 (3):1071–1102.

Lehmann, M. Christian and Daniel T.R. Masterson. 2020. "Does Aid Reduce Anti-refugee Violence? Evidence from Syrian Refugees in Lebanon." *American Political Science Review* 114:1335–1342.

Lischer, Sarah Kenyon. 2003. "Collateral damage: Humanitarian assistance as a cause of conflict." *International Security* 28 (1):79–109.

Luo, Jiao and Aseem Kaul. 2019. "Private action in public interest: The comparative governance of social issues." *Strategic Management Journal* 40 (4):476–502.

Lyall, Jason, Yang Yang Zhou, and Kosuke Imai. 2020. "Can Economic Assistance Shape Combatant Support in Wartime? Experimental Evidence from Afghanistan." *American Political Science Review* 114:126–143.

Maren, Michael. 2009. *The road to hell*. Simon and Schuster.

Maunder, Nick, Neil Dillon, Gabrielle Smith, Sharon Truelove, and Victoria De Bauw. 2016. "Evaluation of the Use of Different Transfer Modalities in ECHO Humanitarian Aid Actions 2011–2014."

McDermott, Rose and Peter K. Hatemi. 2020. "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.

McDonough, Siobhan. 2022. "Giving people cash is usually better than shipping them food." *Vox* .

McGahan, Anita M and Leandro S Pongeluppe. 2023. "There is no planet B: Aligning stakeholder interests to preserve the Amazon rainforest." *Management Science* 69 (12):7860–7881.

McIntosh, Craig and Andrew Zeitlin. 2022. "Using household grants to benchmark the cost effectiveness of a USAID workforce readiness program." *Journal of Development Economics* 157:102875.

MercyCorps. 2022. "Beyond Meeting Immediate Needs: The Impact of Electronic Cash Transfer Approaches on Disaster Recovery and Financial Inclusion."

Mikulak, Magdalena. 2018. "Cost-effectiveness in humanitarian work: cash-based programming."

Moussa, Wael, Alexandra Irani, Nisreen Salti, Rima Al Mokdad, Zeina Jamaluddine, Jad Chaaban, and Hala Ghattas. 2022. "The impact of cash transfers on Syrian refugee children in Lebanon." *World Development* 150.

Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar. 2016. "Building state capacity: Evidence from biometric smartcards in India." *American Economic Review* 106 (10):2895–2929.

Muralidharan, Karthik, Paul Niehaus, Sandip Sukhtankar, and Jeffrey Weaver. 2021. "Improving Last-Mile Service Delivery Using Phone-Based Monitoring." *American Economic Journal: Applied Economics* 13:52–82.

Narang, Neil. 2015. "Assisting uncertainty: how humanitarian aid can inadvertently prolong civil war." *International Studies Quarterly* 59 (1):184–195.

Nunn, Nathan and Nancy Qian. 2014. "US food aid and civil conflict." *American economic review* 104 (6):1630–1666.

O'Donnell, Lynne. 2022. "The Taliban Are Abusing Western Aid." *Foreign Policy* .

———. 2023. "The Taliban Have "Infiltrated" U.N. Deliveries of Aid." *Foreign Policy* .

OECD. 2022. "States of Fragility 2022." URL [https://www.oecd-ilibrary.org/development/states-of-fragility-2022\\_c7fedf5e-en](https://www.oecd-ilibrary.org/development/states-of-fragility-2022_c7fedf5e-en).

Olken, Benjamin A. 2015. "Promises and Perils of Pre-Analysis Plans." *Journal of Economic Perspectives* 29 (3):61 – 80.

Organization for Economic Cooperation and Development (OECD). 2022. *States of Fragility 2022*. URL <https://www.oecd-ilibrary.org/content/publication/c7fedf5e-en>.

Overseas Development Institute. 2015. "Doing cash differently: How cash transfers can transform humanitarian aid." URL [www.odi.org/twitter](http://www.odi.org/twitter).

Pazarbasioglu, Ceyla, Alfonso Garcia Mora, Mahesh Uttamchandani, Harish Natarajan, Erik Feyen, and Mathew Saal. 2020. "Digital Financial Services."

Peachey, Karen. 2020. "Home Blog The 90Bn Question: Can we reach 700 million people in response to COVID-19?"

Perlez, Jane. 1992. "Theft of Food Aid Is a Business in Starving Somalia."." *New York Times* 4.

Politico. 2023. "Ukraine uncovers corruption scheme implicating top officials." *Politico* .

Polman, Linda. 2010. *The crisis caravan: what's wrong with humanitarian aid?* Metropolitan Books.

Pople, Ashley, Ruth Hill, Stefan Dercon, and Ben Brunckhorst. 2021. "Anticipatory Cash Transfers in Climate Disaster Response \* CSAE Working Paper WPS/2021-07." URL <https://doi.org/10.1257/rct.6576-1.1>.

Premand, Patrick and Oumar Barry. 2022. "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.

Puri, Jyotsna, Anastasia Aladysheva, Vegard Iversen, Yashodhan Ghorpade, and Tilman Brück. 2017. “Can rigorous impact evaluations improve humanitarian assistance?” *Journal of Development Effectiveness* 9:519–542.

Quattrochi, John, Ghislain Bisimwa, Peter van der Windt, and Maarten Voors. 2022. “Cash-like vouchers improve psychological well-being of vulnerable and displaced persons fleeing armed conflict.” *PNAS Nexus* 1.

Reinikka, Ritva and Jakob Svensson. 2006. “Using Micro-Surveys to Measure and Explain Corruption.” *World Development* 34 (2):359–370.

Riegg Cellini, Stephanie and James Edwin Kee. 2015. “Cost-effectiveness and cost-benefit analysis.” *Handbook of practical program evaluation* :636–672.

Riley, Emma. 2022. “Resisting social pressure in the household using mobile money: Experimental evidence on microenterprise investment in Uganda.” .

Roberts, Leslie. 2023. “Taliban decree deepens Afghanistan’s health crisis.” *Science (New York, NY)* 379 (6629):222–223.

Salti, Nisreen, Jad Chaaban, Wael Moussa, Alexandra Irani, Rima Al Mokdad, Zeina Jamaluddine, and Hala Ghattas. 2022. “The impact of cash transfers on Syrian refugees in Lebanon: Evidence from a multidimensional regression discontinuity design.” *Journal of Development Economics* 155.

Samii, Cyrus. 2023. “WIDER Working Paper 2023/26-Revisiting community-driven reconstruction in fragile states.” URL <https://doi.org/10.35188/UNU-WIDER/2023/334-5>.

Sandström, Susanna and Levan Tchatchua. 2010. “Do cash transfers improve food security in emergencies? Evidence from Sri Lanka.”

Schwab, Benjamin. 2020. “In the Form of Bread? A Randomized Comparison of Cash and Food Transfers in Yemen.” *American Journal of Agricultural Economics* 102:91–113.

Schwab, Benjamin, Amy Margolies, and John Hoddinott. 2013. “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* .

Sen, Amartya. 1999. “Development as freedom.” *The globalization and development reader: Perspectives on development and global change* 525.

Senate Armed Services Committee. 2023. “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* .

Shleifer, Andrei and Robert W Vishny. 1993. “Corruption.” *The quarterly journal of economics* 108 (3):599–617.

Smith, Gabrielle, Ruth McCormack, Alex Jacobs, Arushi Chopra, Aarsh Vir Gupta, and Thomas Abell. 2018. "The State of the World's Cash Report: Cash Transfer Programming in Humanitarian Aid."

Sopko, John F. 2023. "Examining Afghanistan Since the U.S. Withdrawal." *Testimony to the House Committee on Foreign Affairs, 117th Cong.*

Sossouvi, Kokoévi. 2013. "E-transfers in emergencies: Implementation support guidelines."

Special Inspector General for Afghanistan Reconstruction. 2023a. "2023 high-risk list." *Special Inspector General for Afghanistan Reconstruction Report*.

———. 2023b. "Quarterly Report to the United States Congress - October 30, 2023." .

Suri, Tavneet, Jenny Aker, Catia Batista, Michael Callen, Tarek Ghani, William Jack, Leora Klapper, Emma Riley, Simone Schaner, and Sandip Sukhtankar. 2023. "Mobile money." *VoxDevLit* 2 (2):3.

Suri, Tavneet and William Jack. 2016. "The long-run poverty and gender impacts of mobile money." *Science* 354 (6317):1288–1292.

Tappis, Hannah and Shannon Doocy. 2018. "The effectiveness and value for money of cash-based humanitarian assistance: a systematic review." *Journal of Development Effectiveness* 10:121–144.

