The COVID Cash Transfer Study: The Impacts of an Unconditional Cash Transfer on the Wellbeing of Low-Income Families

Natasha V. Pilkauskas †
Brian A. Jacob †
Elizabeth Rhodes ‡
Katherine Richard †
H. Luke Shaefer †

May 2022

Special thanks to all partner organizations who made this study possible, including GiveDirectly, Propel, OpenResearch, Stand for Children and the Schusterman Family Philanthropies. Particular thanks to Alex Nawar, Michael Cooke, Miriam Laker, Farheen Rizvi, Jeff Kaiser, Cherie Chung, Jeremy Guardiola, Marty Moore, Kaitlin Raimi, and Jonah Edelman. Greg Duncan, Jonathan Morduch, and Johannes Haushofer provided excellent feedback on design and implementation. An extra special thank you to Karen Kling for excellent project management and to Poverty Solutions and Youth Policy Lab at the University of Michigan. More information can be found in our AEA Pre-Trial Registry (https://doi.org/10.1257/rct.5852-1.0).

† University of Michigan, Gerald R. Ford School of Public Policy, Natasha Pilkauskas (npilkaus@umich.edu), Brian Jacob (bajacob@umich.edu), Katherine Richard (krzrich@umich.edu), Luke Shaefer (lshaefer@umich.edu)
‡OpenResearch, Elizabeth Rhodes (elizabeth@openresearchlab.org)

Abstract

There is growing interest in the use of unconditional cash transfers as a means to alleviate poverty, yet little is known about the effects of such transfers in the U.S. This paper reports on the results of a randomized controlled study of a one-time $1,000 unconditional cash transfer to low-income households in the U.S., fielded in May 2020. Respondents were surveyed at the time of the transfer, and one and three months post transfer. We examine the impact of the cash transfer on five pre-registered outcomes (hardship, mental health, parenting, child behavior, partner relationships) and several secondary outcomes (hardship avoidance, consumption, employment, benefit use). We find no effects of the cash transfer on any outcomes for the full sample. In pre-specified exploratory analyses, we find robust evidence of reduced material hardship among those with the lowest levels of economic resources. These impacts are strongest among those with less than $500 of earnings in the previous month, roughly the bottom 50 percent of the study sample.

The U.S. social safety net broadly defined has expanded in the past few decades, but access to cash assistance for low-income families in crisis has declined dramatically since the 1990s (Parolin, 2021), resulting in a major gap in the safety net, particularly for the poorest households (e.g., Edin & Shaefer, 2015; Paxson & Waldfogel, 2003; Shaefer et al., 2020). As a result, there has been growing interest in unconditional cash transfers as a tool to fight poverty (Gennetian et al., 2021; National Academies of Science, 2019).

Despite widespread interest in cash transfers as a means to alleviate poverty and hardship, in the U.S. context, many questions regarding how to design and allocate these transfers, and their causal impacts, remain unanswered. Few, if any, rigorous experiments have been conducted in the U.S. on the impact of one-time cash transfers. Most studies of unconditional cash transfers have been conducted outside the U.S.; these experiments often involve very large transfers and recurring payments that may not translate to the U.S. setting (e.g., Bastagli et al. 2016; Davis & Handa 2015). Understanding how one-time unconditional cash transfers affect families in the U.S. can help inform future potential cash transfer policies.

This paper details the results of a randomized controlled study of a one-time unconditional cash transfer to low-income families with children that occurred in May 2020. In response to the onset of the COVID-19 pandemic in the U.S. in March of 2020, the charitable organization GiveDirectly (GD) began providing low-income individuals who were receiving or had recently received federal food assistance (Supplemental Nutrition Assistance Program – SNAP – benefits) a one-time, lump-sum, unconditional cash transfer of $1,000. The goal of the program was to help mitigate the financial distress associated with the COVID-19 pandemic, and as a one-time transfer, the payment did not impact recipients' eligibility for public assistance.

Despite sufficient statistical power (the experiment was powered to detect effect sizes of 0.09 standard deviations), we find no statistically significant effects of the $1,000 cash transfer among the full study sample on any of our five pre-specified primary outcomes (see AEA registry, https://doi.org/10.1257/rct.5852-1.0): material hardship, mental health challenges, partner conflict, child behavior problems, or parenting behaviors. In pre-specified secondary, or exploratory, analyses we find robust evidence of reduced material hardship among those at the lower end of the income distribution (particularly among those with less than $500 of earnings in the previous month, roughly the bottom 50 percent of the sample). The null effects are not explained by behavioral responses (reduced hardship avoidance techniques, changes in consumption, labor supply, or public benefit use). Nor do the results appear to be driven by the sensitivity of the measures of hardship.

This transfer took place amid unprecedented circumstances, both in terms of a global pandemic and an enormous influx of federal cash transfers to buffer families from poverty (Parolin et al., 2020; Han, Meyer & Sullivan, 2020). Many studies show that the government transfers (e.g., stimulus payments or expanded unemployment insurance) increased total income among low-income families, buffering families against hardships and declines in expenditures (Bachas et al., 2020; Chetty et al. 2020; Evangelist, Wu & Shaefer, 2021). The result of this support was that many households with low incomes experienced a relatively large influx of cash in the early months of the pandemic, when our study took place, complicating the interpretation of our findings. Nonetheless, this study provides some of the first evidence of a one-time, unconditional cash transfer for low-income families in the U.S. (two additional studies are in progress: Hauser et al., 2020 and a second study of this intervention [Jacob, Pilkauskas, Rhodes, Richard & Shaefer, 2022]).

Our findings offer insights into the efficacy of one-time cash transfers, which were used extensively by the federal government and private charities during the COVID-19 pandemic. We also add to a body of research that concludes the lowest income families in the U.S. benefit the most from additional cash transfers. It is possible that a $1,000 cash transfer is not large enough to move the outcomes we measured, and that is why we do not observe impacts in the full sample. It is also possible that the effect of this intervention was diluted because families with low incomes in the U.S. were broadly exposed to an unprecedented level of government cash transfers in the months surrounding our study. Finally, it is possible that recurring cash transfers are more effective in impacting the outcomes we measure than a one-time cash transfer. While our results do not allow us to arbitrate between these possibilities, that the $1,000 cash transfer did not reduce hardship or improve mental health among the full sample of low-income families surprised us, and we think means considerably more research is warranted to understand the impacts of such policies.

The rest of the paper proceeds as follows. In the background section we summarize the literature on cash transfers, briefly discuss the U.S. economic context during our study, and provide some benchmarking for anticipated effects. We then describe the experimental design, data, and empirical strategy. We follow with our results and conclusion.

BACKGROUND

Cash Transfers and Wellbeing in the U.S.

In this section we discuss the literature on cash transfers, distinguishing between lump-sum and recurring as well as conditional versus unconditional payments. Our overview focuses largely on the U.S., as transfers in low and middle income countries tend to be larger relative to average household income than is likely feasible in the U.S., and social policy contexts are quite different; however, we note that an extensive literature has examined the effects of unconditional and conditional cash transfers outside the U.S. context (see Bastagli et al., 2016 for a review; Banerjee et al., 2015; Paxson & Shady, 2010; Haushofer & Shapiro, 2016; Haushofer et al., 2020).

The U.S. social safety net relies heavily on *in-kind transfers* such as Supplemental Nutrition Assistance Program (SNAP or food stamps), Medicaid, or housing assistance. Studies of the U.S. social safety net programs find that these in-kind transfers reduce material hardship (Shaefer & Gutierrez, 2013; McKernan, Ratcliffe & Braga, 2021), increase consumption (Jacob, Kapustin & Ludwig, 2015) and can improve long-term wellbeing (e.g., Hoynes, Schanzenbach & Almond, 2016). However, in-kind transfers may lead to unintended changes in behavior as families over-consume based on what benefits are available (Hammond & Orr, 2016), and are typically more expensive to administer (Haushofer & Shapiro, 2016; Wiederspan et al., 2015). Unlike in-kind transfers, cash is fungible and allows families more freedom to efficiently allocate benefits to address their specific needs (Shaefer et al., 2018). Moreover, an advantage of cash is that it can reach individuals quickly, while other forms of assistance (e.g., housing) may require additional infrastructure that delays aid (Gennetian et al., 2021).

Where the U.S. government does provide cash, such as with Temporary Assistance for Needy Families (TANF) or refundable tax credits, aid is typically *conditioned on work* or seeking employment. One of the key cash transfer programs in the U.S. is the Earned Income Tax Credit (EITC), a *lump-sum cash transfer* provided annually to lower income working families. Most studies of the EITC rely on quasi-experimental methods, although two experimental studies of the EITC found that an additional $1,500 improved the mental and physical health for women (Courtin et al., 2020; Courtin et al., 2021). Other quasi-experimental studies find the EITC is linked with reduced housing hardship (Pilkauskas & Michelmore, 2019), reduced medical hardship (Kondratjeva et al., 2021), and improved health (e.g., Braga, Blavin, & Gangopadhyaya, 2020; Evans & Garthwaite, 2014). A recent study exploiting within-year timing of EITC receipt found no short-term changes in health and psychological distress (Collin et al., 2020), while another found a small, short-term reduction in food insecurity (Batra & Hamad, 2021). Delays in EITC receipt have also been linked with increased food insecurity (Kondratjeva et al., 2022) and delayed consumption (Aladangady et al. 2022). Although these studies are very informative, the EITC differs from the one-time transfer provided by GiveDirectly in that it is a highly anticipated refund administered through the tax system that families rely on each year to relieve financial distress (Sykes et al., 2015). While a related study of a one-time *conditional* cash transfer found emergency provision of money for housing reduced the likelihood of homelessness (Evans, Sullivan & Wallskog, 2016), we are unaware of any experimental studies that test the impact of a ***one-time lump-sum unconditional*** cash transfer to low-income families with children in the United States (although one study is currently in progress [Hauser et al. 2020], as is a second study of this program [Jacob et al., 2022]).

In contrast to one-time transfers, several studies have examined the effects of ***recurring unconditional cash transfers*** in both the U.S. and Canada. A recent randomized controlled study examined the impacts of recurring cash transfers on infants' brain activity and found some evidence that children in the treatment group had more brain activity in areas correlated with language, cognition and socio-emotional wellbeing (Troller-Renfree et al., 2022). The Negative Income Tax experiments from several decades ago also found regular transfers were linked with many positive outcomes, including improvements in health, educational outcomes, and homeownership (e.g., Maynard & Murnane, 1979). Similarly, quasi-experimental studies of casino dividends disbursed among U.S. tribes show that receipt of the cash payments is correlated with improved health, education, and parenting outcomes (e.g., Akee et al. 2010; Wolfe et al. 2012). The Canadian child benefit – a monthly cash payment to families with children – improved child and maternal health, children's test scores, and food security (Milligan & Stabile, 2011; Jones, Milligan & Stabile, 2019). Lastly, two recent experiments in the U.S. consider the effects of *recurring conditional* cash transfers (the Family Rewards Program) over three years and found reductions in poverty and hardship (Riccio & Miller, 2016), especially for those who were in "deeper" poverty prior to entry into the experiment (Miller et al., 2016).

In sum, a number of evaluations using associational, quasi-experimental, and experimental methods find that cash or near-cash transfers in a variety of forms can improve the material wellbeing of poor families across a number of domains. Although few experimental studies exist, studies also suggest that increased economic wellbeing improves parenting practices, child behavior, and relationship quality (e.g., Papp, Cummings & Goeke-Morey, 2009; Duncan and Brooks-Gunn, 1997; Schneider, Waldfogel and Brooks-Gunn, 2015). What is less clear is the extent to which a *one-time unconditional* cash transfer of modest size will improve wellbeing in the U.S. The current study addresses this policy-relevant question.

The COVID-19 pandemic, U.S. economic recession, and the economic stimulus

The cash transfer we study was provided amid unprecedented circumstances in May 2020, shortly after the COVID-19 pandemic hit the U.S. in March 2020 (see Appendix Figure 1 for a full timeline of our study and federal policy responses). In late March 2020, Congress passed a bill that provided families with one-time stimulus payments of $1,200 per eligible adult and $500 for each dependent under 17. From April-July 2020, the bill also temporarily expanded Unemployment Insurance (UI) to cover workers who were previously ineligible (including self-employed, contractors, and gig workers) and provided an additional $600 per week to supplement state UI benefits. To put the unprecedented scale of federal cash assistance that coincided with GiveDirectly's cash transfer into perspective, consider the case of a single mother with two children. She would have received a $2,200 stimulus payment and, if eligible for UI, an average of roughly $1000/week. Thus, the $1,000 study transfer arrived at a time when many members of our sample had recently received the stimulus payment plus the expanded UI – a historically large amount of cash assistance that likely impacted our ability to detect the effects of our cash transfer.

What Effect Should We Expect of a $1,000 Unconditional Cash Transfer?

Although no research to date has examined the impacts of a one-time $1,000 unconditional cash transfer in the U.S., theory and prior research provide some insight into what we might anticipate.[1] Research examining the relationship between income and hardship among current and former welfare recipients found that a 10% increase in average income (roughly $1,900 in the study sample) was associated with a 1.1 percentage point decrease in hardship, or a 3.4% reduction (Sullivan, Turner & Danziger, 2007). The literature on the EITC finds that $1,000 (approximately 4% of annual income) is associated with a very small decline (less than 1% of a standard deviation) in short-term food insecurity (Batra & Hamad, 2021), whereas another study found that a $1,000 increase in the EITC reduced moving in with others by 15% but had no effect on homelessness (Pilkauskas & Michelmore, 2019). The EITC is linked with few short-term effects on mental health (Collin et al 2020) but positive longer-term effects on mental health (Evans & Garthwaite, 2014). These studies suggest that cash reduces hardship (and possibly mental health challenges), but the effects differ depending on type of hardship, and may not be large.

Anticipating the effects of the study transfer is challenging given that it was disbursed during an unprecedented pandemic in which direct cash transfers were the centerpiece of the government's response. Nonetheless, we anticipate that a one-time $1,000 unconditional cash transfer is likely to improve families' economic wellbeing and reduce their experience of material hardship. Based on the size of the transfer and prior literature suggesting recurring transfers may be more effective at reducing hardship (Aguero, Carter & Woolard, 2006), we

[1] A review of cash transfer evaluations in sub-Saharan Africa found transfers of less than 20% of household consumption were associated with small effect sizes on a limited set of outcomes (Davis & Handa, 2015); transfers linked with better mental health, assets or consumption are typically very large (e.g., twice annual family income; Haushofer & Shapiro 2016).

expect the effects to be modest. We further posit, based on theory regarding the fungibility of cash, that the transfer is also likely to improve outcomes beyond reducing material hardship, but again we expect the effects to be modest.

In secondary analyses we also examine effects among lower income groups within our sample. Prior evidence suggests that effects of transfers (and other public programs) are often larger and more positive for the lowest income households (e.g., Evans, Sullivan & Wallskog, 2016). Although our transfer would be considered small on an annual level for the full sample (about 9% of income), for those with monthly earnings of less than $500 (about $110 on average), the transfer was closer to 75% of annual income. Although we expect the transfer to have positive impacts on the full sample, we anticipate a larger positive effect on outcomes for the households with lower incomes.

EXPERIMENTAL DESIGN

GiveDirectly (GD) launched Project 100+ in March 2020 with the ambitious goal of providing direct cash transfers to more than 100,000 low-income families struggling to make ends meet during the COVID-19 pandemic. In the first phase of the program, individuals were eligible to receive a $1,000 unconditional cash transfer if they were receiving federal food assistance through the Supplemental Nutrition Assistance Program (SNAP), lived in one of 15 states identified by GD as being particularly hard hit by the coronavirus, and were not receiving supplemental disability insurance (SSI or SSDI).[2] Our study was embedded in a wave of Project 100+ transfers that occurred in late May 2020. We employed the procedure GD used to distribute all Project 100+ transfers, with modifications to accommodate a control group.

To identify and recruit participants who were eligible for payments, GD relied on Fresh EBT (now called Providers), a free mobile application (app) that at the time had over 4 million users. Fresh EBT allows families receiving public assistance, primarily SNAP, to view and manage their benefits. Users of Fresh EBT were presented a banner inviting them to take a short survey, and if the user was deemed eligible based on their responses, they were selected to receive cash from GD and were provided with information on how to collect the award.

As part of its internal operations, Fresh EBT randomly assigns individual users to 1 of 1,000 different groups called segments, which they then use for various diagnostics and user outreach purposes (our own analyses also confirm it is consistent with random assignment). Segments were selected for the treatment and control groups (see Appendix A for more details on the experimental design) and the banner was only displayed to users that had accessed the app in the past 30 days (to ensure they were active users and likely still receiving benefits) and had not previously received a cash payment from GD. We excluded individuals living in three states (DE, NE, MA) where the usage of Fresh EBT was limited from our study.