Townsend, Robert, Dorte Verner, Abimbola Adubi, Jean Saint-Geours, Izabela Leao, Armine Juergenliemk, Tim Robertson, Melissa Williams, Flore de Preneuf, Marketa Jonasova et al. 2021. "Future of Food: Building Stronger Food Systems in Fragility, Conflict, and Violence Settings." *World Bank Report* .

UN International Telecommunication Union. 2022. "Measuring digital development Facts and Figures 2022." *UN International Telecommunication Union Report* .

UN News. 2023. "UN suspends food assistance as Ethiopia wrestles with aid diversions." *UN News* .

United Nations. 2021. "UN Principles for Responsible Digital Payments."

———. 2022. "WFP Financial Inclusion through Mobile Money."

———. 2023a. "Afghanistan: Security Council condemns Taliban's ban on women working for UN." *UN News* .

———. 2023b. "Security Council Emphasizes That Punitive Restrictions on Women's Rights, Escalating Hunger, Insecurity Taking Devastating Toll in Afghanistan." *UN News* .

———. 2023c. "'The world cannot abandon the people': Top humanitarian official in Afghanistan." *United Nations News* .

United Nations Assistance Mission in Afghanistan. 2023. "Cash shipments to the UN in Afghanistan - Info Sheet." *United Nations Assistance Mission in Afghanistan* .

United Nations Capital Development Fund. 2023. "Interoperability of financial services providers Afghanistan." *UNCDF Report* .

Urquhart, Angus, Fran Girling-Morris, Erica Mason, and Suzanna Nelson-Pollard. 2023. "Global Humanitarian Assistance Report 2023." *Development Initiatives* URL <https://devinit.org/resources/global-humanitarian-assistance-report-2023>.

US Department of Health and Human Services. 1979. "The Belmont Report."

White, Philip, Anthony Hodges, and Matthew Greenslade. 2013. "Guidance on measuring and maximising value for money in social transfer programmes—second edition." Tech. rep., UK Department for International Development.

Williamson, Oliver E. 2000. "The new institutional economics: taking stock, looking ahead." *Journal of economic literature* 38 (3):595–613.

Wolfe, Rebecca. 2020. "The Benefits and Challenges of Randomized Control Trials in Conflict Environments: Reflections From a Scholar-Practitioner." In *Researching peace, conflict, and power in the field: Methodological challenges and opportunities*, edited by Yasemin Gülsüm Acar, Sigrun Marie Moss, and Özden Melis Uluğ, chap. 14. 259–280.

Wood, Elisabeth Jean. 2006. "The ethical challenges of field research in conflict zones." *Qualitative Sociology* 29:373–386.

Wood, Reed M and Emily Molfino. 2016. "Aiding victims, abetting violence: the influence of humanitarian aid on violence patterns during civil conflict." *Journal of global security studies* 1 (3):186–203.

Wood, Reed M and Christopher Sullivan. 2015. "Doing harm by doing good? The negative externalities of humanitarian aid provision during civil conflict." *The Journal of Politics* 77 (3):736–748.

World Bank. 2016. "Cash Transfers in Humanitarian Contexts."

———. 2022. "Cash and In-Kind Transfers in Humanitarian Settings." .

———. 2023a. "Afghanistan's Economy Faces Critical Challenges as it Adjusts to New Realities." .

———. 2023b. "Mobile cellular subscriptions - Afghanistan." *World Bank Database* .

World Food Programme. 2020. "Meeting Needs and Empowering People: WFP's Cash-based Food Assistance in Uganda."

———. 2021a. "Cash-based transfers: Lessons from evaluations."

\_\_\_\_\_. 2021b. "WFP Evidence Summary, Cash-based transfers, lessons from evaluations." .

\_\_\_\_\_. 2023a. "Cash-based transfers and commodity vouchers." *World Food Programme* .

\_\_\_\_\_. 2023b. "Cash-Based Transfers Empower People While Building Resilience." *World Food Programme Report* .

\_\_\_\_\_. 2023c. "Harnessing the Power of Money to Help People Survive and Thrive." *World Food Programme Report* .

\_\_\_\_\_. 2023d. "WFP Afghanistan: Situation Report 13 October 2022." *World Food Programme Afghanistan Report* .

\_\_\_\_\_. 2023e. "World Food Programme Cash policy."

Wormald, Audra, Rajshree Agarwal, Serguey Braguinsky, and Sonali K Shah. 2021. "David overshadows Goliath: Specializing in generality for internationalization in the global mobile money industry." *Strategic Management Journal* 42 (8):1459–1489.

Zürcher, Christoph. 2019. "The folly of aid for stabilisation." *Third World Quarterly* 40 (5):839–854.

# Online Appendix

## A Related work

### 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 (Gentilini, 2022). 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 (Bastagli et al., 2019; GiveDirectly, 2024). 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 (Urquhart et al., 2023). 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 (Puri et al., 2017; Tappis and Doocy, 2018; Jeong and Trako, 2022), the evidence is especially thin in contexts facing the most severe and protracted humanitarian crises (Kurtz et al., 2021).

At the same time that humanitarians are increasing their use of cash-based programming, they are simultaneously shifting from physical to digital distribution (World Food Programme, 2023c, 2020; United Nations, 2022). Although physical cash-based programs are unlikely to disappear entirely, digital is clearly becoming the preferred delivery mechanism, with growing policy support (Overseas Development Institute, 2015; International Rescue Committee, 2018; Smith et al., 2018; Agur, Peria, and Ruchon, 2020). 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.

To understand the prevalence of humanitarian, cash-based programs that have been rigorously evaluated, we examined the most comprehensive repository of systematic evaluations (N=332 as of 2 January 2024) of cash-based programs: GiveDirectly's *Cash Evidence Explorer*. We 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. In our review of the database, we found only four digital cash-based programs in humanitarian contexts.

To further map the evidence base specific to our study, we culled details of relevant cash-based transfer programs from systematic academic reviews ([Bastagli et al., 2019](#); [Tappis and Doocy, 2018](#); [Puri et al., 2017](#); [Gentilini, 2016](#)), policy evidence reviews ([Jeong and Trako, 2022](#); [World Food Programme, 2021a](#); [Mikulak, 2018](#); [Department for International Development, 2017](#); [Maunder et al., 2016](#); [World Bank, 2016](#); [Bailey and Harvey, 2015](#)), evaluation databases and clearinghouses ([GiveDirectly, 2024](#); [Active Learning Network for Accountability and Performance in Humanitarian Action, 2023](#); [Gentilini, 2022](#)), and broader policy discussions ([World Food Programme, 2023c](#); [Urquhart et al., 2023](#); [International Rescue Committee, 2016](#); [Smith et al., 2018](#); [Overseas Development Institute, 2015](#)), which together covered hundreds of cash-based transfer programs. 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, 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).

We focused on conflict- and natural-disaster-based humanitarian crises because rigorous evaluations of cash-based programs during health outbreaks have rarely been conducted ([Jeong and Trako, 2022](#)). 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. 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 ([Gentilini, 2022](#)). Furthermore, although the evidence base is growing, rigorous evaluation appears to be following patterns in programming, overwhelmingly confined to wealthier (middle-income) countries.

In all, we identified 23 relevant cash-based transfer programs summarized in [A10](#). Peer-reviewed academic studies account for 16 of the 23, five are academic working papers, and two are evaluation reports. Column 1 reports the country in which the cash-based intervention took place and the countries are ordered by the current level of fragility of the country in which they occur, moving from the least fragile (top) to most fragile (bottom) ([OECD, 2022](#)). Ecuador and Philippines appear at the top because nei-

ther country is listed in the OECD's list of fragile countries. The next two, Lebanon and Sri Lanka, are not on the core OECD list either, but are flagged as showing key early-warning signs. Sometimes the target beneficiaries (Column 2) 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 (C) and vouchers (V), and the most relevant alternative, food assistance (F). Some of these programs include other/combined treatments, which we note in the extended discussions. Column 4 reports the delivery mechanism, whether some form of digital, physical, or a 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 ([de Hoop, Morey, and Seidenfeld, 2019](#)). 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 ([Kurtz et al., 2021](#)), where social protection systems are also extremely underdeveloped ([Peachey, 2020](#)).

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.

Of the 23 RCTs of humanitarian cash-based transfers, 10 are in Freedom House 'not free' countries, three of those 10 have some digital component, and none, that we can find, target large populations that the government actively oppresses? We coded the Freedom House rating for the first year of the respective program if the program occurred in 2013 or later given that the Freedom House 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 ([Freedom House, 2023](#)).

## 2 Work related to outcomes: Usage, food security, well-being, diversion

To contextualize our experimental findings on usage, food security and mental well-being, [A10](#) reports whether the 23 reviewed studies empirically examined usage at any level, denoting examination with a check mark (✓), the same (or similar) outcome categories that we did, including whether they found positive, null, or negative effects. Few studies carried out systematic tests for diversion, so we only include a check mark (✓) if the study made any observation about levels of diversion in their context. To supplement the summary table, [A11](#) provides details about specific study context and findings.

**Usage** Because tech-illiterate populations are largely poor, rural, elderly, and female ([Peachey, 2020](#)), it is important to assess whether digital transfers will reach those most vulnerable and be used effectively. 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% ([Gentilini, 2022](#)). Of the 10 studies in our review 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. The remaining five relied primarily on ATM cards, which are less demanding technologically, and also confer fewer advantages such as decentralization and transparency. 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 ([Lyall, Zhou, and Imai, 2020](#)). 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 ([Grossman, Humphreys, and Sacramone-Lutz, 2014](#)). (See [A10](#), Column 5).

**Food Security** Most of the studies (21 of 23) considered a food security outcome in some form, although there is considerable variation in the indicators each examines. Of the 21, 16 find that cash-based transfers improved food security outcomes, five found no relationship, and none of the studies found any negative effects. Of the 21, 20 examined at least one indicator of food quantity (e.g., food expenditures or calorie con-

sumption). Of the 21, 18 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 food security outcomes. Existing studies most frequently considered basic needs outcomes. In a recent systematic review, eight of 20 studies considered food security (primarily food consumption) and four of 20 studies examined dietary diversity, most of which demonstrated positive impacts ([Jeong and Trako, 2022](#)). Moreover, five studies examined reductions in negative coping strategies (e.g., meal skipping) and three found positive impacts. (See [A10](#), Column 6).

**Mental well-being** In contrast to food security outcomes, [A10](#) shows that limited attention has been given to the possible mental benefits (or costs) of cash-based programs. Only seven of the 23 studies considered some mental well-being outcomes and six found positive effects. An evaluation of a voucher program in DRC found a positive impact (+0.32sd) on an overall mental 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) ([Quattrochi et al., 2022](#)). 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. Perhaps the strongest impact of cash on mental well-being, with documented short-, medium-, and long-term effects of cash relative to control, occurred in Niger ([Bossuroy et al., 2022](#)). Importantly, like the other treatment arms, cash was accompanied by coaching, savings, training, and market facilitation, making the overall bundle difficult to compare to other cash-only programs. The dearth of studies examining mental well-being mirrors that of prominent systematic reviews ([Jeong and Trako, 2022](#)). They only identified two (of 20) studies reporting subjective well-being results: the DRC voucher study that we just discussed ([Quattrochi et al., 2022](#)) 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 another systematic review, there were no rigorously evaluated studies that addressed mental well-being though some mixed-methods evaluations have found isolated positive impacts ([Attah et al., 2016](#); [Doocy and Tappis, 2016](#); [Hagen-Zanker et al., 2017](#)). (See [A10](#), Column 7).

**Diversion** The extent of diversion in these studies is essentially unknown as four of the 23 studies addressed diversion, but only three attempted to measure its extent. Studies measured the amount of the cash and vouchers actually received and found little leakage ([Aker, 2017](#)), how much of the grant recipients had to give to other household and community members and participants who reported less than 1% of the grant ([Blattman et al., 2016](#)), and the occurrence of robbery where participants reported no occurrence ([Lehmann and Masterson, 2020](#)). Another study 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 (Aker et al., 2016). 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 (Sossouvi, 2013; Gordon, 2015; Overseas Development Institute, 2015; Ecker, Maystadt, and Guo, 2019). 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. This finding is especially informative given how extensively the scholarly and policy literatures have identified diversion as a core concern for in-kind assistance. Indeed, anecdotal evidence of diversion has been extensively documented across the globe (Perlez, 1992; Bryer and Cairns, 1997; De Waal, 1997; Anderson, 1999; Goodhand, 2002; Lischer, 2003; Maren, 2009; Polman, 2010; Barnett, 2013; Zürcher, 2019; O'Donnell, 2023) and cross-national inquiries also suggest that aid prolongs conflict in part due to diversion (Nunn and Qian, 2014; Narang, 2015; Wood and Sullivan, 2015; Wood and Molfino, 2016; Findley, 2018; Findley et al., 2023). (See A10, Column 8).

### 3 Costs of Delivery

The last three columns of A10 report whether the programs (a) reported on and delivered aid cost-efficiently or cost-effectively, (b) decentralized their aid distribution, and (c) could transparently track aid receipt and usage.

**Cost Analysis** Of the 23 reviewed programs, 12 report cost-efficiency or cost-effectiveness metrics. (A10 reports the derivable total cost transfer ratios (TCTRs) for nine of the 12; the remaining three, which we tag with a check mark (✓), 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 (and conservative) 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 WFP's or IRC's cost efficiency estimates, but the report of ECHO-funded projects reports a variety of disaggregated statistics (Maunder et al., 2016). 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. 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 (Overseas Development Institute, 2015), which, among other benefits, reduced potentially significant start-up costs. Others have argued that mobile transfers may only be more efficient than physical cash when a mobile network infrastructure is available, convenient, and clear (Aker et al., 2016; Jeong and Trako, 2022), which the private digital payments platform provided in our context.

These 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 (White, Hodges, and Greenslade, 2013), 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.

**Decentralization** Because most cash-based assistance has been centralized with a single donor or small set of donors and can depend critically on a small set of distribution points, which can be costly to establish and maintain, decentralizing assistance has core implications for costs of delivery. Only five of the 23 programs that we reviewed could be considered decentralized, capable of facilitating quicker access and use of funds. Of these 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 value 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, which can be expensive to collect and sustain. 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 ([Overseas Development Institute, 2015](#)), 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.

## B Intervention and 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 Online Appendix D we discuss ethical considerations. The Pre-Analysis Plan (PAP) is registered at:

<https://www.socialscienceregistry.org/versions/160809/docs/version/document>.

## 1 Research design

**Local context** Afghanistan is in its third year of a complex humanitarian crisis combining economic contraction, political repression, and ongoing concerns over stability and security. Its humanitarian appeal of \$4.6 billion in 2023 represented the world's largest single-country appeal to date ([United Nations, 2023c](#)). It is also the WFP's largest recipient by value of cash-based transfers and commodity vouchers ([World Food Programme, 2023a](#)) and though multiple financial service providers exist in the country, digital aid programs are still nascent ([United Nations Capital Development Fund, 2023](#)). Afghanistan's GDP contracted by 30-35% since the Taliban took over in

2021, and ever since the country has been under the threat of “winters of famine” ([World Bank, 2023a](#); [World Food Programme, 2023d](#)).

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 ([United Nations, 2023b](#)). In December 2022, the Taliban banned female Afghan employees from working in non-governmental organizations, prompting major foreign aid groups to suspend operations ([Roberts, 2023](#)). The ban was expanded in April 2023 to include female Afghan employees of the United Nations ([United Nations, 2023a](#)). 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 ([United Nations Assistance Mission in Afghanistan, 2023](#)). Western donors remain concerned that diversion of these substantial flows is enabling the world’s most gender-repressive regime ([O’Donnell, 2022](#); [Special Inspector General for Afghanistan Reconstruction, 2023a](#)) and that the Taliban are providing safe haven to international terrorist groups ([Senate Armed Services Committee, 2023](#)).

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 ([World Bank, 2023b](#)). 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 ([Asia Foundation, 2019](#)). Other studies have documented a substantial gender gap in women’s access to mobile phones ([Broadcasting Board of Governors, 2015](#)) and access to the internet ([Gallup, 2023](#)). For an overview of empirical research on digital payments in Afghanistan over the last decade, see ([Blumenstock et al., 2015](#); [Blumenstock, Callen, and Ghani, 2018](#); [Blumenstock et al., Forthcoming, 2023](#)).

**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 digital 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 digital 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 digital 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 Community Driven Development Organization (CDDO) and representatives from the digital payments provider. We conducted a third pilot with 52 women in 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 to 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 CDDO, an Afghan organization that assists Community Development Councils (CDCs) in a wide-array of local activities. The CDCs were established through local elections as part of the National Solidarity Program starting in 2004 ([Beath et al., 2016](#); [Beath, Christia, and Enikolopov, 2017b, 2013, 2017a, 2018](#)), 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. 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. After participants were identified, they were onboarded as described in the main text.

**Timeline and randomization** 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. The 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. Women in the experimental sample were assigned to one of two groups. The treatment group received four bi-weekly 4,000 Afghani (roughly 45

USD) aid payments via the digital payments platform first, between November 6, 2022 and December 20, 2022. The control group received the same payments between January 1, 2023 and February 12, 2023, after the treatment group had received all four of its payments. Figure 1 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 treatment and control groups, stratifying on two variables. First, they were stratified on the nahia (neighborhood) in which they registered. There were 16 nahias in total across Kabul, Mazar and Herat. Second, they were stratified on 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 dietary diversity. 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 assigned “misfits” independently across strata. Given that we had few strata and a single treatment arm, there were few misfits, so the risk of harming treatment fractions by independently assigning misfits across strata was low (Carril, 2017). We pre-specified this outcome as one of our primary food security outcomes in the PAP. As can be observed in our PAP, the treatment groups are balanced in 17 of the 18 outcome and heterogeneity variables we collected at baseline. The 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.