On May 21, 2020, individuals in both treatment and control segments were shown an in-app banner inviting them to take a survey about the impact of the coronavirus on their family. Neither the banner nor the screening questions (seven identical questions asked of both treatment

[2] States included in this wave of GD transfers were LA, NH, ME, RI, MN, NC, OH, MI, PA, IN, NM, KY, DE, NE, and MA. GD enrolled individuals in other states as well and later included SSI and SSDI recipients.

and control groups) mentioned GD or the cash transfer, although some users were aware of the GD program and likely clicked on the banner in anticipation of the award. Screening questions allowed us to identify households with children, the focus of this study. Neither eligibility for the study nor willingness to share contact information with researchers affected receipt of the cash transfer. After indicating whether they were interested in participating in the research study, individuals in the control group were simply thanked for answering the short set of screener questions and individuals in the treatment group were given the opportunity to receive a $1,000 cash transfer from GD (through a Hyperwallet account).

We pre-registered this study with five primary outcomes at the AEA RCT Registry (https://doi.org/10.1257/rct.5852-1.0). As described in the pre-registration document, we powered the study to detect effect sizes of 0.09 standard deviations on the material hardship, mental health, child behavior and parenting behavior outcomes and 0.11 standard deviations on the partner relationship outcome. Our target enrollment was 10,000 individuals, and we assumed that 90 percent of the treatment group would take up the cash transfer and 60 percent of the full sample would respond to the one-month follow-up survey. We also described planned secondary analyses to look at treatment effect heterogeneity by baseline hardship (among other things), although we did not power the study to detect specific subgroup effects. Given that the entire sample was to be composed of low-income households, we expected to find effects in the main sample.

DATA

We use administrative data from Fresh EBT, screener data from the brief in-app survey, and data from three rounds of online survey data collection in this study. We briefly describe each data source here; more detail is available in Appendix A. From the Fresh EBT app, we obtained information on the individual's state, zip code, year of birth, preferred language (e.g., English or Spanish), date started using the app, average amount of time spent on the app, and the amount of their most recent SNAP benefit. Fresh EBT does not collect information on income or family size. We also received data from GD on when the money was transferred into the recipient's Hyperwallet account.

From the screener questions administered in Fresh EBT, we obtained self-reported information on whether the respondent experienced a loss of income due to the coronavirus, had increased expenses due to coronavirus, and had received a stimulus payment (although the first checks went out in April 2020, many people had not yet received them by late May). Two additional screener questions asked the respondent to rate how worried they were about being able to afford basic necessities over the next month and how difficult COVID-related restrictions had been for their children.

We administered three online surveys via Qualtrics: a baseline survey shortly after the respondents consented to participate in the study but before treatment, and follow-up surveys one and three months after receipt of the cash transfer. Invitations to complete the baseline survey were sent May 21, 2020, to 13,692 individuals (7,915 treatment, 5,777 control) before most treatment participants received the transfer (3% received it the first day, 90% cashed out the second day and 99.4% cashed out before the one-month survey). The baseline survey was completed by 9,433 individuals (68% of those contacted for the study). The one-month follow-up survey was completed by 6,848 individuals (a response rate of 60% treatment/59% control) and the three-month follow-up survey was completed by 5,774 individuals for an overall response rate of 57% (3,430 [59%] treatment, 2,344 [55%] control).

PRIMARY OUTCOMES

We pre-specified five primary outcomes that are linked with income and economic wellbeing in prior literature: material hardship, mental health, relationship with partner, child wellbeing, and parenting practices. We measure material hardship with a 9-item index that includes measure of housing hardship, food insecurity, and inability to pay utility bills, access medical care, or obtain needed transportation (following other studies, e.g., Iceland, Creamer & Kovach, 2021; Pilkauskas, Currie & Garfinkel, 2012; Rodems & Shaefer, 2020). Mental health challenges are measured by an index that includes validated, reliable measures of anxiety and depression (e.g., Levis et al., 2020). Partner conflict (among those who are married or living with a partner) is assessed with a 6-item scale that includes items that measure physical aggression, sexual coercion, psychological aggression, and controlling behaviors as well as positive relationship quality and emotional support. Child behavior problems are assessed with 11 items adapted from the Child Behavior Checklist (Achenbach & Rescorla, 2000; four externalizing, four internalizing, two prosocial, and one sleep). Last, parenting problems are assessed with four items, two that pick up on harsh parenting and two that capture positive parenting behaviors. For all outcomes we create composite measures, following the approach outlined in Anderson (2008) to calculate a weighted mean of the standardized survey items within the domain. All items are coded so that the positive direction indicates “worse” outcomes. We standardize each outcome using the mean and standard deviation of that outcome among the control group. Each composite can be interpreted as a z-score. Appendix A provides additional details on the measures in this study.

EMPIRICAL STRATEGY

Given the random assignment of the intervention, the estimation of treatment effects is straightforward. We estimate the Intent-to-Treat (ITT) effect of the program on outcome *Y* as follows:

$$Y_i = \beta_{ITT} Z_i + \overrightarrow{X_i} + \varepsilon_i$$

where *i* denotes individuals, *Z* is an indicator for treatment assignment, and *X* is a vector of covariates included to increase the statistical precision of our treatment effect estimates. We control for Fresh EBT utilization characteristics, answers to the screener questions, and state of residence (controls include all characteristics in Table 1). In addition, because respondents in the treatment group took the baseline survey shortly after being notified of the cash award, we only include information from the baseline survey that is time-invariant or unlikely to be influenced by a respondent’s knowledge they will imminently receive $1,000 (e.g., race/ethnicity, gender, age).

As shown in Table 1, our sample appears quite balanced on a large set of covariates including race and ethnicity, household composition, age, educational attainment, employment, and public assistance receipt. However, we find moderate imbalance between the geographic composition and Fresh EBT application usage of treatment and control users (standardized differences of 0.05 or more). Individuals in the treatment group had been using the app for a shorter period of time prior to the study (-0.078 SD) and logged onto the app less frequently (0.109 SD difference). Yet there were virtually no differences between groups on a wide range of other demographics and screener questions, such as the extent to which respondents worried about affording basic necessities. We determined that the imbalance is concentrated in a subset of our study sample that enrolled at specific times of the day. When enrollment opened, all segments assigned to treatment had been exposed to previous GD campaigns, compared to one-third of control segments. Efforts to correct for this imbalance meant that a set of treatment segments were exposed to the banner at a different time of day (see Appendix B for a detailed discussion of the enrollment issues). To determine whether this imbalance might influence our results, we calculate treatment effect estimates using a smaller "restricted" sample from segments that were not impacted by the study enrollment issues (i.e., restricting to those who were all previously exposed to GD campaigns). The results from this restricted sample are similar to those in the full sample, leading us to conclude that the small imbalance we find in the main sample is not substantially biasing our estimates (see Appendix Table 1 for restricted sample results).

We present intent-to-treat estimates, as nearly all (99%) of the treatment group members in our study received the cash transfer (only 137 people did not cash out or did so after the one month survey was fielded). In Appendix Table 2 we present the LATE (local average treatment effects) and find nearly identical results. While treatment was assigned at the level of the Fresh EBT user segment, this should not impact statistical inference as it might in most other contexts. Individual app users are randomly assigned to a segment and do not interact with other individuals in the segment, so we do not expect the outcomes of individuals within segments to be correlated in the absence of the intervention. When reporting individual treatment effect estimates, we present heteroskedasticity-robust standard errors with no clustering. For the full sample we present results at one and three months after the transfer. For all exploratory analyses we focus on the one month survey outcomes, as the larger sample and higher response rate provide more power to detect effects.

Because we examine the impact of the treatment on more than one outcome, we took several complementary approaches to adjust for multiple hypothesis tests. For our primary outcomes, we limited the number of tests by constructing composite outcome indices. When conducting inference on these indices, we control for the family-wise error rate (FWER) across our five primary outcome tests — the probability of rejecting at least one true null hypothesis — using the free step-down re-sampling method (Westfall & Young, 1993). We did not pre-specify adjustments for multiple hypothesis testing among exploratory outcomes, but to help contextualize the significance of the exploratory analyses we include Bonferroni p-values (adjusting for the number of tests in each table, see footnotes for details).

RESULTS

Full Sample Intent-to-Treat Estimates

Primary Outcomes. In Table 2 we present the results for our five primary (pre-specified) outcome measures from the one- and three-month follow-up surveys for the full sample of respondents: material hardship, mental health challenges, partner conflict, child behavior problems, and parenting problems. As described earlier, each outcome is a standardized composite index of survey items and should be interpreted in standard deviation units above or below the control mean. The sample differs across measures due to varying degrees of item missingness and over time as some respondents attritted between the one- and three-month surveys. In Appendix Table 3 we show the same estimates for a balanced panel (restricting the sample to respondents who participated in both the one- and three-month surveys) and find similar results.

We find no significant effects of the treatment on any of the composite outcomes in either survey wave; although the material hardship estimate is marginally significant, it is not robust to multiple hypothesis testing. While statistically insignificant, we note that point estimates are negative, in the direction of improved outcomes, except those on child behavior at three months post transfer.

Components of Hardship. If the transfer affects particular types of hardship more than others, the index might obscure those effects. Certain types of material hardships might be more responsive to a one-time cash transfer than others or individuals may use a lump sum payment to address some types of hardships but not others. Thus, in our pre-analysis plan we outlined a set of secondary analyses examining differences in the effect of the transfer by type of material hardship. In Table 3, we show the effects of the cash transfer on each of the individual hardship items at the one-month survey. We find that the transfer was associated with a marginally statistically significant reduction in the likelihood of experiencing both a utility shut off (and most estimates were negatively signed), but the findings are not robust to corrections for multiple hypothesis testing.

Treatment Effects Among Low-Income Groups

Although we anticipated that the cash transfer would improve the material wellbeing of families in our full sample, we also pre-specified secondary analyses examining the effects of the transfer among lower income groups because numerous prior studies find that cash transfers have bigger effects on the well-being of the lowest income households. The transfer represents a larger share of total income for those at the bottom of the income distribution, so we expect to see larger improvements in outcomes for this population. We focus our discussion on those with lower incomes here, but results for the higher income groups are available in Appendix Table 4 and depict largely null effects.

In Table 4 we show the effects of the cash transfer on the five primary outcomes at the one-month survey using four different measures of economic disadvantage. Our goal is to examine the effects of the cash transfer on the most disadvantaged households in the study sample, considering different potential operationalizations of economic disadvantage. Our primary specification relies on a measure of present income, restricting to those who earned less than $500 in the previous month (mean $110). This represents roughly the bottom 50 percent of our study sample. Current monthly earnings were obtained in the one-month survey and could, in theory, be affected by the intervention, but supplemental analyses found no effect of the treatment on monthly earned income and no correlation between income and inclusion in the study. We also examine results based on 2019 household income (median $13,000), restricting our analysis to those reporting annual income below the median. Finally, we consider two additional proxies for financial disadvantage both measured at the baseline (before the transfer): the effects among those who were unemployed or not in the labor force, and the effects among those receiving public assistance other than SNAP—WIC, TANF or housing assistance—because recipients of these programs generally come from households with low incomes that experience high rates of hardship.

Among respondents with less than \$500 in earnings the previous month, we find that receipt of the \$1,000 cash transfer was associated with a 0.166 standard deviation (SD) decrease in the composite material hardship measure (significant at $p<0.001$). We also find that the transfer reduced mental health challenges by 0.076 SD at the one-month survey for this group; however, this effect is not robust to multiple hypothesis testing. For partner conflict, child behavior problems, and parenting problems we observe no effects that are robust to multiple hypothesis testing.

When we look at the alternative approaches to identifying lower income groups, we see a similar pattern of results. The transfer significantly reduced the incidence of material hardship among those who were not employed and who were on public assistance. The estimate is only marginally significant for those whose 2019 household income was below the median, but the effect is nearly identical in magnitude to the estimate for the sample that was not employed/receiving public assistance. Although we find negatively signed effects of the transfer on mental health problems for these economically disadvantaged groups, none are statistically significant and the point estimates are much smaller than those for the sample of respondents with less than \$500 in monthly earnings. We observe a marginally significant decline in partner conflict (and negatively signed estimates across all groups) but this is not robust to multiple hypothesis testing. No significant effects are found for child behavior or parenting. We also conduct a supplemental analysis (see Appendix Table 5) examining the effect of the transfer on material hardship stratified by predicted level of hardship (endogenous stratification, following Abadie, Chingos & West, 2018). Consistent with the other analyses by economic groups, we find the transfer reduces hardship the most among the group predicted to have the highest level of hardship, although the effect is only marginally significant.

The effects of the transfer were most robust for the less than \$500 in earnings group. Thus, for the remaining analyses we also examine effects among this group as well as the full sample. For completeness, we show the individual hardship items (Appendix Table 6) and the three-month effects (Appendix Table 7) for the less than \$500 group. Although all estimates are negatively signed, we find the transfer significantly improved respondents' ability to pay rent, reduced utility shut-offs, and reduced medical hardship. We also find that the transfer significantly reduced material hardship three-months post-transfer (-0.111 SDs) and reduced mental health challenges (a marginally significant decline of -0.101 SDs after correcting for multiple hypothesis testing).

Behavioral Responses

Although we find some evidence that the transfer reduced material hardship among those with the lowest levels of economic resources, we expected to observe effects in the full sample. In this section we explore whether other behavioral responses might explain the null results for material hardship in the full sample, a set of secondary analyses we outlined in our pre-analysis plan. We also present the results for those with earnings under $500 to examine if the behavioral responses differ for the group for whom we observed the largest treatment effect.

Hardship Avoidance Techniques

Research shows that low-income families regularly employ strategies, such as borrowing money from friends and family or using food banks, to avoid experiencing hardships. One reason why we did not observe a decline in material hardship in the full sample might be because families in our sample engaged in fewer hardship avoidance strategies after receiving the cash transfer, essentially offsetting the transfer. To assess this possibility, we asked respondents whether they sold something, cut back spending, or used savings (or a retirement account) *to make ends meet*. We also asked if they borrowed money (from friends/family or from a financial institution) or put more on a credit card to help pay bills, though only 26% reported having a credit card. Lastly, respondents were asked if they had received free food from food banks. Table 5 displays estimates for the effects of the transfer on these hardship avoidance strategies.

We find no effects of the transfer on hardship avoidance techniques in the full sample. In the sample with monthly earnings of less than $500, we find that respondents were less likely to use food banks (4 percentage points) in response to the transfer, but no other statistically significant effects. These behavioral changes are unlikely to explain the null effects in the full sample, as the only significant behavioral change we observe is a decline in food pantry use for the lower income group and yet we still see a reduction in material hardship among this group. In a supplemental analysis we examined if the transfer affected ownership of cars or computers (considering the selling of assets as a measure of hardship avoidance) and found no significant results.

Use of Public Assistance

Similar to changes in hardship avoidance behaviors, it is possible that families who received the transfer substituted it for public assistance, offsetting any positive impacts on material hardship. In Appendix Table 8, we find little evidence this was the case; in fact, we find the transfer was associated with a statistically significant increase in the likelihood of being on SNAP. This was the case in both the full sample and the lowest monthly earnings group.

Labor Supply

A frequent concern raised in the literature on unconditional cash transfers is the potential for negative impacts on labor supply. Although it is not clear we should expect a one-time $1,000 transfer to impact employment on the extensive margin, it could potentially affect employment at the intensive margin, or through reduced job search activities in the short-term. If transfer recipients substituted the cash for earnings and worked fewer hours, leaving total income unchanged, we would be unlikely to observe any improvements in material hardship. However, we find no significant associations between the transfer and any of the labor supply measures (employment, hours worked, or hours spent looking for work) for either the full sample or the lower earnings group (see Appendix Table 9).

Consumption – How Did Respondents Use the Cash Transfer?

In addition to the primary outcome measures, the one month follow up survey included questions about consumption and a qualitative question on how the treatment group spent the cash transfer. Even if the transfer did not impact primary outcomes in the full sample, it may have affected consumption patterns.

In Table 6, we show the results from the consumption questions batched into larger categories for the full sample and the group with less than $500 in earnings the previous month. The outcomes are fractions of items within a category on which respondents report spending any money in the past 30 days (e.g., 1 = spending in all categories, 0.5 = spending in 50% of categories). In the full sample, we find no statistically significant effects. In the group with earnings of less than $500, we observe marginally significant increases in nonessential nondurable spending, durable spending, and paying bills, but none of these findings are robust to multiple hypothesis testing. Although we followed measurement approaches taken in other surveys, high quality consumption data are challenging to obtain in a short retrospective survey instrument. Additionally, the fungible nature of cash means that individuals may spread their spending over many categories, making it difficult to detect effects. In an open-ended question asked of those who received the transfer, most reported using the money for basic necessities (see Appendix Table 10): 68% mentioned paying bills, 28% indicated that they spent the transfer on food, and 27% used it to pay rent.