**Sample characteristics** A6 reports summary statistics for the experimental sample. 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. Using the 2015 Demographic and Health Survey (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.

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. Mean household 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). This is in line with recent evidence: A nationwide Gallup survey conducted just after the Taliban took control in August 2021 indicates that 94% of Afghans rate life satisfaction below 4 ([Gallup, 2022b](#)).

In addition, the digital 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 have some education (88%) and were, on average, 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.

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 digital payments platform, and only 2 already had accounts with the digital 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 digital 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 treatment group received its first payment to a couple of weeks after the control group received its last payment (see Figure 1).

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 [A12](#), but there is a slightly higher non-response rate among control participants in the second round of surveys (the difference is 6 surveys).

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 (Duflo et al., 2020), that two research assistants could take the data, and the PAP, and separately produce identical analyses.

## 2 Analysis

**Estimation strategy of baseline results** We estimate intent-to-treat treatment effects for pre-specified outcomes based on the following specification:

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

where  $Y_{itn}$  is the outcome of woman  $i$  in survey round  $t$  in nahia  $n$ . 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. 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 randomization.

We conduct our analysis in accordance with our PAP. For each of the three families of outcomes, 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 questions, 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 income, and employment status of the head of the household. The medicine 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 treatment group is doing better health-wise due to the aid payments), we also calculate Lee bounds for this variable following (Lee, 2009). The 95% confidence interval for the Lee bounds goes from 0.014 to 0.068.

To increase power and precision, for each family of outcomes, we also summarized the primary measures into a single summary index following (Kling, Liebman, and Katz, 2007), which we denote the Kling, Liebman, and Katz "KLK" index. The 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. 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 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 [A5](#), 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 treatment group has received 3-4 aid payments, compared to the first round of follow up surveys, after the treatment group has received only 1-2 aid payments. We do so by estimating the following specification:

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

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. Participants might feel uncomfortable answering questions about informal taxation by community leaders and government officials, and so they might 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 we can estimate differences in the shares of respondents reporting they experienced a 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 ([Boudreau et al., 2023](#)).

We implemented a list experiment to gauge the prevalence of informal taxation in the population. There are two items 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 control group was receiving their 3<sup>rd</sup> and 4<sup>th</sup> 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 told 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 applied to them.

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

$$\text{Number of Statements}_{isn} = \rho_0 + \rho_1 \text{Treated List}_{isn} + \rho_2 X_{isn} + v_{isn} \quad (3)$$

where  $s$  corresponds to the surveyor in charge of the survey. We control for strata fixed effects,  $X_{isn}$ . We do this combining the whole sample, for the control group only (which has been receiving aid payments for over a month) and for the treatment group only (which has not received any aid payments in over a month). Panel C of Table 3 shows the results of this exercise. Expanding the regression equation by adding an interaction term between the main treatment assignment (early vs. late aid payments) and the list experiment’s treatment assignment does not change results, with the coefficient on the interaction term equal to 0.0338 (SE = 0.058).

**Checking treatment impact heterogeneity** We look for differential treatment effects on outcomes by subgroup by estimating the following specification:

$$\begin{aligned} Y_{itn} = & \mu_0 + \mu_1 \mathbb{1}[\text{Treatment Group}]_{in} \times \mathbb{1}[\text{Heterogeneity}]_{in} + \mu_2 \mathbb{1}[\text{Heterogeneity}]_{in} \\ & + \mu_3 \mathbb{1}[\text{Treatment Group}]_{in} + \mu_4 X_{i0n} + \mu_5 Y_{i0n} + \mu_6 \mathbb{1}[t = 2] + v_{itn} \end{aligned} \quad (4)$$

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, and 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 A2.

**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 [A3](#). 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 treatment group has a slightly higher preference for digital aid than the control group, potentially because they have had more experience receiving the aid digitally, although responses are very similar across the two groups.

Outside the study, the digital payments 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.

### 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. 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.

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 (De Quidt, Haushofer, and Roth, 2018). 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:

$$Y_{isn} = \psi_0 + \psi_1 \text{Primed}_{isn} + \psi_2 X_{isn} + \psi_3 Y_{is0n} + \omega_{isn} \quad (5)$$

where we control for strata fixed effects (surveyor and treatment 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 randomization.

In addition, we test whether the prime affected participants’ responses differently depending on whether they belonged to the treatment or control group, by estimating the specification:

$$\begin{aligned} Y_{isn} = & \eta_0 + \eta_1 \text{Primed}_{isn} + \eta_2 \text{Primed}_{isn} \times \mathbb{1}[\text{Treatment Group}]_{isn} + \eta_3 \mathbb{1}[\text{Treatment Group}]_{isn} \\ & + \eta_4 X_{isn} + \eta_5 Y_{is0n} + \omega_{isn} \end{aligned} \quad (6)$$

The results of the experimenter demand effect analysis are shown in A4. 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 control group,  $\hat{\eta}_1$ , and column 5 shows the (overall) experimenter demand effects for the treatment 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 Afghanis 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 treatment 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 treatment group women, when they were receiving the aid payments, helped the control group women. 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).

## 4 Deviations from PAP

As mentioned before, our PAP is registered at

<https://www.socialscienceregistry.org/versions/160809/docs/version/document>. All the analysis was conducted after the submission of this PAP.

We adhered 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 do not show Lee Bounds results. We do not report results on participants' satisfaction with the digital payments platform, HesabPay, but there were only 2 instances of participants stating that they were "somewhat dissatisfied" with HesabPay (out of 2393 responses in the treatment 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 A5). As a robustness check, we also include an alternative way of creating summary indices following (Anderson, 2008), 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 (Kling, Liebman, and Katz, 2007), the KLK index. Upon further consideration, and after computing the results, we

decided to additionally show results using a dummy for whether the respondent answered yes to any of the four diversion tables (column 6 in Table 3), which was not pre-specified. This is because, due to the lack of positive responses to the informal taxation questions and the way in which the index is created (standardising first each component variable, and then the final index) leads to extremely large values for those who answered yes to multiple questions. For reference, out of the 59 instances of participants answering yes to any of the four questions in any of the two survey rounds, 49 said yes to only one question, 9 said yes to two questions, and only one said yes to three questions (e.g. index value for this woman is 34, a clear outlier), with no one saying yes to all four. For completeness and robustness, A7 shows the 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 (Anderson, 2008) (column 3). Moreover, we also present results dropping the one outlier that answered yes to three of the taxation questions in one round and received a KLK Index value above 34 standard deviations (column 4). None of these alternative indices was pre-specified. As can be observed, while the results using the pre-specified KLK index in Table 3 are only significant at the 10% level, the other four alternative indices are all statistically insignificant at conventional levels. This suggests that the one outlier is what leads to the significant results in the main results. We also do not report the  $p$ -values adjusted for multiple hypothesis testing for the individual diversion measures, but they are all above 0.25.

## 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 (Bete, 2022), roughly equivalent to two biweekly direct aid transfers of 4000 AFA and a monthly survey incentive of 350 AFA. This basket was composed of 89 kg wheat flour, 21 kg domestic rice, 7 liters vegetable oil, 9 kg puSIses, and 1 kg salt at prevailing exchange rates in August 2022. For reference, the Minimum Expenditure Basket including food, health-care, shelter, and other components totaled almost twice as much at 181.36 USD. 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 food security outcomes, 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 past week declined after each payment (Figure 2, 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 treatment 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%) indicated saving some food for future consumption, but typically only mentioned having one or two weeks' 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 (Lehmann and Masterson, 2020). Two other programs found positive effects, but the differing outcome measurements complicate effect size comparisons (Pople et al., 2021; Bedoya et al., 2019). 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 (Alik-Lagrange et al., 2019; Chowdhury et al., 2017), vouchers in D.R.C. did not reduce days relying on a variety of coping strategies that included meal skipping (Quattrochi et al., 2022), 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 (MercyCorps, 2022).

## C Cost analysis

### 1 Cost-efficiency

We estimated three cost-efficiency quantities: the total cost-transfer ratio (TCTR), the cost-transfer ratio (CTR), and 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 a household (Riegg Cellini and Edwin Kee, 2015). All estimates are based on the variable operational costs of the digital payments provider, NGO, & community development organization. We excluded fixed costs, which consist of the digital provider's platform origi-

ination and standard maintenance costs. 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.