Hardship Timing and Intensity

One explanation for the lack of full sample findings could be that the hardship measures were insufficiently sensitive to capture the exact timing or intensity of material hardship. Thus, between the one- and three-month surveys, we developed additional hardship items that considered the timing and intensity of hardships.

Timing

To address the timing of material hardships, we asked three-month survey respondents to indicate whether they experienced the hardship (e.g., inability to pay rent) in each month between February (pre-pandemic) and July (the transfer occurred in May). Appendix Table 11 shows the full sample analyses and Appendix Table 12 shows the results for respondents with less than $500 in earnings. We find no significant results for the full sample. In the lower income sample, results suggest that hardships were most likely to be reduced in the months around the transfer (April, May, June). Thus, it appears that respondents are generally able to recall the timing of their hardships, suggesting this is not likely driving the null results in the full sample.

Intensity

We may not have observed reductions in material hardship in the full sample because our measures of hardship were not sufficiently sensitive to detect effects. For example, if a respondent paid off some back owed rent, they may still experience an inability to pay the *full rent* that month even though they were able to pay *some rent*. To test this theory, we added questions about back owed debt (e.g., back rent or back utilities) to the three-month survey. In

Appendix Table 13 we show the effect of the treatment on reporting having a particular type of back owed debt (and, conditional on having that debt, the amount of the debt) for both the full sample and the respondents with less than $500 in earnings.

We find some evidence that the treatment is linked with *more* types of back owed debt – to friends/family, and the total number of back owed debts (but not the amount) in the full sample. In contrast, among respondents who earned less than $500 in the prior month, we find no significant associations between the treatment and owing different types of back debt; however, among those with rent debt, we observe a reduction in back owed rent/mortgage and an increase in medical debt (among those with medical debt). Our hypothesis that families are paying down back owed debts was not supported.

CONCLUSION

Prior research has found that higher income is linked with reduced material hardship, fewer mental health problems, improved parenting, improved child behavior, and better relationship quality. Thus, we hypothesized that a one-time $1,000 unconditional cash transfer would improve all of these outcomes, if only modestly. Although recipients expressed that the money was incredibly important to them in open ended response questions, we detected no statistically significant impacts on any of our primary outcomes for the full study sample. Consistent with work in other areas (e.g., Evans et al., 2016), we did find evidence that the $1,000 cash transfer reduced material hardship among those at the lower end of the income distribution, most robustly among those with less than $500 in family earnings in the previous month. The transfer reduced the likelihood of not paying rent by 13% and of having utilities shut off, not paying phone/internet bills, or experiencing medical hardships by 10%. The effects on this lower income group may reflect the relative size of the transfer; $1,000 represented a nearly 75% increase in annual income for the lowest income group, compared to an 8% increase on average for the full sample. This finding is in keeping with research from low and middle income countries that found that large unconditional cash transfers (relative to income) have large effects, but smaller transfers are less likely to improve outcomes (Haushofer et al. 2020).

Nonetheless, we were surprised by the null effects in the full sample given that all respondents' income was low enough to qualify for public assistance, nearly 100% of individuals assigned to the treatment group were treated (received the $1,000), and the study was well powered to detect effects. To better understand our findings, we explored a number of potential explanations for the full sample null results. We found little evidence that the transfer affected hardship avoidance behaviors (e.g., selling items, borrowing money from friends/family), use of public assistance, labor supply, or consumption. We also hypothesized that perhaps the hardship measures were not sufficiently sensitive to assess both the timing and depth of hardships experienced by respondents. We developed some new measures to test this hypothesis in the three-month survey, but we find little evidence that suggests that families could not recall the timing of hardships. Our hypothesis that the treatment group might put some money towards many bills without completely wiping out debt or hardship was not supported.

This study has some limitations. It was conducted on a sample of families with low incomes, all of whom were currently on, or had recently been on, SNAP (or Pandemic-EBT). This is a population of particular interest to policymakers and those who study poverty; however, our sample is not representative of all families on SNAP. The study was conducted in only 12 states and recruited participants through a mobile app. Parents who are less tech savvy or who have limited access to the internet may have been less likely to participate, and respondents who opted into the study may differ in unobservable ways from those who did not.

The randomization process also met with a few challenges (see Appendix B), but findings for the restricted sample unaffected by randomization issues generally confirm findings for the full sample. Perhaps the biggest limitation of this study is that the cash transfers were disbursed at the beginning of a global pandemic that coincided with an enormous hike in unemployment and large cash transfers delivered by the federal government including stimulus checks and dramatically increased unemployment insurance eligibility and generosity. Thus, it is hard to know the extent to which our findings were influenced by the social and economic environment, or other seasonal variation, and to predict the degree to which our findings might generalize beyond the extraordinarily unique context in which the study was conducted.

Despite the limitations, this study has several policy implications and points to important areas for future research on unconditional cash transfers. A number of studies show that COVID-era cash transfer payments improved the financial wellbeing of many Americans (e.g., Chetty et al. 2020). Similar to the stimulus payments, the cash transfer in this study was a one-time lump sum transfer; however, unlike the stimulus payments, it was unanticipated, which may have affected spending decision patterns. Although recipients of the cash reported using the money to pay bills and buy food, we did not see discernable patterns to corroborate this in questions about household expenditures. We used standard expenditure questions, but the fungible nature of cash makes it challenging to track and to differentiate from other streams of income, like the large government checks or UI payments, that many participants received around the same time. Policymakers wishing to understand the effects of unconditional cash transfer programs (such as the Child Tax Credit, Earned Income Tax Credit, or others) may need to employ new approaches to track expenditures (e.g., administrative data) and consider ways to better understand recipients' intended use of the cash transfer to improve outcome evaluation (Shaefer, Jacob, Pilkauskas, Rhodes & Richard, 2022, for additional areas for research).

Additionally, while the transfer in this study was only provided to some families, stimulus payments and refundable tax credits are typically more universal. It is plausible that a narrowly targeted transfer may be less effective in reducing hardship if individuals receiving the transfer are embedded in social networks that include non-recipients with great needs. Further research is needed to better understand the diffusion of transfers across social networks and whether the degree of universality affects how recipients use the transfers and the impact of the transfer on outcomes.

We also found that respondents owed significant amounts of back debts (and many types of back debt), despite the significant federal influx of cash. Unlike student loans, for example, this form of debt is rarely discussed in policy-making circles, and our findings suggest that a one-time $1,000 transfer to poor households is not sufficient to meaningfully impact the wellbeing of these families. Future policies and programs interested in really improving economic wellbeing may need to provide larger and more sustained, recurring transfers. Much of the strongest evidence linking cash or near cash transfers to improved outcomes are of transfers with recurring payments. There may be differences in the impacts of one-time cash transfers versus recurring payments, even if they are of similar amounts.

In sum, our study sheds light on an important policy question – the impacts of an unexpected, one-time cash transfer to a low-income population of families. Although we find no effects of the cash transfer on material hardship, mental health, relationship with partner, child wellbeing, or parenting practices in the full sample, we do find that the transfer reduced material hardship for the most economically disadvantaged families in the study. Our results also highlight the need for more research to understand how the frequency of unconditional cash transfers (one-time vs. monthly or annually), transfer size, level of saturation, and whether the transfers are anticipated versus unanticipated affect wellbeing in the U.S.

REFERENCES

Abadie, A., Chingos, M. M., & West, M. R. (2018). Endogenous stratification in randomized experiments. *Review of Economics and Statistics*, 100(4), 567-580.

Achenbach, T.M., & Rescorla, L.A. (2000). Manual for the ASEBA Preschool Forms and Profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth & Families.

Aguero, J., Carter, M., & Woolard, I. (2006). The impact of unconditional cash transfers on nutrition: The South African Child Support Grant.

Akee, R., Copeland, W. E., Keeler, G., Angold, A., & Costello, E. J. (2010). Parents' Incomes and Children's Outcomes: A Quasi-Experiment Using Transfer Payments from Casino Profits. *American Economic Journal Applied Economics*, 2(1), 86–115.

Aladangady, A., Aron-Dine, S., Cashin, D., Dunn, W., Feiveson, L., Lengermann, P., Richard, K. , Sahmn, C., (2022) Spending Responses to High-Frequency Shifts in Payment Timing: Evidence from the Earned Income Tax Credit. *American Economic Journal: Economic Policy (Forthcoming).*

Anderson, M. L. (2008). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, *103*(484), 1481–1495.

Bachas, N., Ganong, P., Noel, P. J., Vavra, J. S., Wong, A., Farrell, D., & Greig, F. (2020). *Initial Impacts of the Pandemic on Consumer Behavior: Evidence from Linked Income, Spending, and Savings Data* (NBER Working Paper No. 27617). National Bureau of Economic Research. https://www.nber.org/papers/w27617.

Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., Pariente, W., Shapiro, J., … Udry, C. (2015). A Multi-faceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries. *Science,* 348(6236).

Bastagli, F., Hagen-Zanker, J., Harman, L., Barca, V., Sturge, G., Schmidt, T. & Pellerano, L. (2016). Cash transfers: what does the evidence say? *Overseas Development Institute.*

Batra, A., & Hamad, R. (2021). Short-term effects of the earned income tax credit on children's physical and mental health. *Annals of Epidemiology, 58*, 15-21.

Braga, B., Blavin, F., & Gangopadhyaya, A. (2020). The long-term effects of childhood exposure to the earned income tax credit on health outcomes. *Journal of Public Economics*, *190*, 104249.

Chetty, R., Friedman, J. N., Hendren, N., Stepner, M., & the Opportunity Insights Team. (2020). *How Did COVID-19 and Stabilization Policies Affect Spending and Employment? A New Real-Time Economic Tracker Based on Private Sector Data.* Opportunity Insights. https://opportunityinsights.org/wp-content/uploads/2020/05/tracker_paper.pdf.

Collin, D. F., Shields-Zeeman, L. S., Batra, A., Vable, A. M., Rehkopf, D. H., Machen, L., & Hamad, R. (2020). Short-term effects of the earned income tax credit on mental health and health behaviors. *Preventive Medicine*, *139*, 106223.

Courtin, E., Aloisi, K., Miller, C., Allen, H. L., Katz, L. F., & Muennig, P. (2020). The Health Effects Of Expanding The Earned Income Tax Credit: Results From New York City: Study examines the health effects of the New York City Paycheck Plus program that increases the Earned Income Tax Credit for low-income Americans without dependent children. *Health Affairs*, 39(7), 1149-1156.

Courtin, E., Allen, H. L., Katz, L. F., Miller, C., Aloisi, K., & Muennig, P. A. (2021). Effect of Expanding the Earned Income Tax Credit to Americans without Dependent Children on Psychological Distress (Paycheck Plus): a Randomized Controlled Trial. *American Journal of Epidemiology.*

Davis, B., & Handa, S. (2015). *How Much Do Programmes Pay? Transfer size in selected* national cash transfer programmes in sub-Saharan Africa. UNICEF Office of Research Innocenti Research Briefs No. 2015-05.

Duncan, G., & Brooks-Gunn, J., Eds. (1997). *Consequences of Growing Up Poor*. New York: Russell Sage.

Edin, K., & Shaefer, H. L. (2015). *$2.00 a Day: Living on Almost Nothing in America*. Boston, MA: Houghton Mifflin Harcourt.

Evans, W., & Garthwaite, C. (2014). Giving Mom a Break: The Impact of Higher EITC Payments on Maternal Health. *American Economic Journal: Economic Policy*, 6, 258-290.

Evans, W. N., Sullivan, J. X., & Wallskog, M. (2016). The impact of homelessness prevention programs on homelessness. *Science*, 353(6300), 694-699.

Evangelist, M., Wu, P. & Shaefer, H. L. (2021). Emergency unemployment benefits and health care spending during COVID. *Health Services Research*.

Gennetian, L. A., Shafi, E., Aber, J. L., & De Hoop, J. (2021). Behavioral insights into cash transfers to families with children. *Behavioral Science & Policy,* 7(1), 71-92.

Hammond, S., & Orr, R. (2016). *Toward a Universal Child Benefit.* Washington, D.C.: Niskanen Center.

Han, J., Meyer, B. D., & Sullivan, J. X. (2020). *Income and Poverty in the COVID-19 Pandemic (*NBER Working Paper No. 27729). National Bureau of Economic Research.

Hauser, O., Jachimowicz, J., Jamison, J., & Jaroszewicz, A. (2020).A randomized controlled trial varying unconditional cash transfer amounts in the United States." AEA RCT registry. July 21. https://doi.org/10.1257/rct.6149-1.0

Haushofer, J., Chemin, M., Jang, C., & Abraham, J. (2020). Economic and psychological effects of health insurance and cash transfers: Evidence from a randomized experiment in Kenya. *Journal of Development Economics*, *144*, 102416.

Haushofer, J., & Shapiro, J. (2016). The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya. *Quarterly Journal of Economics*, 131(4), 1973–2042.

Hoynes, H., Schanzenbach, D. W., & Almond, D. (2016). Long-run impacts of childhood access to the safety net. *American Economic Review*, *106*(4), 903-34.

Iceland, J., Kovach, C., & Creamer, J. (2021). Poverty and the Incidence of Material Hardship, Revisited. *Social Science Quarterly*, 102(1), 585-617.

Jacob, B. A., Kapustin, M., & Ludwig, J. (2015). The impact of housing assistance on child outcomes: Evidence from a randomized housing lottery. *The Quarterly Journal of Economics*, *130*(1), 465-506.

Jacob, B.A., Pilkauskas, N.V., Rhodes, E., Richard, K. and Shaefer, H.L. (2022). The COVID cash transfer study II: The hardship and mental health impacts of an unconditional cash transfer to low-income individuals. *Poverty Solutions Working Paper.*

Jones, L. E., Milligan, K., & Stabile, M. (2019). Child cash benefits and family expenditures: Evidence from the National Child Benefit. *Canadian Journal of Economics,* 52, 1433-1463.

Kondratjeva, O., Roll, S. P., Despard, M., & Grinstein-Weiss, M. (2021). The impact of state earned income tax credit increases on material and medical hardship. *Journal of Consumer Affairs,* 55(3), 872-910.

Kondratjeva, O., Roll, S. P., Despard, M., & Grinstein-Weiss, M. (2022). The Impact of Tax Refund Delays on the Experience of Hardship Among Lower-Income Households. *Journal of Consumer Policy*, 1-42.

Levis, B., Sun, Y., He, C., Wu, Y., Krishnan, A., Bhandari, P. M., et al. (2020). Accuracy of the PHQ-2 Alone and in Combination With the PHQ-9 for Screening to Detect Major Depression. *Journal of the American Medical Association*, *323*(22), 2290–11.

Maynard, R., & Murnane, R. (1979). Effects of a negative income tax on school performance: Results of an experiment. *Journal of Human Resources*, 14(4), 463–476.

McKernan, S. M., Ratcliffe, C., & Braga, B. (2021). The effect of the US safety net on material hardship over two decades. *Journal of Public Economics,* 197, 104403.

Miller, C., Miller, R., Verma, N., Dechausay, N., Yang, E., Rudd, T., et al. (2016). *Effects of a Modified Conditional Cash Transfer Program in Two American Cities*. New York: MDRC.

Milligan, K., & Stabile, M. (2011). Do Child Tax Benefits Affect the Well-being of Children? Evidence from Canadian Child Benefit Expansions. *American Economic Journal: Economic Policy*, 3(3), 175–205.

National Academies of Sciences, Engineering, and Medicine. (2019). *A Roadmap to Reducing Child Poverty*. Washington, DC: National Academies Press.

Papp, L., Cummings, E., & Goeke-Morey, M. (2009). For richer, for poorer: Money as a topic of marital conflict in the home. *Family Relations*, 58(1), 91–103.

Parolin, Z. (2021). Decomposing the decline of cash assistance in the United States, 1993 to 2016. *Demography, 58(3),* 1119-1141.

Parolin, Z., Curran, M., Matsudaira, J., Waldfogel, J., & Wimer C. (2020). Monthly *Poverty Rates in the United States during the COVID-19 Pandemic*. Center on Poverty and Social Policy, Columbia University.

Paxson, C., & Schady, N. (2010). Does Money Matter? The Effects of Cash Transfers on Child Development in Rural Ecuador. *Economic Development and Cultural Change*, 59, 187-229.

Paxson, C., & Waldfogel, J. (2003). Welfare reforms, family resources, and child maltreatment. *Journal of Policy Analysis and Management,* 22(1), 85–113.