[A13](#) reports raw costs and estimated cost-efficiency quantities. Panel A reports raw costs (a) disaggregated by category, (b) including/excluding onboarding, and (c) by treatment and control group status, which both received transfers but at different times. Panel B reports the CPB, including/excluding the transfer value, and including/excluding onboarding. 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).

To contextualize our cost-efficiency estimate of 6.7 cents, we estimate how many more individuals could have been served by the WFP in 2022 had their cash-based portfolio been delivered at 6.7 cents rather than 17 cents for each dollar. Using reasonable assumptions, we estimate that had WFP delivered all \$357M digitally, the savings would have been sufficient to support an additional 77,000 households or 538,075 individuals for the four month lean season. More conservatively, if we assume that WFP's costs were lower by 1 cent (CTR=0.16) and ours were higher by the same amount (CTR=0.077), then 432,250 individuals would have been supported through the four month lean season. More ambitiously, if we assume that WFP's costs were higher by 1 cent (CTR=0.18) and ours were lower by the same amount (CTR=0.057), then 641,613 additional individuals would have been supported. The replication materials contain our inputs, assumptions, calculations, and can be adjusted to explore different scenarios.

## 2 Cost-effectiveness

We compute cost-effectiveness ratios (CERs) by dividing costs by treatment effects aggregated over participants and time ([Riegg Cellini and Edwin Kee, 2015](#)), which are interpreted as the number of dollars required to achieve a defined impact. All CER estimates appear in [A14](#).

We first estimated the cost required to achieve a reduction of a day with skipped meals, one of our key preregistered outcomes. The 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. Most conservatively, the average cost to achieve a reduction of a day with skipped meals is \$31.58 USD:  $\$231,936.08 / (1208 \text{ beneficiaries} * 0.760 \text{ fewer days with skipped meals per week} * 8 \text{ weeks})$ . 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 (Panel C).

Reducing days with skipped meals is important for many reasons, especially for improving the health of household members. We draw on the CVWG estimates of calorie consumption per person-household-day ([Bete, 2022](#)) 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 (Panel D, control mean). With this benchmark, we can estimate the cost required to increase the number of calories by a fixed amount, here by 1,000. Averaging across the full eight-week program period (Panel D), we estimate that 2.38 USD in digital aid results in an increase of 1,000 calories:  $\$231,936.08 / (1208 \text{ beneficiaries} * (97,323,825 \text{ fewer forgone calories} * 8 \text{ weeks}))$ . Based on estimates from the second half of the program (Panel E), \$1.73 USD results in a 1,000 calorie increase. Based on estimates only from the program's final week (Panel F), \$1.28 USD results in a 1,000 calorie increase.

Because digital aid continued to have an effect after the program ended (see Figure 2), 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. For the post-program period, the treatment group is compared to the control group's baseline value (2.615) because in weeks 9–16 the control group is now receiving transfers and no longer constitutes a control. The treatment group's

12- and 16-week values could be compared to those of the control group just before receiving treatment (2.546), but by that point the control 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 treatment group's own baseline value (2.626) or an average of the treatment and control groups' baseline values. In practice, all three are similar and do not qualitatively change the inferences. 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 treatment 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 ( $\$231,936.08 / (209,950,567 / 1000)$ ).

By improving the food security 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, 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 ([Clarke and Darcy, 2014](#)).

## D 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).

As 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. Additionally, 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 three 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 consortium 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. Because 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 ([Burton, 2020](#); [World Food Programme, 2021a](#); [United Nations, 2021](#)).

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 Belmont Report principles ([US Department of Health and Human Services, 1979](#)), and further address considerations specific to insecure, humanitarian crisis environments ([Wood, 2006](#); [Campbell, 2019](#); [Puri et al., 2017](#); [Wolfe, 2020](#)).

### 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 Online Appendix 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 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.

Treatment and control 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.

## 2 Beneficence

*Risks:* We prioritized the long-standing principle of “do no harm” ([US Department of Health and Human Services, 1979](#); [Wood, 2006](#)). In addition to designing the study

## ONLINE APPENDIX

to minimize the risk of potential harm (See Online Appendix 1), 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 included: (a) implementation through CDCs comprised of 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 2% of participants reported any attempts at 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 individuals and types of IPV (Hidrobo, Peterman, and Heise, 2016; Buller et al., 2018; Bastagli et al., 2019; Baranov et al., 2021; Blofield et al., 2022). For example, there may be isolated negative impacts for women with similar or higher education than their partner (Hidrobo and Fernald, 2013) or when transfers are relatively large (Angelucci, 2008). They may also be more pronounced for emotional violence and threats rather than physical violence (Bobonis, 2011; Bobonis, González-Brenes, and Castro, 2013). We deliberately targeted female-headed households as this would help miti-

gate 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 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, 98% of women reported that they had not been approached by anyone for money, and 93% 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 ([McDermott and Hatemi, 2020](#); [American Political Science Association, 2020](#)). 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 CDCs and community members, and had full legal approval to operate. It collected data for the baseline survey as part of the registration and onboarding. For follow-up surveys, the enumeration team operated entirely remotely, reducing the security risks associated with having a physical presence in communities and at the homes of beneficiaries. 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 befall 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 ([US Department of Health and Human Services, 1979](#)). In practice, quantifying direct and indirect benefits can be challenging and perhaps even misleading ([Baele, 2013](#); [Humphreys, 2015](#)), 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 transfer, 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 con-

ducting a randomized trial during a humanitarian crisis (Wolfe, 2020). 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.

### 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 CDCs, 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 three 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.

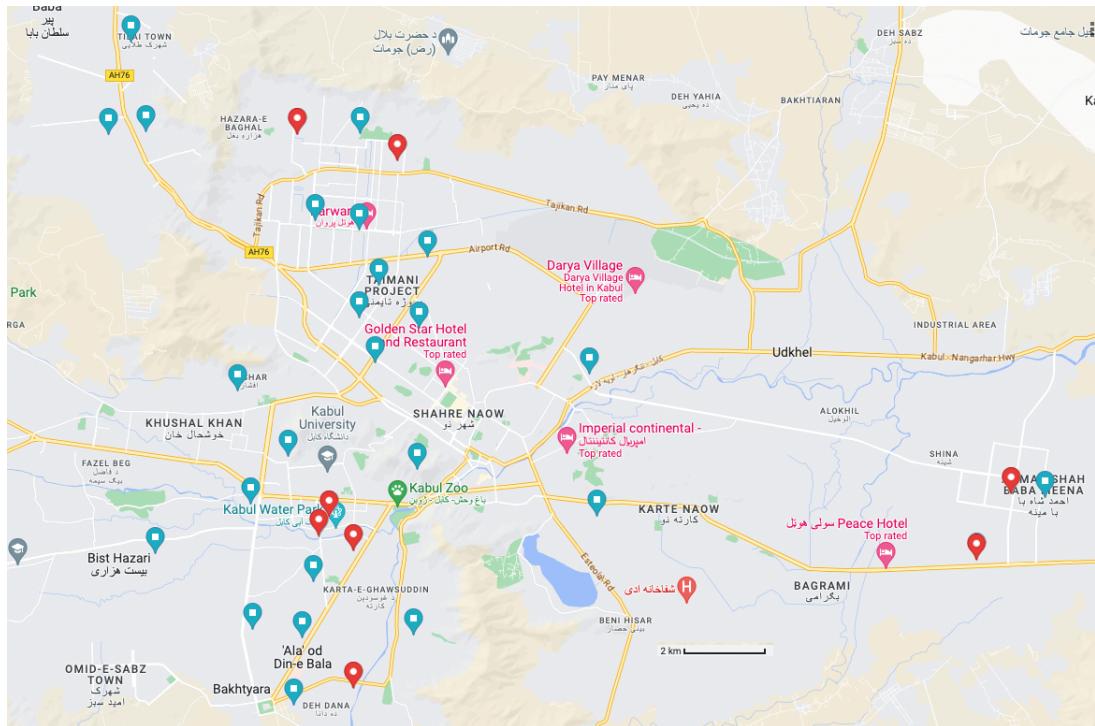
Beneficiaries were randomly assigned to treatment and control using a phase-in/waitlist design so that program benefits were not denied to anyone deemed eligible through the targeting and recruitment process. Although all study participants eventually received the program, a relevant question is whether 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 5 and further note that in contrast to the relatively long timeline of many randomized programs, 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 treatment or control group, they all received the survey incentive each time that they participated. As such, those in the control group did not accrue costs to their time that the treatment 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.”