Pilkauskas, N. V., Currie, J. M., & Garfinkel, I. (2012). The great recession, public transfers, and material hardship. *Social Service Review*, 86(3), 401-427.

Pilkauskas, N., & Michelmore, K. (2019). The Effect of the Earned Income Tax Credit on Housing and Living Arrangements. *Demography*, *56*, 1303–1326.

Riccio, J., & Miller, C. (2016). *New York City's First Conditional Cash Transfer Program: What Worked, What Didn't*. New York: MDRC.

Rodems, R., & Shaefer, H. L. (2020). Many of the kids are not alright: Material hardship among children in the United States. *Children and Youth Services Review*, 112(2020), 104767.

Schneider, W., Waldfogel, J., & Brooks-Gunn, J. (2015). The great recession and behavior problems in 9-year old children. *Developmental psychology*, *51*(11), 1615.

Shaefer, H. L., Collyer, S., Duncan, G., Edin, K., Garfinkel, I., Harris, D., et al. (2018). A Universal Child Allowance: A Plan to Reduce Poverty and Income Instability Among Children in the United States. *RSF: the Russell Sage Foundation Journal of the Social Sciences*, *4*(2), 22.

Shaefer, H. L., Edin, K., Fusaro, V., & Wu, P. (2020). The Decline of Cash Assistance and the Well-Being of Poor Households with Children. *Social Forces*, *98*(3), 1000–1025.

Shaefer, H. L., & Gutierrez, I. A. (2013). The Supplemental Nutrition Assistance Program and material hardships among low-income households with children. *Social Service Review*, 87(4), 753-779.

Shaefer, H.L., Jacob, B.A., Pilkauskas, N.V., Rhodes, E., and Richard, K (2022). The COVID cash transfer studies: Key findings and future directions. *Poverty Solutions Research Brief*.

Sullivan, J. X., Turner, L., & Danziger, S. H. (2007). The Relationship between Income and Material Hardship. *Journal of Policy Analysis and Management,* 27(1), 63–81.

Sykes, J., Križ, K., Edin, K., & Halpern-Meekin, S. (2015). Dignity and dreams: What the Earned Income Tax Credit (EITC) means to low-income families. *American Sociological Review*, *80*(2), 243-267.

Troller-Renfree, S. V., Costanzo, M. A., Duncan, G. J., Magnuson, K., Gennetian, L. A., Yoshikawa, H., ... & Noble, K. G. (2022). The impact of a poverty reduction intervention on infant brain activity. *Proceedings of the National Academy of Sciences*, 119(5).

Westfall, P. H., & Young, S. S. (1993). *Resampling-Based Multiple Testing: Examples and Methods for p-value Adjustment*. New York: John Wiley & Sons.

Wiederspan, J., Rhodes, E., & Shaefer, H.L. (2015). Expanding the Discourse on Antipoverty Policy: Reconsidering a Negative Income Tax. *Journal of Poverty,* 19(2), 218-238.

Wolfe, B., Jakubowski, J., Haveman, R., & Courey, M. (2012). The Income and Health Effects of Tribal Casino Gaming on American Indians. *Demography,* 49(2), 499–524.

Table 1: Covariate Balance of Control and Treatment Baseline Respondents

	Control Mean	Treament Mean	T-Statisic	Standardized Difference
Race/Ethnicity				
Hispanic	0.065	0.064	-0.189	-0.004
Non-Hispanic White	0.223	0.234	1.264	0.026
Non-Hispanic Black	0.650	0.629	-2.180	-0.046
Other race/ethnicity	0.026	0.029	0.804	0.017
Other Demographics				
Age	32.696	32.631	-0.426	-0.009
Female	0.943	0.941	-0.453	-0.009
Education				
Less than high school	0.115	0.121	0.965	0.020
High school	0.426	0.411	-1.467	-0.030
Some college	0.309	0.317	0.794	0.017
Associates or more	0.129	0.131	0.350	0.007
Employment				
Full-time	0.185	0.180	-0.660	-0.014
Part-time	0.120	0.122	0.393	0.008
Temporarily laid off	0.357	0.338	-1.972	-0.041
Unemployed	0.203	0.210	0.825	0.017
Other employment status	0.100	0.103	0.511	0.011
Public Assistance				
WIC	0.415	0.393	-1.855	-0.045
Housing	0.262	0.261	-0.074	-0.002
Unemployment Insurance	0.417	0.415	-0.184	-0.004
TANF	0.205	0.181	-2.436	-0.059
Household Composition				
Household size	4.072	4.040	-0.901	-0.019
Married	0.160	0.161	0.094	0.002
Cohabiting	0.134	0.145	1.402	0.030
Lives with own children	0.991	0.995	2.173	0.041
Lives with other children	0.090	0.088	-0.378	-0.008
Lives with other relatives	0.105	0.109	0.691	0.014
Lives with non-relatives	0.015	0.013	-0.875	-0.018
Number of children				
Under age 6	1.043	1.032	-0.515	-0.011
Ages 6-12	1.011	1.035	1.050	0.022
Ages 13-17	0.526	0.494	-1.817	-0.037
Total number of children	2.566	2.538	-0.908	-0.019
Application Data				
# of days using Fresh EBT	329.650	300.147	-3.799	-0.078
Joined before Feb 1, 2020	0.526	0.490	-3.453	-0.072
Joined on May 21, 2020	0.023	0.037	3.821	0.088
# of App Uses	204.094	175.899	-5.453	-0.109
SNAP benefit amount	352.146	354.169	0.333	0.008
Spanish language	0.016	0.011	-2.318	-0.045

Table 1: Covariate Balance of Control and Treatment Baseline Respondents

	Control Mean	Treament Mean	T-Statisic	Standardized Difference
State				
Indiana	0.030	0.035	1.183	0.025
Kentucky	0.024	0.034	3.074	0.070
Louisiana	0.042	0.046	0.918	0.019
Maine	0.002	0.001	-0.276	-0.006
Michigan	0.258	0.230	-3.126	-0.064
Minnesota	0.010	0.010	-0.191	-0.004
North Carolina	0.164	0.169	0.628	0.013
New Hampshire	0.001	0.002	0.247	0.005
New Mexico	0.003	0.006	1.805	0.043
Ohio	0.284	0.310	2.726	0.057
Pennsylvania	0.178	0.154	-3.059	-0.062
Rhode Island	0.004	0.004	0.030	0.001
Application Screener				
Lost income due to COVID	0.953	0.965	2.713	0.054
Expenses increased due to COVID	0.906	0.905	-0.170	-0.004
Received stimulus (CARES) payment	0.694	0.692	-0.195	-0.004
Worried about affording basic	4.559	4.580	1.426	0.029
How difficult has COVID been for	4.452	4.473	1.318	0.027
N	4069	5370		

Note: Data come from administrative records, initial eligibility screening and the baseline survey data. T-statistic comes from a bi-variate regressions of each variable on treatment status, with standard errors clustered by segment. Standardized difference divides the t-statistic by the standard deviation of the control group.

Table 2: Primary Outcome Intent-to-Treat Estimates in Full Sample at One and Three Months Post Cash Transfer

	Material Hardship	Mental Health Challenges	Partner Conflict	Child Behavior Problems	Parenting Problems
One-Month					
Treatment	-0.041+	-0.010	-0.071	0.011	-0.024
(SE)	(0.024)	(0.025)	(0.044)	(0.025)	(0.026)
Observations	6520	6299	2139	6375	6363
R2	0.106	0.060	0.049	0.067	0.026
Control Mean	0.002	0.000	0.005	-0.001	-0.001
Control SD	0.998	1.000	1.005	1.001	1.001
Bonferroni	0.448	1.000	0.448	1.000	1.000
FWER	0.385	0.880	0.385	0.880	0.694
Three-Month					
Treatment	0.001	-0.002	-0.078+	0.027	-0.043
(SE)	(0.026)	(0.028)	(0.043)	(0.028)	(0.028)
Observations	5530	5244	2051	5269	5266
R2	0.106	0.070	0.058	0.044	0.037
Control Mean	-0.002	-0.001	0.001	-0.000	-0.003
Control SD	1.002	0.999	0.996	1.004	1.000
Bonferroni	1.000	1.000	0.346	1.000	0.488
FWER	0.996	0.996	0.317	0.699	0.409

Note: Outcomes are standardized. Treatment is the $1,000 cash transfer. All analyses include the full set of controls. Standard errors are robust. Bonferroni and Free-step down resampling p-values correct for multiple hypothesis testing in a family of five tests. We resample 10,000 times to compute the Family Wise Error Rate (FWER) p-value.

+ $p < 0.10$, $*$ $p < 0.05$, $**$ $p < 0.01$, $***$ $p < 0.001$

Table 3: Effects of Treatment on Individual Hardship Items in Full Sample

	Not Pay Utilities	Utilities Shutoff	Not Pay Phone/ Internet	Not Pay Rent	Double-up, Shelter or Eviction	Worried Food Run Out	Food Did Not Last	Medical Hardship	Transportation Hardship
Treatment	0.002	-0.017+	-0.015	-0.018	0.006	-0.013	-0.010	-0.015	-0.005
(SE)	(0.013)	(0.010)	(0.013)	(0.013)	(0.007)	(0.012)	(0.013)	(0.010)	(0.012)
Observations	6422	6442	6450	6271	6510	6451	6250	6522	6536
Control Mean	0.591	0.223	0.562	0.484	0.070	0.609	0.486	0.206	0.335
Control SD	0.492	0.416	0.496	0.500	0.255	0.488	0.500	0.404	0.472
Bonferroni p-value	7.886	0.885	2.119	1.536	3.611	2.475	3.842	1.190	5.819

Note: Each hardship item is a binary yes/no variable, equal to one if a respondent experienced a given hardship at the one-month survey. Treatment is the $1,000 cash transfer. All analyses include the full set of controls. Standard errors are robust. Bonferroni p-values are computed for the family of nine tests.

\+ $p < 0.10$, $*$ $p < 0.05$, $**$ $p < 0.01$, $***$ $p < 0.001$

Table 4: Alternative Measures of Financial Wellbeing: Intent-to-Treat Effects on Main Outcomes One Month Post Cash Transfer

	Material Hardship	Mental Health Challenges	Partner Conflict	Child Behavior Problems	Parenting Problems
Monthly Earnings Less than $500					
Treatment	-0.166***	-0.076*	-0.166+	-0.026	-0.044
(SE)	(0.037)	(0.039)	(0.087)	(0.039)	(0.039)
Observations	2969	2843	722	2902	2885
Control Mean	0.248	0.098	0.135	0.040	-0.060
Control SD	1.042	1.048	1.120	1.049	1.032
Bonferroni	0.000	0.247	0.289	2.555	1.296
Less than Median 2019 Income					
Treatment	-0.067+	-0.026	-0.094	0.036	-0.012
(SE)	(0.038)	(0.037)	(0.091)	(0.037)	(0.037)
Observations	3134	3017	761	3060	3063
Control Mean	0.104	0.002	0.063	-0.041	-0.082
Control SD	1.071	1.028	1.136	1.011	1.021
Bonferroni	0.369	2.459	1.499	1.659	3.756
Not employed					
Treatment	-0.067*	-0.001	-0.045	0.004	-0.001
(SE)	(0.029)	(0.031)	(0.053)	(0.031)	(0.031)
Observations	4423	4284	1496	4336	4340
Control Mean	0.070	0.024	-0.025	0.006	-0.053
Control SD	1.006	1.008	0.975	1.009	0.963
Bonferroni	0.110	4.814	1.979	4.473	4.880
Receiving public assistance					
Treatment	-0.074*	-0.001	-0.018	0.049	-0.001
(SE)	(0.033)	(0.035)	(0.063)	(0.035)	(0.037)
Observations	3484	3357	1066	3400	3384
Control Mean	0.015	-0.019	-0.020	-0.033	0.003
Control SD	1.014	1.003	0.987	1.007	1.062
Bonferroni	0.129	4.860	3.891	0.822	4.939

Note: One-month survey respondents. Outcomes are standardized. Treatment is the $1,000 cash transfer. All analyses include the full set of controls. Standard errors are robust. Bonferroni p-values adjust for five tests in each sample split.

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Hardship Avoidance Intent-to-Treat Estimates, One Month Post Transfer, Full Sample and Monthly Earnings Less than $500

	Sold Something	Cut Spending	Used Savings	Borrowed from Family or Friends	Free food from bank or pantry
Full Sample					
Treatment	-0.005	0.006	0.006	-0.009	-0.015
(SE)	(0.012)	(0.009)	(0.011)	(0.012)	(0.011)
Observations	6559	6568	6535	6465	6760
Control mean	0.513	0.850	0.226	0.599	0.308
Control SD	0.500	0.357	0.418	0.490	0.462
Bonferroni p-value	1.000	1.000	1.000	1.000	0.877
Monthly Earnings Less than $500					
Treatment	-0.016	-0.018	-0.010	-0.021	-0.037*
(SE)	(0.018)	(0.013)	(0.014)	(0.018)	(0.017)
Observations	2982	2987	2974	2938	3048
Control mean	0.574	0.878	0.180	0.672	0.353
Control SD	0.495	0.328	0.385	0.470	0.478
Bonferroni	1.000	0.805	1.000	1.000	0.169

Note: One-month survey respondents. Outcomes are binary indicators. Treatment is the $1,000 cash transfer. All analyses include the full set of controls. Standard errors are robust. Bonferroni p-values are computed over a family of 5 tests for each subsample.

+ $p < 0.10$, $*$ $p < 0.05$, $**$ $p < 0.01$, $***$ $p < 0.001$

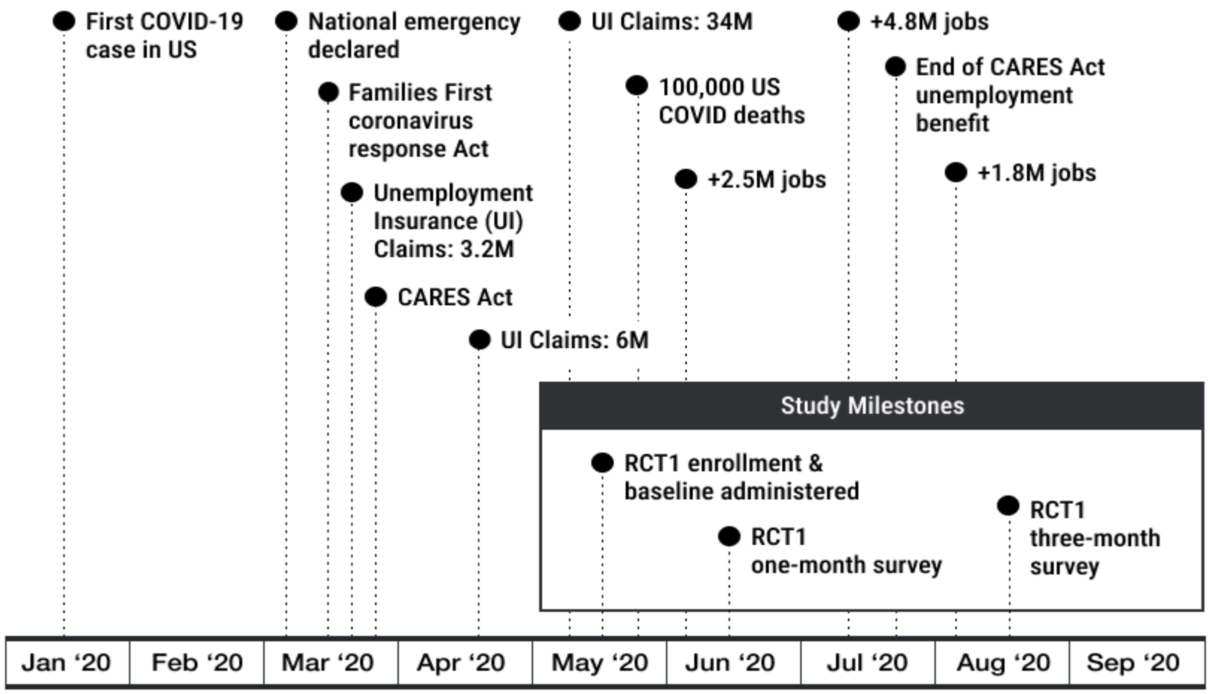
Table 6: Consumption Category Intent-to-Treat Estimates One-Month Post Transfer, Full Sample and Monthly Earnings Less than $500

	Nonessential Nondurables	Essential Nondurables	Durables	Spending on Kids	Paying Bills	Savings
Full Sample						
Treatment	0.006	0.006	0.012	0.010	0.009	0.002
(SE)	(0.006)	(0.007)	(0.007)	(0.009)	(0.007)	(0.008)
Observations	6848	6848	6848	6848	6848	6848
Control Mean	0.307	0.841	0.339	0.305	0.693	0.230
Control SD	0.273	0.278	0.301	0.358	0.278	0.321
Bonferroni	1.00	1.00	0.636	1.00	1.00	1.00
Monthly Earnings Less than $500						
Treatment	0.017+	0.017	0.019+	0.012	0.021+	0.003
(SE)	(0.009)	(0.011)	(0.011)	(0.013)	(0.011)	(0.010)
Observations	3084	3084	3084	3084	3084	3084
Control Mean	0.253	0.818	0.301	0.292	0.639	0.166
Control SD	0.259	0.298	0.297	0.353	0.303	0.279
Bonferroni	0.429	0.649	0.528	1.00	0.308	1.00

Note: One-month survey respondents. All analyses include the full set of controls. Treatment is the $1,000 cash transfer. Standard errors are robust. Bonferroni p-values are computed over a family of 6 tests. Each outcome is a the fraction of items in a category that one-month respondents spent any money on in the past thirty days. If an item has a missing response we assume they had no spending on that category. Nonessential nondurables includes spending on restaurant food, alcohol, tobacco, entertainment and adult education. Essential nondurables includes spending on groceries, transportation and personal care goods. Durables includes spending on vehicle purchase, vehicle repair, household repair and maintenance, electronics, appliances, and furniture. Spending on kids includes childcare and children's education expenses. Bills includes payments for rent, mortgage, medical expenses, phone, internet, and utilities. Savings refers to paying down debt and depositing money in savings or retirement accounts.

+ $p < 0.10$, $*$ $p < 0.05$, $**$ $p < 0.01$, $***$ $p < 0.001$

Appendix Figure 1: Timeline of U.S. events and study milestones.



Appendix Table 1: Primary Outcome Intent-to-Treat Estimates in Restricted Sample at One and Three Months Post Cash Transfer

	Material Hardship	Mental Health Challenges	Partner Conflict	Child Behavior Problems	Parenting Problems
One-Month					
Treatment	-0.054	-0.006	-0.134+	-0.048	-0.078+
(SE)	(0.038)	(0.041)	(0.071)	(0.041)	(0.043)
Observations	2956	2853	1024	2887	2881
R2	0.132	0.067	0.084	0.083	0.038
Control Mean	0.005	0.002	0.041	0.073	0.073
Control SD	0.987	0.974	0.985	1.000	1.029
Bonferroni	0.484	0.876	0.290	0.486	0.290
FWER	0.442	0.874	0.303	0.442	0.303
Three-Month					
Treatment	-0.004	0.004	-0.112+	-0.041	-0.127**
(SE)	(0.041)	(0.045)	(0.065)	(0.047)	(0.047)
Observations	2437	2314	935	2329	2322
R2	0.129	0.092	0.092	0.056	0.052
Control Mean	-0.020	0.016	0.011	0.067	0.088
Control SD	0.956	0.978	0.930	1.031	1.022
Bonferroni	1.000	1.000	0.336	1.000	0.034
FWER	0.993	0.993	0.326	0.763	0.054

Note: "Restricted" sample that was unaffected by any implementation challenges. Outcomes are standardized. Treatment is the $1,000 cash transfer. All analyses include the full set of controls. Standard errors are robust. Bonferroni and Free-step down resampling p-values correct for multiple hypothesis testing in a family of five tests. We resample 10,000 times to compute the Family Wise Error Rate (FWER) p-value.

+ $p < 0.10$, $* \ p < 0.05$, $** \ p < 0.01$, $*** \ p < 0.001$

Appendix Table 2: Primary Outcome Local Average Treatment Effects in Full Sample at One and Three Months Post Cash Transfer

	Material Hardship	Mental Health Challenges	Partner Conflict	Child Behavior Problems	Parenting Problems
One-Month					
Indicator if User Claimed GD Transfer	-0.041+	-0.010	-0.073	0.011	-0.024
(SE)	(0.024)	(0.025)	(0.044)	(0.026)	(0.026)
Observations	6520	6299	2139	6375	6363
Control Complier Mean	0.006	0.006	0.005	-0.002	0.000
Bonferroni	0.448	1.000	0.448	1.000	1.000
FWER	0.382	0.879	0.382	0.879	0.689
Three-Month					
Indicator if User Claimed GD Transfer	0.001	-0.002	-0.079+	0.027	-0.043
(SE)	(0.026)	(0.028)	(0.043)	(0.028)	(0.028)
Observations	5530	5244	2051	5269	5266
Control Complier Mean	0.005	0.002	-0.002	0.000	-0.007
Bonferroni	1.000	1.000	0.346	1.000	0.488
FWER	0.997	0.997	0.324	0.703	0.413

Note: Sample of all users who responded to both the one and three month surveys. We estimate 2SLS models and instrument for treatment with an indicator equal to one if a user successfully cashed out their transfer. All models include our standard set of controls. Standard errors are robust. Control Complier Means are estimated by restricting to the subsample of respondents that did not cash out the transfer, running a 2SLS model of each outcome as described above, and saving the coefficient of the indicator equal to one if a user did not cash out their transfer. Bonferroni and Free-step down resampling p-values correct for multiple hypothesis testing in a family of five tests. We resample 10,000 times to compute the Family Wise Error Rate (FWER) p-value.

+ $p < 0.10$, $* \ p < 0.05$, $** \ p < 0.01$, $*** \ p < 0.001$

Appendix Table 3 : Primary Outcomes Intent-to-Treat Estimates in Full Balanced Panel Sample at One and Three Months Post Cash Transfer

	Material Hardship	Mental Health Challenges	Partner Conflict	Child Behavior Problems	Parenting Problems
One-Month					
Treatment	-0.046+	-0.03	-0.052	-0.009	-0.042
(SE)	(0.027)	(0.029)	(0.049)	(0.029)	(0.029)
Observations	4957	4701	1460	4753	4739
R2	0.12	0.066	0.062	0.076	0.029
Control Mean	-0.009	0.041	-0.049	0.022	0.003
Control SD	0.977	1.006	0.929	0.998	0.993
Bonferroni	0.233	1.000	0.325	1.000	1.000
FWER	0.213	0.816	0.283	0.896	0.697
Three-Month					
Treatment	-0.001	-0.011	-0.058	0.031	-0.034
(SE)	(0.028)	(0.029)	(0.046)	(0.029)	(0.029)
Observations	4957	4701	1460	4753	4739
R2	0.115	0.075	0.074	0.044	0.041
Control Mean	-0.004	0.005	-0.054	0.008	-0.002
Control SD	1.007	0.994	0.894	0.982	0.979
Bonferroni	1.000	1.000	0.652	1.000	0.652
FWER	0.969	0.969	0.517	0.829	0.517

Note: Sample is restricted to those who responded to both the one- and three-month surveys. Outcomes are standardized. Treatment is the $1,000 cash transfer. All analyses include the full set of controls. Standard errors are robust. Bonferroni and Free-step down resampling p-values correct for multiple hypothesis testing in a family of five tests, separately for the one and three month samples. We resample 10,000 times to compute the Family Wise Error Rate (FWER) p-value.

+ $p < 0.10$, $* \ p < 0.05$, $** \ p < 0.01$, $*** \ p < 0.001$

Appendix Table 4: Alternative Measures of Financial Wellbeing (Top) : Intent-to-Treat Effects on Main Outcomes One Month Post Cash Transfer

	Material Hardship	Mental Health Challenges	Partner Conflict	Child Behavior Problems	Parenting Problems
Monthly earnings more than $500					
Treatment	0.069*	0.045	-0.042	0.035	-0.010
(SE)	(0.031)	(0.033)	(0.051)	(0.033)	(0.034)
Observations	3551	3456	1417	3473	3478
Control Mean	-0.210	-0.082	-0.062	-0.037	0.050
Control SD	0.907	0.949	0.934	0.957	0.972
Bonferroni	0.116	0.863	2.050	1.414	3.904
More than median 2019 Income					
Treatment	-0.015	0.003	-0.074	-0.013	-0.032
(SE)	(0.030)	(0.034)	(0.050)	(0.035)	(0.036)
Observations	3386	3282	1378	3315	3300
Control Mean	-0.093	-0.001	-0.027	0.035	0.075
Control SD	0.916	0.974	0.923	0.991	0.977
Bonferroni	3.063	4.673	0.705	3.532	1.856
Employed					
Treatment	0.015	-0.034	-0.127	0.024	-0.075
(SE)	(0.041)	(0.044)	(0.086)	(0.044)	(0.048)
Observations	2097	2015	643	2039	2023
Control Mean	-0.143	-0.050	0.074	-0.017	0.113
Control SD	0.966	0.981	1.070	0.984	1.073
Bonferroni	3.623	2.172	0.692	2.978	0.588
Not receiving public assistance					
Treatment	-0.004	-0.016	-0.114+	-0.026	-0.046
(SE)	(0.035)	(0.037)	(0.063)	(0.037)	(0.036)
Observations	3034	2940	1073	2973	2977
Control Mean	-0.014	0.023	0.031	0.035	-0.006
Control SD	0.979	0.996	1.024	0.993	0.925
Bonferroni	4.590	3.353	0.348	2.428	0.994

Note: One-month survey respondents. Outcomes are standardized. Treatment is the $1,000 cash transfer. All analyses include the full set of controls. Standard errors are robust. Bonferroni p-values adjust for five tests in each sample split.

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix Table 5: Endogenous Stratification of Sample on Predicted Material Hardship

	Predicted Material Hardship		
	Bottom Tercile	Middle Tercile	Top Tercile
Treatment at One-Month	-0.027	-0.012	-0.081+
(SE)	(0.035)	(0.041)	(0.042)

Note: Outcomes are standardized. Treatment is the $1,000 cash transfers. Estimates use the endogenous stratification approach described in Abadie et al. (2018), including the same set of controls for the treatment effect estimates. We report the repeated split sample results from 100 bootstraps. We predict material hardship using predictors selected by LASSO prediction. Analyses include the following, all of which were measured at baseline: self-assessed expected hardship, an indicator for having received a stimulus payment, an indicator if respondents borrowed money from family or friends, and indicators if respondents report using food pantries or receive unemployment insurance.

+ $p < 0.10$, $* \ p < 0.05$, $** \ p < 0.01$, $*** \ p < 0.001$

Appendix Table 6 : Effects of Treatment on One Month Survey Individual Hardship Items, Monthly Earnings Less than $500

	Not Pay Utilities	Utilities Shutoff	Not Pay Phone/ Internet	Not Pay Rent	Double-up, Shelter or Eviction	Worried Food Run Out	Food Did Not Last	Medical Hardship	Transportation Hardship
Treatment	-0.037*	-0.048**	-0.051**	-0.066***	-0.014	-0.043*	-0.050**	-0.045**	-0.025
(SE)	(0.019)	(0.017)	(0.019)	(0.019)	(0.011)	(0.017)	(0.019)	(0.016)	(0.018)
Observations	2923	2920	2937	2856	2958	2951	2869	2968	2976
Control Mean	0.624	0.299	0.618	0.553	0.100	0.693	0.588	0.247	0.427
Control SD	0.485	0.458	0.486	0.497	0.301	0.461	0.492	0.431	0.495
Bonferroni p-value	0.408	0.039	0.059	0.006	1.00	0.101	0.062	0.033	1.00

Note: Each hardship item is a binary yes/no variable, equal to one if a respondent experienced a given hardship at the one-month survey. Treatment is the $1,000 cash transfer. All analyses include the full set of controls and standard errors are robust. Bonferroni p-values are computed for the family of nine tests.

+ $p < 0.10$, $*$ $p < 0.05$, $**$ $p < 0.01$, $***$ $p < 0.001$

Appendix Table 7: Primary Outcome Intent-to-Treat Estimates Three Months Post Cash Transfer, Monthly Earnings Less than $500

	Material Hardship	Mental Health Challenges	Partner Conflict	Child Behavior Problems	Parenting Problems
Three-Month					
Treatment	-0.111**	-0.101*	-0.093	-0.024	-0.055
(SE)	(0.043)	(0.044)	(0.086)	(0.046)	(0.045)
Observations	2253	2153	656	2166	2162
Control Mean	0.236	0.084	0.090	0.059	-0.035
Control SD	1.063	1.035	1.082	1.062	1.042
Bonferroni	0.050	0.093	0.643	0.643	0.643
FWER	0.055	0.089	0.512	0.596	0.512

Note: Three-month survey respondents. Outcomes are standardized. Treatment is the $1,000 cash transfer. All analyses include the full set of controls. Standard errors are robust. Bonferroni p-values adjust for five tests in each sample split.

+ $p < 0.10$, $*$ $p < 0.05$, $**$ $p < 0.01$, $***$ $p < 0.001$

Appendix Table 8: Intent-to-Treat Estimates on Benefit Receipt One Month Post Transfer, Full Sample and Monthly Earnings Less than $500

	Unemploy-ment Insurance	TANF	SNAP	SSDI	Medicaid, SCHIP	Other Assistance	WIC	Housing	Assistance from a charity
Full Sample									
Treatment	0.009	-0.005	0.020*	-0.006	0.012	0.010	-0.003	-0.001	0.154***
(SE)	(0.006)	(0.005)	(0.008)	(0.004)	(0.012)	(0.008)	(0.005)	(0.005)	(0.007)
Observations	6760	6769	6734	6725	6723	6739	6767	6759	6731
Control Mean	0.313	0.113	0.856	0.032	0.528	0.109	0.314	0.199	0.025
Control SD	0.464	0.316	0.351	0.176	0.499	0.311	0.464	0.399	0.157
Bonferroni	1.00	1.00	0.111	1.00	1.00	1.00	1.00	1.00	0.000
Monthly Earnings Less than $500									
Treatment	0.010	-0.014	0.032**	-0.006	0.025	0.004	-0.000	0.001	0.145***
(SE)	(0.008)	(0.008)	(0.011)	(0.006)	(0.018)	(0.012)	(0.008)	(0.007)	(0.010)
Observations	3047	3053	3047	3041	3041	3047	3053	3053	3048
Control Mean	0.288	0.146	0.872	0.026	0.505	0.108	0.321	0.244	0.024
Control SD	0.453	0.353	0.335	0.160	0.500	0.310	0.467	0.430	0.153
Bonferroni	1.00	0.72	0.052	1.00	1.00	1.00	1.00	1.00	0.000

Note: One-month survey respondents. Outcomes are binary indicators. Treatment is the $1,000 cash transfer. All analyses include the full set of controls. Standard errors are robust. Bonferroni p-values are computed over a family of 9 tests for each income subgroup.

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix Table 9: Intent-to-Treat Estimates on Labor Supply, One Month Post Transfer, Full Sample and Monthly Earnings Less than $500

	Employed Part- or Full-time	Hours Worked	Looking for Work
Full Sample			
Treatment	-0.000	-0.021	0.012
(SE)	(0.010)	(0.345)	(0.012)
Observations	6608	6608	6533
Control Mean	0.371	10.167	0.576
Control SD	0.483	16.647	0.494
Bonferroni	1.000	1.000	0.872
Monthly Earnings Less than $500			
Treatment	0.003	0.035	0.005
(SE)	(0.013)	(0.316)	(0.017)
Observations	2998	2998	2960
Control Mean	0.194	3.470	0.695
Control SD	0.396	9.567	0.461
Bonferroni	1.000	1.000	1.000

Note: One-month survey respondents. Employment/looking for work are binary indicators. Hours worked is a continuous measure. Treatment is the $1,000 cash transfer. All analyses include the full set of controls. Standard errors are robust. Bonferroni p-values adjust for a family of three tests for each subsample.

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix Table 10: Qualitative Responses to "Please tell us how you used the $1,000 from Give Directly/ Hyperwallet /Fresh EBT", One-Month Survey

How did you use the Money?	
Bills	68
Rent	27
Food	28
Home items/needs/repairs	15
Child spending	15
Car related expenses	13
Child clothing	12
Put money into savings	3
Clothes	3
Health care/health	2
Child school	2
Other (leisure, job, daily expenses)	2
Loans	1
Necessities	1
Observations	3694

Note: Responses are restricted to those who were in the treatment group and who provided a response to these questions. Responses are not mutually exclusive. N=48 reported not getting the $$.