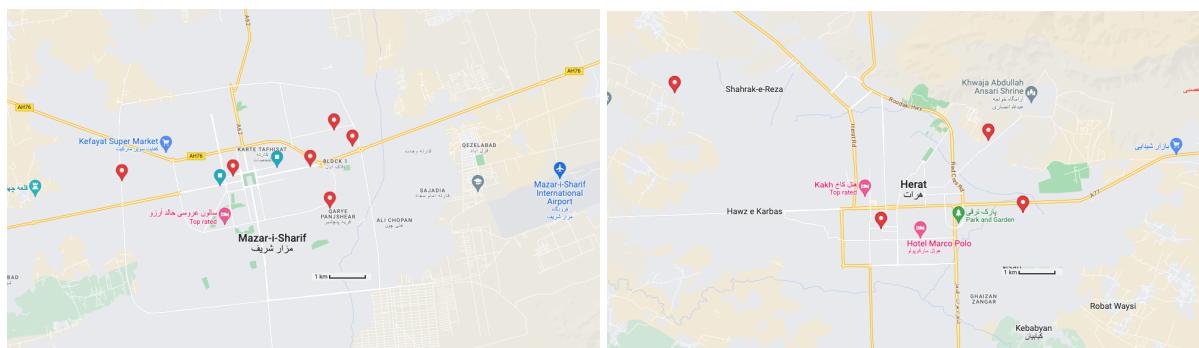
## ONLINE APPENDIX

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).

## E Supplementary material figures and tables



(a) Kabul



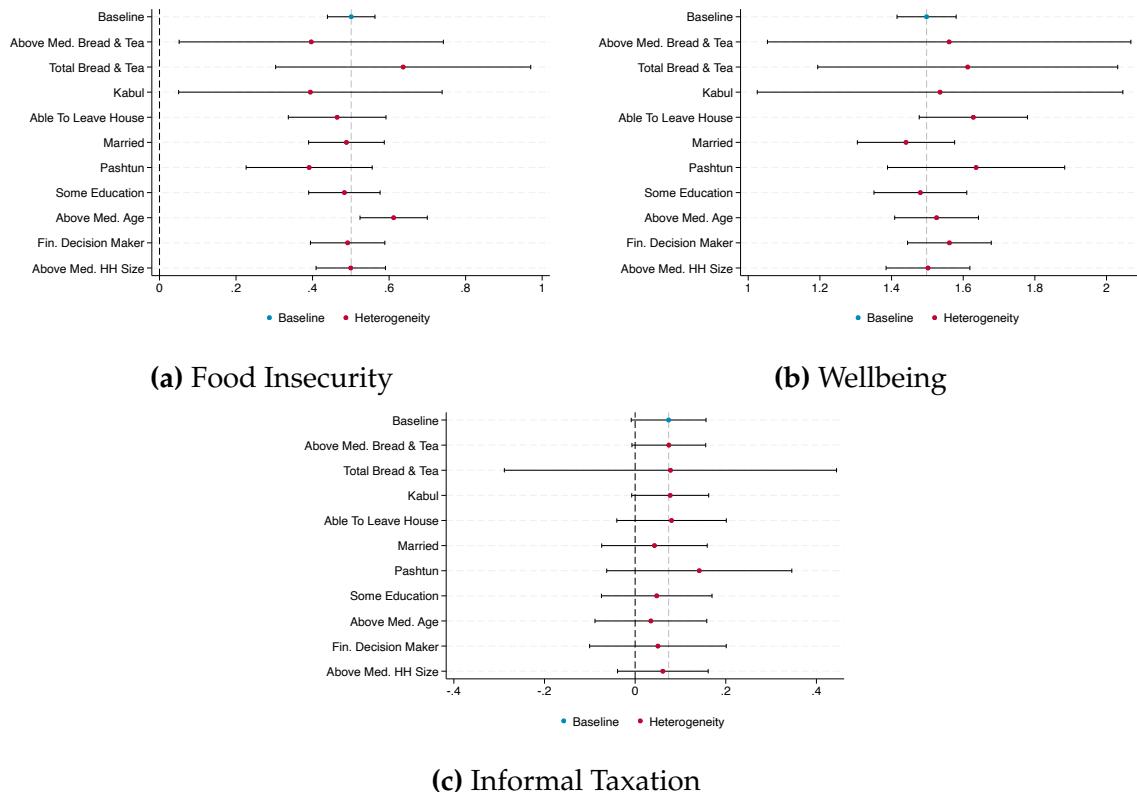
(b) Balkh

(c) Herat

**Figure A1: Location of Merchants**

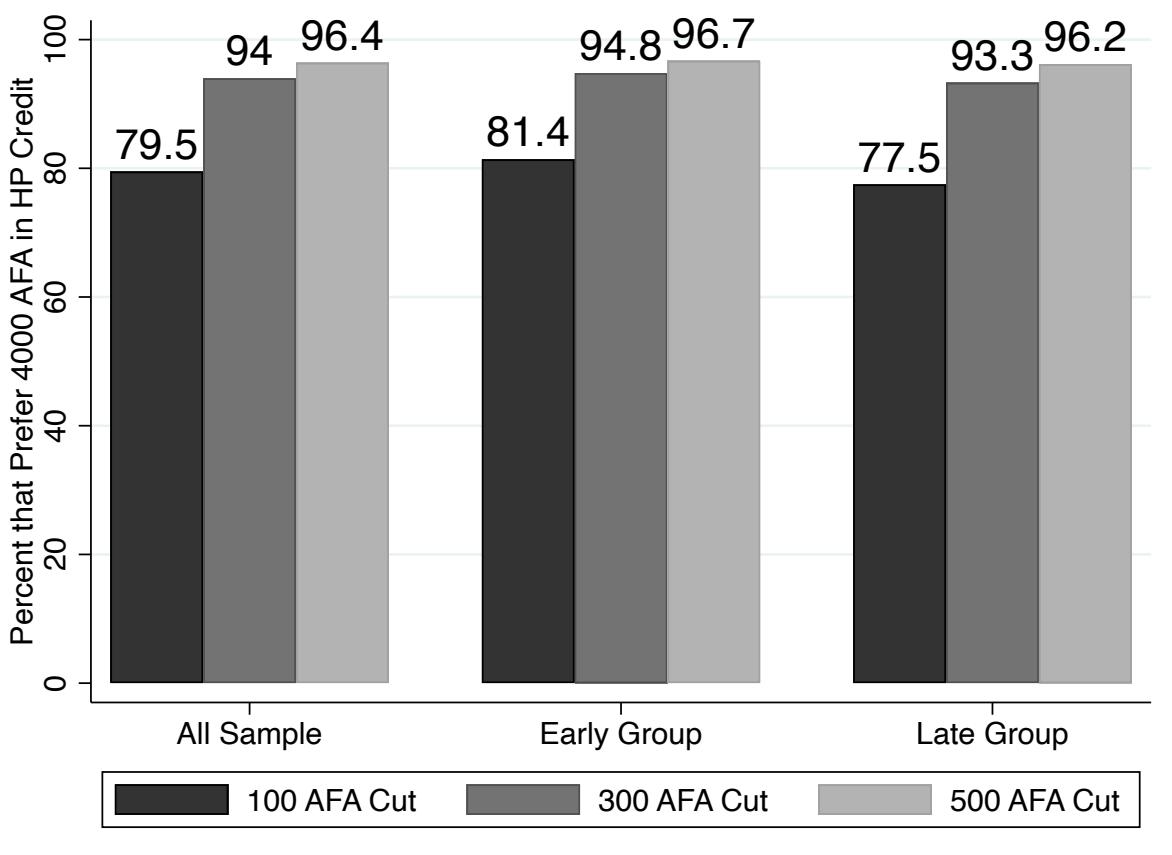
**Notes:** The map shows the location of merchants that participants visited at some point between the start of the program and December 31, 2022. Red pins are those that belong to merchants who participated in the onboarding sessions, while blue pins are those of other merchants. Note that not all merchants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

## ONLINE APPENDIX



**Figure A2: Heterogeneity of Treatment Effects**

**Notes:** The blue dots and the gray-shaded vertical lines correspond to the baseline estimate  $\hat{\gamma}_1$  from Equation (1). The red dots correspond to  $\hat{\mu}_1 + \hat{\mu}_2 + \hat{\mu}_3$  from Equation (4). Bars show 95% confidence intervals.



**Figure A3: 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.” We 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.