+ $p < 0.10$, $* \ p < 0.05$, $** \ p < 0.01$, $*** \ p < 0.001$

Appendix Table 11: Intent-to-Treat Estimates of the Probability of Experiencing Hardships in Each Month, Full Sample at Three Months Post Transfer

2020	February	March	April	May	June	July	# of Months (0-6)
Paid rent/mortgage bill in full							
Treatment	-0.006	-0.004	0.005	-0.002	0.011	-0.010	-0.009
(SE)	(0.010)	(0.012)	(0.015)	(0.015)	(0.015)	(0.015)	(0.060)
Observations	4232	4247	4251	4268	4290	4296	4352
Control Mean	0.880	0.809	0.678	0.624	0.585	0.555	4.044
Control SD	0.326	0.393	0.467	0.485	0.493	0.497	2.018
Bonferroni	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Paid gas/electric bill in full							
Treatment	-0.011	-0.010	-0.009	-0.007	-0.020	-0.016	-0.065
(SE)	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)	(0.015)	(0.069)
Observations	4259	4258	4275	4294	4306	4319	4352
Control Mean	0.761	0.701	0.580	0.505	0.461	0.411	3.358
Control SD	0.427	0.458	0.494	0.500	0.499	0.492	2.253
Bonferroni	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Paid phone bill in full							
Treatment	-0.001	-0.010	0.009	-0.010	0.009	0.001	0.001
(SE)	(0.011)	(0.012)	(0.014)	(0.014)	(0.015)	(0.015)	(0.064)
Observations	4254	4260	4270	4279	4301	4307	4329
Control Mean	0.848	0.815	0.728	0.690	0.645	0.616	4.288
Control SD	0.359	0.389	0.445	0.463	0.479	0.486	2.082
Bonferroni	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Paid cable/internet in full							
Treatment	-0.015	-0.028	-0.015	-0.021	-0.020	-0.013	-0.120
(SE)	(0.014)	(0.015)	(0.016)	(0.016)	(0.016)	(0.017)	(0.072)
Observations	3590	3590	3604	3626	3655	3664	3689
Control Mean	0.802	0.763	0.676	0.624	0.579	0.533	3.906
Control SD	0.398	0.425	0.468	0.485	0.494	0.499	2.150
Bonferroni	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Unable to see a doctor							
Treatment	-0.016	-0.012	-0.015	-0.007	-0.015	-0.010	-0.075
(SE)	(0.008)	(0.008)	(0.008)	(0.009)	(0.010)	(0.010)	(0.043)
Observations	5477	5483	5480	5484	5478	5478	5502
Control Mean	0.103	0.099	0.107	0.112	0.150	0.177	0.744
Control SD	0.304	0.299	0.309	0.315	0.357	0.382	1.593
Bonferroni	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Unable to buy prescriptions							
Treatment	0.00	0.00	-0.01	-0.01	-0.01	-0.01	-0.054
(SE)	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	(0.045)
Observations	5309	5313	5313	5325	5335	5343	5366
Control Mean	0.09	0.09	0.10	0.12	0.14	0.16	0.706
Control SD	0.29	0.29	0.31	0.32	0.34	0.37	1.640
Bonferroni	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Note: Sample comes from the three-month survey. Monthly outcomes are binary indicators, number of months ranges from 0-6. Treatment is the $1,000 cash transfer which occurred in May. All analyses include the full set of controls. Standard errors are robust. Bonferroni p-values are computed over a family of 42 tests.

$+\ p < 0.10$, $*\ p < 0.05$, $**\ p < 0.01$, $***\ p < 0.001$

Appendix Table 12: Intent-to-Treat Estimates of the Probability of Experiencing Hardships in Each Month, Monthly Earnings of Less than $500, Three Months Post Transfer

2020	February	March	April	May	June	July	# of Months (0-6)
Paid rent/mortgage bill in full							
Treatment	0.013	0.030	0.074**	0.064*	0.062*	0.039	0.260*
(SE)	(0.019)	(0.022)	(0.025)	(0.025)	(0.025)	(0.025)	(0.102)
Observations	1589	1590	1595	1606	1613	1613	1643
Control Mean	0.840	0.746	0.554	0.482	0.457	0.430	3.419
Control SD	0.367	0.436	0.497	0.500	0.499	0.496	2.057
Bonferroni	1.000	1.000	0.146	0.493	0.554	1.000	0.473
Paid gas/electric bill in full							
Treatment	-0.011	-0.005	0.026	0.032	-0.003	0.011	0.059
(SE)	(0.022)	(0.024)	(0.025)	(0.025)	(0.024)	(0.024)	(0.112)
Observations	1641	1640	1651	1660	1664	1666	1676
Control Mean	0.747	0.670	0.522	0.432	0.400	0.347	3.064
Control SD	0.435	0.471	0.500	0.496	0.490	0.476	2.206
Bonferroni	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Paid phone bill in full							
Treatment	0.002	-0.012	0.031	0.029	0.049	0.057*	0.147
(SE)	(0.020)	(0.021)	(0.024)	(0.025)	(0.025)	(0.025)	(0.107)
Observations	1653	1658	1656	1659	1670	1671	1684
Control Mean	0.810	0.774	0.661	0.600	0.549	0.517	3.858
Control SD	0.393	0.419	0.474	0.490	0.498	0.500	2.116
Bonferroni	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Paid cable/internet in full							
Treatment	-0.012	-0.013	0.005	0.008	0.009	0.008	0.019
(SE)	(0.024)	(0.026)	(0.028)	(0.028)	(0.028)	(0.027)	(0.120)
Observations	1348	1349	1356	1366	1378	1381	1387
Control Mean	0.772	0.719	0.607	0.542	0.490	0.460	3.507
Control SD	0.420	0.450	0.489	0.499	0.500	0.499	2.181
Bonferroni	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Unable to see a doctor							
Treatment	-0.021	-0.024	-0.039**	-0.024	-0.039*	-0.043*	-0.190**
(SE)	(0.013)	(0.014)	(0.014)	(0.014)	(0.016)	(0.017)	(0.070)
Observations	2226	2227	2227	2232	2227	2225	2238
Control Mean	0.118	0.119	0.133	0.135	0.184	0.207	0.892
Control SD	0.323	0.324	0.340	0.342	0.388	0.405	1.714
Bonferroni	1.000	1.000	0.234	1.000	0.683	0.465	0.298
Unable to buy prescriptions							
Treatment	-0.004	-0.021	-0.029*	-0.024	-0.015	-0.034*	-0.126
(SE)	(0.014)	(0.014)	(0.015)	(0.015)	(0.016)	(0.017)	(0.076)
Observations	2131	2135	2134	2144	2152	2150	2161
Control Mean	0.111	0.119	0.136	0.147	0.161	0.198	0.864
Control SD	0.315	0.324	0.343	0.355	0.367	0.399	1.803
Bonferroni	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Note: Sample comes from the three-month survey. Monthly outcomes are binary indicators, number of months ranges from 0-6. Treatment is the $1,000 cash transfer which occurred in May. All analyses include the full set of controls. Standard errors are robust. Bonferroni p-values are computed over a family of 42 tests.

+ $p < 0.10$, $* \ p < 0.05$, $** \ p < 0.01$, $*** \ p < 0.001$

Appendix Table 13: Intent-to-Treat Estimates on Amounts and Types of Back Owed Debts Three Months Post Transfer, Full Sample and Monthly Earnings Less than $500

	Rent/ Mortgage	Gas/ Electric	Phone	Cable/ Internet	Medical	Family/ Friends	Other Debt	Car	Total # of Categories
Full Sample									
Owe Back Debt									
Treatment	0.004	0.022	0.002	0.031	-0.000	0.039**	-0.020	-0.016	0.137**
(SE)	(0.015)	(0.015)	(0.015)	(0.017)	(0.013)	(0.014)	(0.018)	(0.021)	(0.049)
Observations	4202	4259	4231	3599	5508	5521	2819	2189	5774
R2	0.098	0.074	0.060	0.086	0.070	0.058	0.092	0.079	0.097
Control Mean	0.468	0.592	0.360	0.466	0.357	0.515	0.695	0.366	2.597
Control SD	0.499	0.492	0.480	0.499	0.479	0.500	0.461	0.482	1.865
Bonferroni	1.00	1.00	1.00	0.551	1.00	0.039	1.00	1.00	0.044
Amount in $ Back Debt Owed									
Treatment	-120.198*	-36.368	-17.042	2.556	346.949	-52.034	-349.164	-0.051	-14.878
(SE)	(61.165)	(24.383)	(12.306)	(8.524)	(390.590)	(47.367)	(770.632)	(79.223)	(323.153)
Observations	1950	2570	1525	1741	1929	2921	1868	772	4952
R2	0.095	0.090	0.072	0.070	0.055	0.091	0.093	0.109	0.061
Control Mean	1613.944	606.688	269.463	214.574	4134.746	1027.654	9769.802	901.843	6671.227
Control SD	1339.810	624.285	244.237	177.349	7905.868	1247.482	1.6e+04	1083.035	1.1e+04
Bonferroni	0.446	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Monthly Earnings Less than $500									
Owe Back Debt									
Treatment	-0.048	-0.007	-0.032	0.016	-0.014	0.041	-0.038	-0.001	0.088
(SE)	(0.025)	(0.024)	(0.025)	(0.028)	(0.020)	(0.021)	(0.029)	(0.037)	(0.076)
Observations	1575	1642	1640	1346	2248	2259	986	755	2349
R2	0.125	0.080	0.062	0.101	0.089	0.053	0.108	0.125	0.113
Control Mean	0.594	0.654	0.447	0.544	0.332	0.580	0.770	0.389	2.714
Control SD	0.491	0.476	0.498	0.499	0.471	0.494	0.422	0.488	1.898
Bonferroni	0.506	1.00	1.00	1.00	1.00	0.492	1.00	1.00	1.00
Amount in $ Back Debt Owed									
Treatment	-245.376**	-30.531	-28.370	-13.930	1370.934*	-110.911	413.032	-92.862	414.406
(SE)	(93.347)	(42.144)	(19.167)	(14.129)	(679.083)	(72.339)	(1263.338)	(135.521)	(498.885)
Observations	869	1058	688	732	707	1326	713	284	2048
R2	0.149	0.104	0.127	0.148	0.118	0.127	0.153	0.252	0.103
Control Mean	1742.711	632.441	269.672	238.425	3473.419	1043.126	9136.568	1012.598	6173.622
Control SD	1413.878	672.565	267.713	199.682	7097.339	1279.521	1.5e+04	1311.963	1.0e+04
Bonferroni	0.079	1.00	1.00	1.00	0.395	1.00	1.00	1.00	1.00

Note: Sample restricted to three-month survey. Back owed measures are all binary outcomes. Amount of back owed debt is conditional on reporting debt in a given category. Standard errors are robust. Bonferroni p-values are computed over families of 9 tests.

+ $p < 0.10$, $* \; p < 0.05$, $** \; p < 0.01$, $*** \; p < 0.001$

Appendix A: Study Design

Study enrollment and experimental design

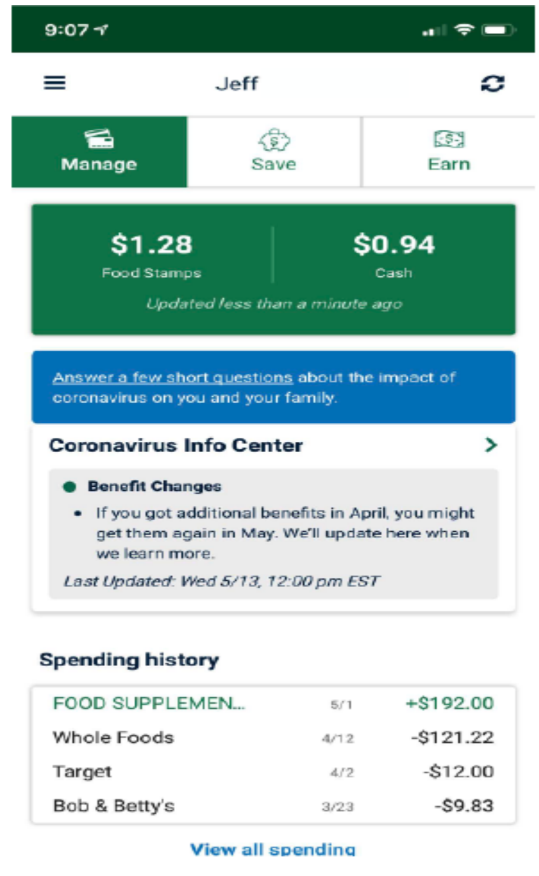
This appendix provides additional information on the study design. As described in the text, GiveDirectly (GD) launched Project 100+ in March 2020 with the goal of providing direct cash transfers to more than 100,000 low-income families struggling to make ends meet during the COVID-19 pandemic. In the first phase of the program, individuals were eligible to receive a $1,000 unconditional cash transfer if they were receiving federal food assistance through the Supplemental Nutrition Assistance Program (SNAP). GD identified 15 states (LA, NH, ME, RI, MN, NC, OH, MI, PA, IN, NM, KY, DE, NE, and MA) that had been hit hard by the coronavirus and began providing transfers in those states to individuals who were not receiving Supplemental Security Income or Social Security Disability Insurance benefits (SSI or SSDI).[1] As a one-time cash transfer, the transfer did not affect recipients' eligibility for public assistance. In later waves of Project 100+ GD expanded to include other states and later also included SSI and SSDI recipients. We partnered with GD to evaluate their program and launched our study with a wave of cash transfers that went out on May 21, 2020.

GD partnered with Propel, the makers of the free Fresh EBT mobile application (app) that enables individuals receiving public assistance to manage and review their benefits. Fresh EBT (now called Providers) had more than 4 million users at the time of this study. To disburse the Project 100+ cash transfers, Fresh EBT presented a banner (a color differentiated section on the app; see Figure A1 below) on the app to select users, inviting them to take a short eligibility survey. Users who clicked on this banner, answered the screener questions, and were deemed eligible based on their responses were informed that they were selected to receive a cash payment from GD. To collect the award, individuals were required to acknowledge they were over 18 years of age, provide their first and last names, email address, and phone number, and sign a waiver of claims against GD. Recipients received an email from GD within three days with instructions for accessing the funds. To access funds individuals had to complete the enrollment form and agree to GD's terms. The money was then transferred into a Hyperwallet account. Recipients used the Hyperwallet platform to choose how to receive their payments; options included PayPal, cash pickup, virtual or physical prepaid cards, ACH bank account transfer, or paper check.

---

[1] GD relied on a report projecting high unemployment in the first 15 target states: https://tcf.org/content/commentary/new-data-show-true-march-jobless-rate-near-20-percent/

Figure A1: Example In-App Banner for Study Enrollment



We partnered with GD to study the effect of the cash transfers on recipients. To do so, we employed the procedure GD used to distribute all Project 100+ transfers, with modifications to accommodate a control group. As part of its internal operations, Fresh EBT randomly assigns individual users to 1 of 1,000 different groups called segments, which they then use for various diagnostics and user outreach purposes (our own analyses also confirm the distribution of users across segments is consistent with random assignment). For our study, users in segments 0-300 and 801-1000 were selected for the treatment group and users in segments 500-800 were assigned to a control group (see Table A1).

Eligibility for the cash transfer and study was limited to individuals who had accessed the app in the past 30 days (to ensure they were active users and likely still receiving benefits) and to those who had not previously received a cash payment from GD. We excluded individuals living in three states (DE, NE, MA) where usage of Fresh EBT was low, leading to concerns that the app users would not be representative of the SNAP population in those states. We further limited the study to individuals with children because financial hardships are particularly acute in this population (Rodems & Shaefer, 2020), and because the majority of Fresh EBT users include families with children (about 80% according to Fresh EBT). Although roughly two-thirds of SNAP benefits go to families with children (CBPP, 2017), only 40% of households receiving SNAP in 2018 had children (USDA, 2019). Thus, our sample is not representative of all SNAP households, but rather focused on families with children receiving SNAP in the 12 sample states.

Table A1: Study Segment Assignment and Times of Day

Segments	Randomized Group	Dates used in prior campaigns	Time exposed on 5/21/20	Time closed on 5/21/20
0-100	T	3/30/20 and 4/27/20	7:00 AM	10:50 AM
101-200	T	3/30/20 and 4/27/20	7:00 AM	10:50 AM
201-300	T	4/6/20 and 5/4/20	9:00 AM	10:50 AM
500-600	C	4/20/20 and 5/4/20	7:00 AM	10:05 AM
601-700	C		7:00 AM	10:05 AM
701-800	C		9:00 AM	10:05 AM
801-900	T		11:20 AM	12:06 PM
901-1000	T		11:20 AM	12:06 PM

Note: Fresh EBT assigns users to segments 1-1000 used to display content to subsets of the application population. At 7:00 am on 5/21/20, Fresh EBT administrators exposed segments 0-200 to our study banner to be enrolled in our treatment group, and segments 500-700 to be enrolled as our control group. At 9:00 am, after realizing that segments 0-200 and 500-600 had already been exposed to prior Project 100 banners, administrators exposed additional segments to our study banner. At 11:20 am, given the limited exposure of application users in our treatment group to the study banner, administrators exposed an additional set of segments to our study banner.

Study enrollment began the morning of May 21, 2020. Starting at 7am EST, individuals in both treatment and control segments were shown an in-app banner inviting them to take a survey about the impact of the coronavirus on their family. The banner did not explicitly mention the potential $1,000 cash transfer, but many app users were aware of the GD program by this time and so it is likely that at least some of them clicked on the banner in anticipation of the award. Within roughly five hours, a sufficient number of treatment individuals had claimed the cash award that the available funding for that day's campaign had been expended. At this point, Fresh EBT removed the banner from the app for both treatment and control users. Fresh EBT did not contact users in any way, so only users who logged onto the app during this five-hour window could have seen the banner and been included in our study.

Figure A2: Study Enrollment Process

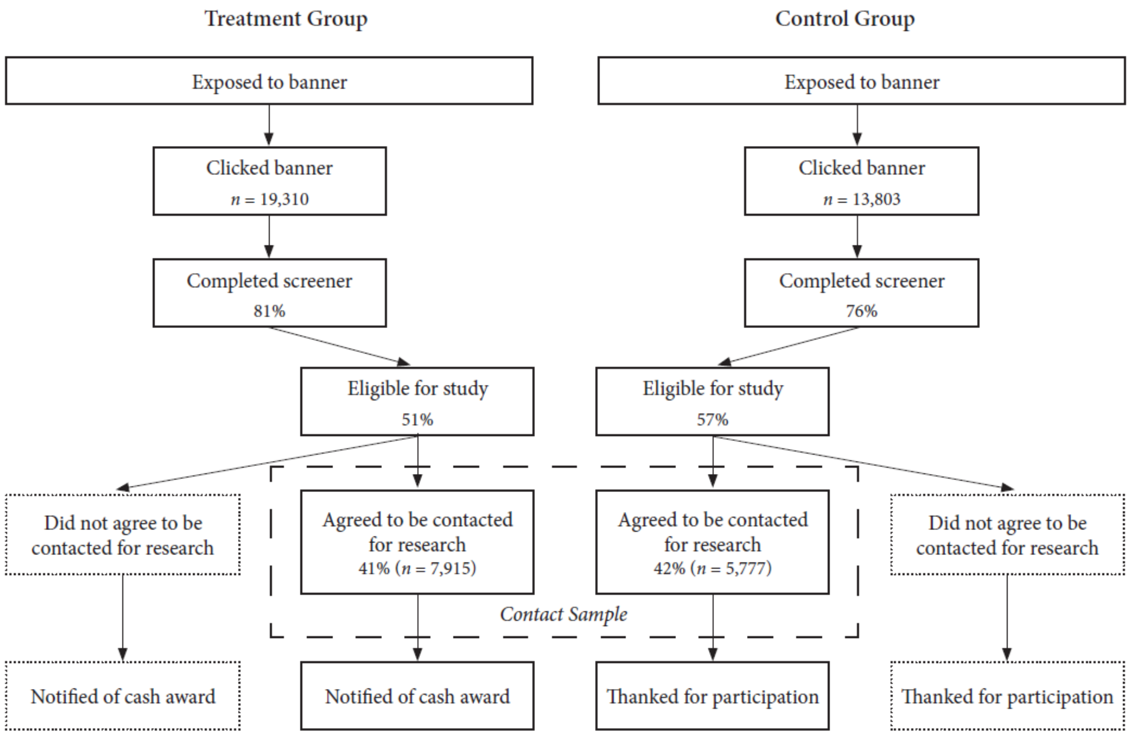


Figure A2 illustrates the flow of individuals through the study enrollment process. The treatment and control banners were identical (the number of clicks in the treatment group is driven by the different number of segments exposed, see Table A1). Individuals who clicked on either treatment or control banners were asked seven identical screener questions, including one that allowed us to identify families with children. Neither the banner nor the screener mentioned GD or the cash transfer. A little over half of the respondents (51% treatment, 57% control) were eligible and, of those who were eligible, roughly 70 and 69 percent (treatment and control group, respectively) agreed to share contact information with researchers to participate in a research study about how the coronavirus pandemic is affecting families.

Neither eligibility for the study nor willingness to share contact information with researchers affected receipt of the cash transfer. After indicating whether they were or were not interested in participating in the research study, individuals in the control group were simply thanked for answering the short set of screener questions and individuals in the treatment group were given the opportunity to receive a $1,000 cash transfer from GD. As soon as individuals completed the enrollment form and agreed to GD's terms, the money was transferred in as little as 24 hours to a Hyperwallet account. Recipients used the Hyperwallet platform to choose how to receive their payments; options included PayPal, cash pickup, virtual or physical prepaid cards, ACH bank account transfer, or paper check. As described in more detail below, we later reached out to the 13,692 individuals (7,915 treatment, 5,777 control) who agreed to share contact information with researchers, asking them to complete the baseline survey.

We pre-registered this study with five primary outcomes at the AEA RCT Registry (blinded). As described in the pre-registration document, we powered the study to detect effect sizes of 0.09 standard deviation on our material hardship, mental health, child well-being and parenting behavior outcomes, and to detect sizes of 0.11 standard deviations for the partner relationship outcome. Our target enrollment was 10,000 individuals, assuming a 90 percent take-up rate and a 60 percent response rate to our one-month follow-up survey. We also described planned secondary analyses to look at treatment effect heterogeneity by baseline hardship (among other things), although we did not power the study to detect specific subgroup effects. Given that the entire sample was to be composed of low-income households (by construction), we expected to find effects in the main sample.

Data and Measures

We use three sources of data for our study: 1) administrative data from Fresh EBT; 2) screener data from the brief in-app survey; and 3) data from three rounds of online survey data collection in this study.

From the Fresh EBT app, we obtained administrative data on the individual's state, zip code, year of birth, preferred language (e.g., English or Spanish), number of days since first use of the Fresh EBT app, the number of unique app uses, and the amount of their most recent SNAP benefit (the balance of SNAP benefit on their EBT cards). The data that Fresh EBT collects is quite limited, thus we did not have information on income or family size. We also created a treatment indicator for user segments 0-300 and 801-999 (those in the treatment group). We received administrative data from GD specifying when the money was transferred into the recipient's Hyperwallet account in order to determine if the treatment participants had cashed out/received their money.

From the screener questions administered in Fresh EBT, we obtained self-reported information on whether the respondent experienced a loss of income due to the coronavirus, had increased expenses due to coronavirus, and had received a CARES Act stimulus payment (although the first checks went out in April, many people had not yet received them by late May). We created binary indicators for each of these variables. The two final screener questions asked users how worried they were about being able to afford necessities and how difficult the pandemic had been for their children. Responses ranged from 1 to 5, where larger numbers indicate more worry or difficulty.

We administered three online surveys via Qualtrics: a baseline survey shortly after the respondents consented to participate in the study but before treatment, and follow-up surveys one and three months after receipt of the cash transfer. Surveys were offered in Spanish and English. We invited individuals to participate using the email address and cell phone number they provided in Fresh EBT. Respondents were sent a $10 gift card via Tango Rewards for each completed survey as compensation. Reminder emails and text messages were sent regularly to help ensure high response rates, and we increased the incentive amount to $20 during the final days of the survey. We also employed some phone calling to increase response rates at the three-month survey (note, results are not sensitive to restricting to those who did not get calls or receive additional incentive). We ran two pilot surveys (n=100 each) to test the flow between Fresh EBT, the survey, and incentive distributions before study enrollment.

The surveys included questions on respondent demographics, household composition, employment, and benefit receipt, along with instruments measuring the main outcomes described below (survey instruments are available upon request). The questionnaire varied slightly for each survey. The baseline survey was very concise and focused on key demographics and outcome measures, whereas the one-month survey included a few additional measures (like consumption and a qualitative question on how the treatment group spent the cash transfer). The three-month survey included a few additional measures aimed at capturing the intensity of material hardship and hardship avoidance techniques (pilot tested for clarity of measures with about 100 respondents, including 5 in-depth qualitative interviews).

Figure A3: Survey Response Rates by Wave

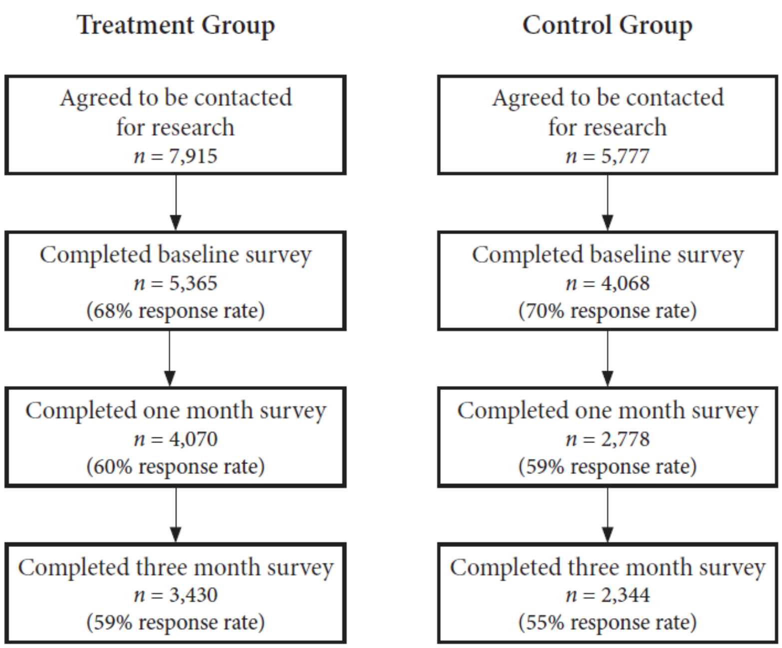


Figure A3 shows the survey response rates by each survey wave. Invitations to complete the baseline survey were sent on May 21st, 2020, shortly after the respondents agreed to be contacted by researchers but before most treatment participants received the transfer (3% received it the first day, 90% cashed out the 2nd day and 99.4% had cashed out before the one-month survey). The baseline survey was open for 8 days, during which 9,433 individuals (68% of those contacted for the study) completed the survey.

The one-month follow-up survey was sent on June 18th, 2020 to all of the respondents who had taken the baseline survey and was open for 9 days. The one-month follow-up survey was completed by 6,848 individuals, for a response rate of 60% and 59% of the invited treatment and control samples respectively. The three-month follow-up survey was launched on August 21st, exactly 3 months after the notification of the cash transfer (only those who responded to either the baseline or one-month survey were invited to take the three-month survey). The three-month survey was open for 14 days. The overall response rate was 57% or N=5,774 (3,430 [59%] treatment, 2,344 [55%] control).

A number of individuals agreed to be contacted for the study after completing the screener questions but then did not take the baseline survey once they were contacted.[2] We have app data and the preliminary screener question information on these individuals. In Table A2, we compared their data to those who took the baseline survey. In general, the characteristics are similar, but there are some significant differences. Those who did not take the survey used the application less often and reported slightly fewer problems in the screener questions (fewer hardships, less child difficulty, lower difficulty with increased expenses).

Table A2: Respondents who provided contact information: Differences between those who took the baseline survey and those who did not

	Did not take baseline survey	Took baseline survey	T-Statistic
Application Data			
# of days using Fresh EBT	312.67	312.87	0.03
Joined before Feb 1, 2020	0.53	0.51	-2.44
Joined on May 21, 2020	0.04	0.03	-3.18
# of App Uses	162.92	188.05	5.75
SNAP benefit amount	332.93	353.27	3.89
Missing_Benefits	0.28	0.32	5.41
Spanish language	0.02	0.01	-1.88
State			
Indiana	0.03	0.03	1.47
Kentucky	0.03	0.03	0.53
Louisiana	0.06	0.04	-2.90
Maine	0.00	0.00	-0.97
Michigan	0.24	0.24	-0.39
Minnesota	0.01	0.01	0.51
North Carolina	0.17	0.17	-0.48
New Mexico	0.01	0.00	-2.52
Ohio	0.26	0.30	4.73
Pennsylvania	0.19	0.16	-3.50
Rhode Island	0.01	0.00	-1.23
Application Screener			
Lost income due to COVID	0.95	0.96	1.88
Expenses increased due to COVID	0.89	0.91	3.48
Received stimulus (CARES) payment	0.70	0.69	-1.02
Worried about affording basic necessities/bills (5-point scale)	4.53	4.57	3.11
How difficult has COVID been for children (5 point scale)	4.40	4.46	4.50
N	4253	9439	

[2] Note, these individuals who did not complete include those whose contact information was bad or incorrect as well as those who simply opted not to participate.

To address potential concerns of differential attrition, we test if attrition from each survey sample is statistically associated with treatment status. We find that control users were 3.9 percentage points more likely to leave our sample between the baseline and one-month survey; however, we find no significant association between treatment assignment and attrition between our one- and three-month surveys (see Table A3). Moreover, when we examine treatment and control respondents who attritted, we see they are balanced on observable baseline characteristics (see Table A4), suggesting similar respondents attritted across both groups.

Table A3: Association between Attrition and Treatment Status

	Attritted between Baseline Survey and One-Month Follow-up	Attritted between One-Month Survey and Three-Month Follow-up	Attritted between One-Month Survey and Three-Month Follow-up, Low Income Sample
Treatment	-0.039***	-0.012	-0.023
(SE)	(0.009)	(0.010)	(0.015)
Observations	9433	6848	3084
R2	0.121	0.076	0.075

Note: We regress an indicator equal to one if a respondent left our sample (after completing the baseline or one month survey respectively) onto an indicator for treatment assignment.
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A4: Covariate Balance of Treatment and Control Groups that Responded to Baseline but No Follow-Up Surveys – Attriters

	Control Mean	Treatment Mean	T-Statistic	Standardized Difference
Race/Ethnicity				
Hispanic	0.048	0.052	0.456	0.019
Non-Hispanic White	0.142	0.142	0.008	0.000
Non-Hispanic Black	0.711	0.696	-0.841	-0.034
Other Race/Ethnicity	0.029	0.032	0.470	0.019
Female	0.917	0.908	-0.840	-0.034
Education				
Less than high school	0.130	0.137	0.501	0.020
Some college	0.252	0.270	1.000	0.040
Associates or more	0.098	0.080	-1.562	-0.060
Age	32.238	31.721	-1.643	-0.068
Employment				
Full-time	0.169	0.159	-0.670	-0.027
Part-time	0.119	0.101	-1.382	-0.054
Temporarily laid off	0.345	0.358	0.683	0.027
Unemployed	0.207	0.200	-0.433	-0.017
Other employment status	0.093	0.089	-0.325	-0.013
Public Assistance				
WIC	1.000	1.000		
Housing	1.000	1.000		
Unemployment Insurance	1.000	1.000		
TANF	1.000	1.000		
Household composition				
Household size	3.923	3.978	0.822	0.034
Married	0.121	0.104	-1.328	-0.052
Cohabiting	0.094	0.121	2.180	0.092
Lives with own children	0.988	0.995	1.789	0.062
Lives with other children	0.092	0.089	-0.252	-0.010
Lives with other relatives	0.086	0.093	0.597	0.024
Lives with non-relatives	0.011	0.008	-0.929	-0.034
Number of children				
Under age 6	1.055	1.080	0.597	0.025
Ages 6-12	1.041	1.068	0.534	0.022
Ages 13-17	0.516	0.444	-1.986	-0.077
Total number of children	2.584	2.571	-0.216	-0.009
N	1170	1331		

Note: Sample of all respondents who completed the baseline survey but did not complete the one month survey. T-statistics come from bi-variate regressions of each variable on treatment status, with robust standard errors. The standardized difference divides the t-statistic by the standard deviation of the control group.

Although all users had to indicate that they had children in order to be sent our initial survey, not all users appear to live with their children or other people's children. As such, we drop 176 users from our study sample who do not appear to reside with any children. Additionally, if a user indicates that they live with children and a partner, we verify that their household size accounts for themselves in addition to other members.

Outcome Measures

Material Hardship. Material hardship is a consumption-based indicator of economic wellbeing that has been linked with poorer outcomes for both children (e.g., Zilanawala & Pilkauskas, 2012) and adults (Heflin & Iceland, 20019). Following many earlier studies (e.g., Shaefer & Guttierez, 2013; Pilkauskas, Currie & Garfinkel, 2012) our 9-item measure of material hardship assesses housing hardships (inability to pay rent, eviction, doubling up), food insecurity, inability to pay utility bills, and inability to access needed medical care. We also include a measure of transportation insecurity (Gould-Werth, Griffin & Murphy, 2018). Questions were drawn from the Survey of Income and Program Participation, the Transportation Security Index and the Fragile Families and Child Wellbeing Study (FFCWS). In addition to the composite measure, we examine each individual item.