## ONLINE APPENDIX

**Table A1:** 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.

## ONLINE APPENDIX

**Table A2: Baseline Balance Check**

Variable	Whole	Sample	Treatment	Control	p-value
	Mean (1)	SD (2)	Mean (3)	Mean (4)	Difference (5)
<b>Panel A. Outcome Vars</b>					
1. Total bread-tea meals	13.76	2.56	13.74	13.77	0.79
2. Skipped meals	2.62	1.52	2.63	2.62	0.86
3. Inf tax gov (others)	0.003	0.058	0.002	0.004	0.474
4. Inf tax leaders (others)	0	0.02	0	0.001	0.317
5. Inf tax gov (you)	0	0	0	0	1
6. Inf tax leaders (you)	0.002	0.041	0	0.003	0.045
7. Life satisfaction	3.53	1.17	3.51	3.54	0.59
8. Happy	0.01	0.1	0.01	0.01	0.99
<b>Panel B. Heterogeneity Vars</b>					
1. Married	0.34	0.47	0.34	0.34	0.88
2. Pashtun	0.1	0.3	0.11	0.09	0.15
3. Some education	0.37	0.48	0.37	0.36	0.37
4. Balkh	0.33	0.47	0.33	0.33	0.77
5. Herat	0.33	0.47	0.33	0.33	0.85
6. Kabul	0.34	0.47	0.34	0.34	0.95
7. Above median age	0.56	0.5	0.56	0.56	0.81
8. Fin. decision maker	0.66	0.47	0.66	0.66	0.99
9. Above median HH size	0.64	0.48	0.64	0.65	0.6
10. Able to leave house	0.65	0.48	0.64	0.66	0.56
Number of individuals	2409		1208	1201	

*Notes:* The table shows, for different pre-specified variables at baseline, the overall mean and standard deviation (columns 1 and 2), the mean in the treatment group (column 3) and the mean in the control group (column 4). Column 5 shows the p-value of the difference between the means in the treatment and control group, adjusting for robust standard errors.

## ONLINE APPENDIX

**Table A3: Treatment Effects by Survey Round – Secondary Outcomes**

	(1)	(2)	(3)	(4)	(5)
<b>Panel A. Types of Food</b>	Rice	Beans	Vegetables	Chicken	Dairy
$\beta_1$ : Treated $\times$ Round 2	0.366*** (0.061)	0.240*** (0.050)	0.014 (0.071)	0.001 (0.011)	0.032 (0.023)
$\beta_2$ : Round 2	-0.051 (0.040)	-0.124*** (0.030)	0.164*** (0.051)	-0.008 (0.006)	-0.065*** (0.015)
$\beta_3$ : Treated	0.413*** (0.049)	0.373*** (0.041)	-0.011 (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	0.088	0.084	0.013	0.012	0.014
<b>Panel B. Other Economic Outcomes</b>	Income	Employed	Fin. Decision-Maker	Medicine Purchase	
$\beta_1$ : Treated $\times$ Round 2	36.972 (151.108)	0.050** (0.023)	0.014 (0.022)	0.008 (0.019)	
$\beta_2$ : Round 2	-165.374** (64.319)	0.032* (0.017)	0.030** (0.015)	-0.046*** (0.012)	
$\beta_3$ : Treated	109.985 (149.582)	-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	0.012	0.024	0.064	0.015	

*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:* \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

## ONLINE APPENDIX

**Table A4:** Summary Table – Experimenter Demand Effects

Variable	Control Mean (1)	Baseline Estimate (2)	Overall Estimate (3)	Experimenter Demand		
				FWER p-value (4)	Control Estimate (5)	Treatment Estimate (6)
<b>Panel A. Food Security</b>						
Days skipping meals (past week)	2.615	-0.76*** (0.051)	-0.031 (0.062)	0.7225 (0.087)	-0.082 (0.087)	0.041 (0.077)
Children skipping meals (=1)	0.873	-0.117*** (0.012)	0.016 (0.016)	0.3913 (0.021)	0.028 (0.021)	0.009 (0.024)
Regularly eat twice a day	0.501	0.093*** (0.015)	-0.001 (0.014)	0.925 (0.02)	-0.004 (0.02)	-0.001 (0.02)
Total bread and tea meals (past week)	13.639	-1.608*** (0.121)	-0.186 (0.161)	0.3897 (0.212)	-0.059 (0.212)	-0.267 (0.217)
<i>Food Security - KLK Index</i>	0	0.501*** (0.032)	0.005 (0.039)		-0.011 (0.05)	0.007 (0.051)
<b>Panel B. Informal Taxation</b>						
Inf. tax gov. off. (others)	0.004	0.001 (0.001)	-0.002 (0.002)	0.4468 (0.002)	-0.002 (0.002)	-0.001 (0.003)
Inf. tax comm. leader (others)	0.001	0.002 (0.002)	-0.002 (0.003)	0.4513 (0.002)	0 (0.002)	-0.005 (0.005)
Inf. tax gov. off. (you)	0	0.002** (0.001)	-0.001 (0.001)	0.4513 (0.001)	-0.001 (0.001)	-0.001 (0.003)
Inf. tax comm. leader (you)	0.003	0.002 (0.003)	-0.006** (0.003)	0.2026 (0.004)	-0.006 (0.004)	-0.006 (0.004)
<i>Informal tax. - KLK Index</i>	0	0.074* (0.042)	-0.1* (0.057)		-0.075 (0.05)	-0.127 (0.093)
<b>Panel C. Wellbeing</b>						
Better economic situation	0.048	0.335*** (0.011)	0.006 (0.019)	0.9754 (0.017)	0.016 (0.026)	-0.012 (0.026)
Satisfied with fin. situation	0.133	0.263*** (0.012)	-0.003 (0.018)	0.9754 (0.02)	0.005 (0.02)	-0.017 (0.026)
Happy	0.009	0.28*** (0.014)	-0.005 (0.017)	0.9754 (0.02)	-0.014 (0.022)	-0.003 (0.022)
Life satisfaction (std)	0.011	1.682*** (0.058)	0.002 (0.089)	0.9754 (0.098)	0.043 (0.111)	-0.081 (0.111)
<i>Economic/Wellbeing - KLK Index</i>	0	1.498*** (0.042)	0.004 (0.071)		0.027 (0.064)	-0.059 (0.086)

*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). Primary outcomes show FWER-adjusted p-values within each family outcome (following Romano & Wolf, 2005, using 5000 repetitions). 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.

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

## ONLINE APPENDIX

**Table A5: 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)
<b>Panel A. Primary Outcomes</b>							
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	0.124	0	0.0002	4412
<i>Food Security - KLK Index</i>	-0.016	1.01	0.511	0.033	0		4412
<i>Food Security - Anderson Index</i>	0.002	1.058	0.494	0.034	0		4412
Inf. tax gov. off. (others)	0.002	0.043	0.001	0.001	0.563	0.6818	4412
Inf. tax comm. leader (others)	0.002	0.043	0.003	0.002	0.116	0.108	4412
Inf. tax gov. off. (you)	0.001	0.03	0.002	0.001	0.163	0.1524	4412
Inf. tax comm. leader (you)	0.005	0.071	0.001	0.002	0.609	0.6818	4412
<i>Informal tax. - KLK Index</i>	-0.007	0.923	0.075	0.041	0.067		4412
<i>Informal tax. - Anderson Index</i>	-0.016	0.763	0.048	0.031	0.115		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	2	0.07	0	0.0002	4412
<i>Economic/Wellbeing - KLK Index</i>	0.001	1	1.51	0.043	0		4412
<i>Economic/Wellbeing - Anderson Index</i>	0.096	1.167	1.326	0.045	0		4412
<b>Panel B. Secondary Outcomes</b>							
Days eating rice (past week)	0.672	1.05	0.616	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	0.991	0.415	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	0.002	0.004	3344
Involved in fin. decisions	0.672	0.47	0.019	0.016	0.228	0.15	4412
Total household income (past month)	1093.634	1792.27	156.196	94.148	0.097	0.07	4412
Household's head employed (past month)	0.284	0.451	-0.008	0.014	0.573	0.274	4412

*Notes:* Controls for stratification fixed effects, survey round fixed effects, and baseline values of dependent variables, if available, are included. Standard errors are clustered at the 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).

## ONLINE APPENDIX

**Table A6:** 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 for this question at baseline.

**Table A7:** Are Digital Payments Diverted? Indices

	KLK Index Others (1)	KLK Index You (2)	Anderson Index All (3)	Dropping Outlier (4)
<b>Panel A. Baseline</b>				
Treated	0.034 (0.026)	0.043 (0.031)	0.049 (0.036)	0.059 (0.039)
Observations	4,611	4,647	4,648	4,647
Control Mean	-0.003	0.015	0.011	0.012
<b>Panel B. Long-Run</b>				
$\beta_1$ : Treated $\times$ Round 2	0.072 (0.053)	0.024 (0.061)	0.051 (0.071)	0.062 (0.079)
$\beta_2$ : Round 2	-0.048 (0.029)	-0.096*** (0.037)	-0.135*** (0.046)	-0.137*** (0.045)
$\beta_3$ : Treated	-0.073 (0.085)	0.007 (0.109)	-0.027 (0.125)	-0.034 (0.137)
Observations	4,611	4,647	4,648	4,647
$(\beta_1 + \beta_2 + \beta_3) / \beta_3$	0.670	-9.410	4.090	3.230

*Notes:* In Panels A and B, controls for stratification fixed effects, survey round fixed effects, and baseline values of dependent variables, if available, are included. 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). Column 4 uses the KLK Index composed of the four informal taxation questions (as in the baseline results), but dropping the one observation with the highest KLK Index value. Control for surveyor fixed effects are included. Standard errors are clustered at the individual level.