Mental Health Challenges. Poverty and income are linked with mental health challenges (e.g., Heflin & Iceland, 2009; Haushofer & Shapiro, 2016), which in turn is linked with longer-term wellbeing (e.g., Mitra & Jones, 2017). We include two well-validated measures of mental health challenges – the short form of the Patient Health Questionnaire (PHQ-2) for depression and the abbreviated Generalized Anxiety Disorder (GAD-2) to assess anxiety. Both scales are frequently used in other studies and have been shown to be reliable measures of their respective outcomes (e.g., Levis et al., 2020). In addition to the composite index we examined anxiety and depression scales separately.

Partner Conflict. Many studies link income with relationship quality and relationship dissolution (e.g., Papp, Cummings & Goeke-Morey, 2009). We assess partner relationships using questions from the FFCWS that were based on a modified version of the Conflict Tactics Scale. In particular, we include items that assess physical aggression, sexual coercion, psychological aggression and controlling behaviors. We also include indicators of positive relationship quality and emotional support. The relationship quality questions were asked of respondents who were married or co-residing with a partner.

Child Behavior Problems. Child behavior is linked with both short- and longer-term outcomes (e.g., Duncan & Brooks-Gunn, 1997) and the link between child behavior and income is well established (e.g., Duncan, Magnuson & Votruba-Drzal, 2017). Measuring child behavior in a short online survey is challenging as behavior varies by child’s age and by individual child. We did not have the time to ask parents to report on each child; thus, we selected questions used in the FFCWS that were adapted from Achenbach’s Child Behavior Checklist that would be sufficiently broad to cover a wide range of child ages. We included items that pick up on both externalizing (destructive/aggressive behaviors; 4 items) and internalizing (withdrawn/sad behaviors; 4 items) behavior scales as well as prosocial behaviors (2 items). We also included one item that assesses sleep behaviors.

Parenting Problems. Parent-child relationships are key factors in child development, and research finds that income affects parents' stress, which in turn affects parenting behaviors (e.g., Nomaguchi & House, 2013). To assess parenting, we use measures adapted from the FFCWS that pick up on both harsh parenting behaviors (yelling/hitting or spanking; 2 items) and positive parenting behaviors (helping children/spending time with children; 2 items). As is the case with child behavior, measures of parenting typically differ by the age of the child; thus, we adapted the questions to keep them sufficiently general to be applicable to children of all ages.

Composites

For all outcomes we create composite indices, following the approach outlined in Anderson (2008) to calculate a weighted mean of the standardized survey items within the domain. We first switch the signs of all items so that the positive direction always indicates a "worse" outcome (i.e., more material hardship, more parenting problems). We then standardize each outcome using the mean and standard deviation of that outcome among the control group and create a weighted average of these standardized items using all of the outcomes in the domain. To create weights, we use the inverse of the covariance matrix of the transformed outcomes in the domain. If an individual does not respond to a particular item, we assign the item the mean (standardized) value of the individual's responses to the other items in the same domain. We require non-missing responses for at least two thirds of items within a given domain for a respondent to be included in that domain's composite index. Analysis of missing response patterns indicated that most users answered all items within each domain. The proportion of respondents missing data for individual items ranged from less than 1 to 3.3 percent. Finally, we standardize the weighted averages using the mean and standard deviation of the composite among the control group, so that each composite can be interpreted as a z-score.

Control Variables

For the survey data, respondents who left an item blank or responded "Prefer not to answer" were assigned a missing value for the indicator. We construct mutually exclusive indicators for users that are non-Hispanic Black, non-Hispanic White, Hispanic, and "other race" which encompasses the remaining respondents. A respondent's race or ethnicity is only coded as missing if they did not answer both race and ethnicity items. We then constructed a variable that indicates whether the respondent identified as female. To code levels of education, we created indicators for those who specify that they have not graduated from high school, graduated from high school or equivalent, attended some college without completing a degree, and received an associate's degree or more schooling. We also include two measure of household structure: 1) a household size variable (number of people living in the household); and 2) indicator variables for whether the respondent lives with each of the following: their spouse, their non-spouse partner, their own children, other's children, their relatives, or other unrelated adults.

To control for employment, we constructed a series of mutually exclusive indicators for full-time employment, part-time employment, unemployment, other employment statuses, being temporarily laid off, and missing employment status. Because respondents could select multiple employment statuses, we enforce the following restrictions in the order described: if a user indicates any full-time employment, we set all other statuses to zero; if a user reports part-time, other employment, or unemployment, we set temporarily laid off to zero; if a user reports part-time employment, temporarily laid off, or unemployment, we set other employment to zero; if a user indicates any unemployment, all other statuses are set to zero.

References

Anderson, M. L. (2008). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, *103* (484), 1481–1495.

Center on Budget and Policy Priorities. (2017). *SNAP Helps Millions of Children*. https://www.cbpp.org/research/food-assistance/snap-helps-millions-of-children.

Duncan, G., & Brooks-Gunn, J., Eds. (1997). *Consequences of Growing Up Poor*. New York: Russell Sage.

Duncan, G. J., Magnuson, K., & Votruba-Drzal, E. (2017). Moving beyond correlations in assessing the consequences of poverty. *Annual Review of Psychology*, 68, 413-434

Gould-Werth, A., Griffin, J., & Murphy, A. K. (2018). Developing a New Measure of Transportation Insecurity: An Exploratory Factor Analysis. *Survey Practice*, *11*(2), 1–28.

Haushofer, J., & Shapiro, J. (2016). The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya. *Quarterly Journal of Economics*, 131(4), 1973–2042.

Heflin, C. M., & Iceland, J. (2009). Poverty, material hardship, and depression. *Social Science Quarterly*, 90(5), 1051-1071.

Levis, B., Sun, Y., He, C., Wu, Y., Krishnan, A., Bhandari, P. M., et al. (2020). Accuracy of the PHQ-2 Alone and in Combination With the PHQ-9 for Screening to Detect Major Depression. *Journal of the American Medical Association*, *323*(22), 2290–11.

Mitra, S., & Jones, K. (2016). The impact of recent mental health changes on employment: new evidence from longitudinal data. *Applied Economics*, 49(1), 96–109.

Nomaguchi, K., & House, A. N. (2013). Racial-ethnic disparities in maternal parenting stress: The roles of structural disadvantages and parenting values. *Journal of Health and Social Behavior*, 54(3), 386-404.

Papp, L., Cummings, E., & Goeke-Morey, M. (2009). For richer, for poorer: Money as a topic of marital conflict in the home. *Family Relations*, 58(1), 91–103.

Pilkauskas, N. V., Currie, J. M., & Garfinkel, I. (2012). The great recession, public transfers, and material hardship. *Social Service Review*, 86(3), 401-427.

Rodems, R., & Shaefer, H. L. (2020). Many of the kids are not alright: Material hardship among children in the United States. *Children and Youth Services Review*, 112(2020), 104767.

Shaefer, H. L., & Gutierrez, I. A. (2013). The Supplemental Nutrition Assistance Program and material hardships among low-income households with children. *Social Service Review*, 87(4), 753-779.

U.S. Department of Agriculture, Food and Nutrition Service, Office of Policy Support (2019). *Characteristics of Supplemental Nutrition Assistance Program Households: Fiscal Year 2018*. Alexandria, VA: U.S. Department of Agriculture.

Zilanawala, A., & Pilkauskas, N. V. (2012). Material hardship and child socioemotional behaviors: Differences by types of hardship, timing, and duration. *Children and Youth Services Review*, 34(4), 814-825.

Appendix B: Implementation Issues and Balance Tests

In this appendix, we discuss study implementation problems and present results from several statistical tests that shed light on the nature and extent of potential imbalance. We conclude that implementation problems likely caused modest imbalance along several dimensions. However, the magnitude of the imbalance in the sample of survey respondents is small, and we identify a subset of the study sample that was not subject to any implementation problems and appears well balanced. Analyses on this balanced subsample are consistent with those reported in the main paper.

Issues with Implementation

On Thursday, May 21, enrollment via the Fresh EBT mobile app began at 7am EDT. Shortly thereafter Fresh EBT administrators realized that all the segments they assigned to the treatment banner had the opportunity to obtain cash awards in earlier enrollment periods in March and April. In contrast, only one-third of the control segments had previously been offered these opportunities. Individuals who previously received cash payments are a non-random sample of users. Prior award recipients tend to be younger and more active app users than non-recipients. Because we excluded users who had previously received a cash transfer from our study, the imbalanced exposure to previous campaigns between treatment and control may have inadvertently excluded a larger, selected sample of users from our treatment group. As a result, it is possible that treatment users exhibit different average characteristics than control users.

To re-balance the sample later in the morning, Fresh EBT administrators showed the treatment banner to users in segments 801-1000, all of whom had no prior opportunities to obtain a cash award from GiveDirectly. While the effort to balance the sample in terms of prior campaign exposure may have been helpful, these segments were exposed to the banner at a different time of day than the control segments which may have introduced additional sample imbalance. Appendix Table 1 summarizes which segments were shown treatment and control banners during which times of the day.

Finally, Fresh EBT inadvertently neglected to forward to researchers contact information on 345 individuals who agreed to share contact information with researchers but did not complete the process necessary to obtain the cash payment. This includes individuals who did not want to agree to the various terms required by GD, as well as those who may have become suspicious of the offer or confused by the several pages of legal acknowledgments and agreements. As a result, these individuals were not included in the initial survey, but all 345 were invited to take the one-month survey.

Identifying Imbalance

To explore the effects of these implementation issues, we examine balance on two sets of variables. The first set includes information on individuals from the Fresh EBT database. These "app user characteristics" include age, state of residence, preferred language spoken, most recent SNAP benefit amount, the date the individual first signed up for the app, and the number of times the individual used the app prior to May 21, 2020. The second set includes responses to the screener questions.

Users in the treatment and control segments selected into our sample by clicking on the banner, answering the screener questions and agreeing to participate in the study. We do not expect that these individuals will be completely representative of all app users who were shown the banners. For example, more active users will be more likely to see the banner during the enrollment period, and certain individuals will be more likely to click the banner than others for various reasons. Our concern here is whether the treatment and control groups appear balanced in this sample.

To verify balance, we regress different characteristics on an indicator for treatment assignment, and then run a joint test that all coefficients are jointly zero. We report the F-statistic and p-value from this seemingly-unrelated estimation (SUE) framework that reflects the correlation across covariates. We also conduct permutation-based tests by randomly reassigning treatment status and conducting the joint hypothesis test 1000 times. We then compared the F-statistic obtained from the actual sample to the distribution of F-statistics from these permutations. The permutation-based p-value equals the fraction of simulated F-statistics that were greater than the observed F-statistic. Each balance table follows an identical format to those included in the main text and Appendix Tables.

Table B1: Balance Table of All Users that Consented to Research and Provided Contact Information to Propel

	Control Mean	Treatment Mean	T-Statistic	Standardized Difference
Application Data				
# of days using Fresh EBT	330.501	299.886	-4.752	-0.081
Joined before Feb 1, 2020	0.535	0.497	-4.459	-0.077
Joined on May 21, 2020	0.025	0.041	5.324	0.103
# of App Uses	196.593	168.317	-6.793	-0.113
SNAP benefit amount	344.723	348.172	0.698	0.014
Missing Benefits	0.284	0.327	5.438	0.096
Spanish Language	0.016	0.013	-1.288	-0.021
State				
Indiana	0.030	0.033	0.983	0.017
Kentucky	0.023	0.034	4.084	0.078
Louisiana	0.046	0.050	0.946	0.017
Maine	0.002	0.001	-1.341	-0.021
Michigan	0.260	0.230	-4.027	-0.069
Minnesota	0.010	0.009	-0.657	-0.011
North Carolina	0.166	0.169	0.486	0.008
New Hampshire	0.001	0.001	-0.126	-0.002
New Mexico	0.005	0.006	0.818	0.015
Ohio	0.264	0.302	4.933	0.087
Pennsylvania	0.189	0.160	-4.377	-0.074
Rhode Island	0.004	0.004	0.557	0.010
Application Screener				
Lost Income COVID	0.954	0.960	1.856	0.031
Expenses Increased COVID	0.899	0.900	0.270	0.005
Got CARES Stimulus	0.696	0.695	-0.119	-0.002
COVID Hardship	4.540	4.571	2.393	0.041
COVID Kids Difficulty	4.426	4.457	2.307	0.040
N	5777	7915		

Note: The SUEST F statistic is 144.32 with a p-value of 0, which indicates statistically significant imbalance between treatment and control users.

Table B1 presents treatment and control means among the set of roughly 13,700 individuals who were eligible for and agreed to participate in the baseline study, including the 345 individuals who were inadvertently excluded from the baseline survey. We see that individuals in the treatment group started using the Fresh EBT app roughly one month later than the control group, a difference of roughly 0.08 SD that is statistically significant. In addition, individuals in the treatment group were more likely to live in Kentucky and Ohio, and less likely to live in Michigan and Pennsylvania. While users in the treatment and control groups were equally likely to indicate they had lost income and increased expenses as a result of COVID, users in the treatment group reported slightly greater levels of hardship than those in the control group. The mean response to screener questions relating to material hardship and difficulty with children were roughly 0.04 SD higher among the treatment group. The F-statistic (p-value) of the joint hypothesis test that all estimates are zero is 144.32 (0.000), suggesting significant imbalance between the treatment and control groups. Permutation-based p-values generate comparable results.

Unaffected Sample

A subset of users included in Table B1 should be unaffected by the implementation concerns described above ("restricted sample"). Specifically, individuals in segments 0-200 (treatment) and 500-600 (control) had all been exposed to prior GD campaigns and were all exposed to study banners between 7am and 11am. Table B2 shows treatment-control balance among individuals in this restricted subsample who responded to the initial survey.

Table B2: Balance Table of Baseline Survey Respondents, from Restricted Subsample

	Control Mean	Treatment Mean	T-Statistic	Standardized Difference
Application Data				
Days using Fresh EBT	316.474	311.423	-0.334	-0.013
EBT Early	0.512	0.511	-0.057	-0.002
Day of Account	0.024	0.019	-0.925	-0.036
Number of App Uses	173.237	174.008	0.098	0.004
SNAP Benefits	357.431	360.913	0.292	0.014
Missing Benefits	0.300	0.307	0.380	0.015
Spanish Language	0.021	0.015	-1.144	-0.044
State				
Indiana	0.041	0.031	-1.245	-0.049
Kentucky	0.026	0.028	0.413	0.017
Louisiana	0.043	0.045	0.265	0.011
Maine	0.001	0.002	0.760	0.035
Michigan	0.248	0.239	-0.540	-0.022
Minnesota	0.009	0.011	0.500	0.021
North Carolina	0.170	0.175	0.325	0.013
New Hampshire	0.003	0.002	-0.723	-0.027
New Mexico	0.003	0.005	0.665	0.029
Ohio	0.264	0.280	0.855	0.035
Pennsylvania	0.187	0.178	-0.617	-0.025
Rhode Island	0.003	0.003	-0.081	-0.003
Application Screener				
Lost Income COVID	0.934	0.966	3.406	0.128
Expenses Increased COVID	0.900	0.905	0.440	0.018
Got CARES Stimulus	0.724	0.716	-0.471	-0.019
COVID Hardship	4.494	4.525	0.996	0.039
COVID Kids Difficulty	4.435	4.416	-0.623	-0.026
N	859	2121		

Note: The SUEST F statistic is 22.15 with a p-value of 0.51, which indicates no statistically significant imbalance.

In the restricted sample, response rates were similar but not identical across groups: 66.6 percent for control and 70.1 percent for treatment. Here we see very few treatment-control differences. The F-statistic on a joint test is 22.15, with a p-value of 0.51. The permutation-based p-value is 0.81, providing additional evidence of covariate balance. Appendix Table 4 presents results for our five primary outcomes in this restricted subsample. As the restricted sample did not exhibit the imbalances between treatment and control and were unaffected by implementation concerns, we opted to invite the 1,558 individuals who did not respond to the baseline survey to take the one-month follow-up survey.

Implications for Analysis of Balanced Panel

The evidence presented above suggests that several known implementation problems led to modest treatment-control imbalance in the full study sample. If the imbalance is limited solely to these observable characteristics, controlling for these variables will allow us to recover causal estimates of the intervention. Our primary concern is that the imbalance of observables signals additional differences in unobservable characteristics. In theory, one way to test this concern is to compare treatment and control responses on the initial survey. However, individuals completed this survey after learning whether they were awarded the cash payment and, in some cases, after they received the money. While this should not have affected responses to some questions (e.g., household demographics, past benefit receipt), it is possible that exposure to the treatment may have influenced responses to some other measures.