## ONLINE APPENDIX

**Table A8: Experts' Survey Questionnaire**

Statement	Answer Options
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.	
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.	Yes
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?	0-100%
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.)	0-21 Bread and Tea Meals
Q3) What share of beneficiaries do you expect will report efforts to tax or divert their payments?	0-100%
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%.)	0-100 cents

## ONLINE APPENDIX

**Table A9: 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 ([Food Security Information Network, 2023](#)), with the share of population being the share of the population analyzed by FSIN. Active armed conflict ([Davies, Pettersson, and Öberg, 2023](#)). Level of freedom ([Freedom House, 2023](#)). Community development council (CDC) presence coded from various sources (authors' research). Mobile-cellular subscriptions per 100 people ([UN International Telecommunication Union, 2022](#)). Mobile money prevalence index ([Andersson-Manjang, 2021](#)).

**Table A10:** Cash-based transfer programs with rigorous evaluation in crisis contexts

Country/ Study	Target Beneficiaries	Cash Type	Delivery Mechanism	Usage (Mobile)	Food Security	Mental Well-Being	Diversion Discussed	Cost (TCTR)	Decent- ralized	Trans- parent
1. Ecuador ( <a href="#">Hidrobo et al., 2014</a> )	Refugees	C/V/F	ATM/physical		+			1.075	✓	
2. Philippines ( <a href="#">MercyCorps, 2022</a> )	Affected HHs	C	Mobile money	✓	Ø			1.020		
3. Philippines ( <a href="#">Kandpal et al., 2016</a> ) and ( <a href="#">Crost, Felter, and Johnston, 2016</a> )	Poor HHs	C	Physical		+					
4. Lebanon ( <a href="#">Salhi et al., 2022</a> ) and ( <a href="#">Moussa et al., 2022</a> )	Refugees	C	ATM		+	+			✓	
5. Lebanon ( <a href="#">Lehmann and Masterson, 2020</a> )	Refugees	C	ATM		+	Ø	✓		✓	
6. Lebanon ( <a href="#">de Hoop, Morey, and Seidenfeld, 2019</a> )	Refugees	C	Not stated							
7. Sri Lanka ( <a href="#">Sandström and Tchatchua, 2010</a> )	Affected HHs	C/F	Physical		+				✓	
8. Bangladesh ( <a href="#">Pople et al., 2021</a> )	At-risk HHs	C	Mobile money	✓	+	+			✓	✓
9. Uganda ( <a href="#">Blattman et al., 2016</a> )	Poor women	C	Physical		+		✓		✓	
10. Niger ( <a href="#">Aker et al., 2016</a> )	Poor women	C	Mobile/physical	✓	+			1.059		✓
11. Niger ( <a href="#">Hoddinott, Sandström, and Upton, 2018</a> )	Poor HHs	C/F	ATM		+			1.053		
12. Niger ( <a href="#">Bossuroy et al., 2022</a> ) and ( <a href="#">Premand and Barry, 2022</a> )	Poor women	C	Physical		+	+		1.252		
13. Niger ( <a href="#">Langendorf et al., 2014</a> )	Poor children	C/F	Physical		+					
14. Iraq ( <a href="#">Kurtz et al., 2021</a> )	Poor HHs	C	Physical		+					
15. D.R.C. ( <a href="#">Aker, 2017</a> )	Displaced	C/V	Bank/physical		Ø		✓	1.087		
16. D.R.C. ( <a href="#">Bonilla et al., 2017</a> )	Displaced	C/V	Multiple		Ø					
17. D.R.C. ( <a href="#">Quattrochi et al., 2022</a> )	Displaced	V	Physical		Ø	+		1.437		
18. C.A.R. ( <a href="#">Alik-Lagrange et al., 2019</a> )	Poor HHs	C	Physical		Ø	+				
19. Yemen ( <a href="#">Schwab, 2020</a> ) and ( <a href="#">Schwab, Margolies, and Hoddinott, 2013</a> )	Poor HHs	C/F	Physical		+			1.083		
20. Yemen ( <a href="#">Kurdi, 2021</a> )	Poor women	C	Physical		+					
21. Afghanistan ( <a href="#">Bedoya et al., 2019</a> )	Poor women	C	Physical		+	+		2.017	✓	✓
22. Afghanistan ( <a href="#">Lyall, Zhou, and Imai, 2020</a> )	At-risk youth	C	Mobile money	✓						
23. South Sudan ( <a href="#">Chowdhury et al., 2017</a> )	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; Ø =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, 2022](#)). Ecuador and Philippines are not ranked or discussed; Lebanon and Sri Lanka are not ranked, but discussed as exhibiting early-warning signs.

**Table A11:** Related food security, mental well-being, and diversion results

Program	Relevant study context & notes	Selected results
1. 2011 WFP food security & Food security for refugee integration (Ecuador, <a href="#">Hidrobo et al. (2014)</a> ).	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, <a href="#">MercyCorps (2022)</a> ).	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 food security outcomes (coping strategies index, FCS dietary diversity).
3. 2009–11 Pantawid Pamilya poverty reduction program (Philippines, <a href="#">Kandpal et al. (2016)</a> and <a href="#">Crost, Felter, and Johnston (2016)</a> ).	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, <a href="#">Salti et al. (2022)</a> and <a href="#">Moussa et al. (2022)</a> ).	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, <a href="#">Lehmann and Masterson (2020)</a> ).	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, <a href="#">de Hoop, Morey, and Seidenfeld (2019)</a> ).	RD assessment of 441 USD cash / control to 1440 Syrian refugee HHs to offset commuting / foregone income (part of government education initiative).	No food security, mental well-being, or diversion outcomes reported.
7. 2005–6 WFP cash transfers for tsunami relief (Sri Lanka, <a href="#">Sandström and Tchatchua (2010)</a> ).	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, <a href="#">Pople et al. (2021)</a> ).	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, <a href="#">Blattman et al. (2016)</a> ).	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, <a href="#">Aker et al. (2016)</a> ).	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, <a href="#">Hod-dinott, Sandström, and Upton (2018)</a> ).	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, <a href="#">Bossuroy et al. (2022)</a> ).	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, <a href="#">Langendorf et al. (2014)</a> ).	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, <a href="#">Kurtz et al. (2021)</a> ).	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., <a href="#">Aker (2017)</a> )	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., <a href="#">Bonilla et al. (2017)</a> )	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., <a href="#">Quattrochi et al. (2022)</a> )	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., <a href="#">Alik-Lagrange et al. (2019)</a> )	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, <a href="#">Schwab, Margolies, and Hod-dinott (2013)</a> and <a href="#">Schwab (2020)</a> ).	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, <a href="#">Kurdi (2021)</a> ).	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, <a href="#">Bedoya et al. (2019)</a> ).	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, <a href="#">Lyall, Zhou, and Imai (2020)</a> ).	RCT evaluation of 75 USD cash / training / combined to 1,841 youth (756 control) at risk of insurgent recruitment (supported by local government).	No food security, mental well-being, or diversion outcomes reported.

23. 2013–14 Cash / Targeting the Ultra Poor (South Sudan, <a href="#">Chowdhury et al. (2017)</a> ).	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).
--	---	---

## ONLINE APPENDIX

**Table A12: 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	0.017	0.017

*Notes:* Controls for strata fixed effects are included. Standard errors are clustered at the 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 A13:** Cost-efficiency calculations

<b>Panel A: Raw costs</b>	
<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
<b>Panel B: Costs per beneficiary (CPB)</b>	
<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
<b>Panel C: Transfer ratios</b>	
<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.

## ONLINE APPENDIX

**Table A14:** Cost-effectiveness ratios: days with skipped meals and corresponding calorie intake

<b>CERs for Reduction in Days with Skipped Meals</b>				
<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>				
Control Group	2.569	24,682.952	36.700%	
Treatment 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>				
Control Group	2.448	11,760.670	34.973%	
Treatment 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>				
Control Group	2.508	3,012.108	35.823%	
Treatment Group	1.088	1,314.304	15.543%	
Effect	-1.42	-1,715.36	-20.286%	\$16.90
<b>CERs for Reduction in Foregone Calories</b>				
<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>				
Control Group	34,041.82	327,073,797	36.700%	
Treatment 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>				
Control Group	32,439.77	155,840,670	34.973%	
Treatment 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>				
Control Group	33,233.51	39,913,443	35.829%	
Treatment 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 treatment 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 treatment group and control 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 treatment group and control group respectively (1208 or 1201 participants \* 92,757 calories per household week \* 8 weeks).