POLICY RESEARCH WORKING PAPER 9769

Intervention Size and Persistence

Florence Kondylis
John Ashton Loeser

WORLD BANK GROUP

Abstract

Do larger interventions improve longer run outcomes more cost effectively? And should poverty traps motivate increasing intervention size? This paper considers two approaches to increasing intervention size in the context of temporary unconditional cash transfers—larger transfers (*intensity*), and adding complementary graduation program interventions (*scope*). It does so leveraging 38 experimental estimates of dynamic consumption impacts from 14 developing countries. First, increasing intensity decreases cost effectiveness and does not affect persistence of impacts.

This result can be explained by poverty traps or decreasing marginal return on investment in a standard buffer stock model. Second, increasing scope increases impacts and persistence, but reduces cost effectiveness at commonly evaluated time horizons and increases heterogeneity. In summary, larger interventions need not have more persistent impacts, and when they do, this may come at the expense of cost effectiveness, and poverty traps are neither necessary nor sufficient for these results.

Intervention Size and Persistence[*]

Florence Kondylis[†]   John Loeser[†]

Contents

1 Introduction

"Big push" interventions are commonly proposed to generate large, sustained increases in household, community, and national income (Kraay & McKenzie, 2014; Banerjee et al., 2020). At the household level, two approaches to increasing the size of interventions may, in theory, enable households to escape poverty traps and produce persistent decreases in poverty (Ghatak, 2015). When households are in a "scarcity poverty trap", increasing the *intensity* of interventions can push households over a poverty threshold. Alternatively, when households face "frictional poverty traps", increasing the *scope* of interventions can enable households to overcome multiple constraints. However, formally evaluating increasing intervention intensity and scope requires measures of the impact of these approaches on cost effectiveness (Banerjee et al., 2015).

Do larger interventions improve longer run outcomes more cost effectively? And should poverty traps motivate increasing intervention size? The role of intervention size is particularly important in the context of temporary cash transfers, which are increasingly used as a benchmark for cost effective poverty reduction (Haushofer & Shapiro, 2016; McIntosh & Zeitlin, 2018). Cash transfer *intensity* corresponds to increasing transfer size — standard theory (Deaton, 1991; Carroll, 2019) suggests larger cash transfers should increase household consumption relatively more cost effectively over longer time horizons. Cash transfer *scope* corresponds to adding complementary interventions to target multiple constraints[1] — a growing body of evidence suggests adding interventions that target multiple constraints complements temporary cash transfers (Banerjee et al., 2018; Bossuroy et al., 2021). For both approaches to increasing size, evidence on the dynamics, and therefore the persistence, of cost effectiveness is limited. Producing this evidence requires variation in size, timing of estimates, and a large sample for statistical power (Muralidharan et al., 2019).

In the context of temporary cash transfers, we find that increasing intensity reduces cost

---

[1]These notions of size are distinct from *scale*, the impacts of which are reviewed and analyzed at length by Muralidharan & Niehaus (2017).

effectiveness, while increasing scope increases persistence but reduces cost effectiveness at commonly evaluated time horizons. To produce these results, we leverage a deep literature that has experimentally estimated the impacts of temporary unconditional cash transfers and multifaceted graduation programs. We pool across studies, and exploit variation in both intervention size and timing of estimates. We present theory that suggests these results are fully consistent with the presence of both scarcity and frictional poverty traps, or alternatively with decreasing returns to investment, in a standard buffer stock model. This challenges the conventional wisdom that the presence of poverty traps implies larger interventions are more cost effective.

Our analysis includes 38 estimates of impacts on household consumption from 17 randomized control trials (RCTs) of temporary unconditional cash transfers (UCT) and multifaceted graduation programs that include temporary unconditional transfers as a component ("Targeting the Ultra Poor", or TUP) in 14 developing countries. We focus on household consumption for three main reasons. First, economic theory makes strong predictions on the persistence of UCT of varying sizes (Carroll & Kimball, 1996). Second, impacts of UCT on consumption are valuable for disciplining key parameters in economic models of interest (Kaplan & Violante, 2014; Auclert et al., 2018). Third, UCT and TUP programs are commonly motivated as targeting poverty reduction through increasing consumption (Banerjee et al., 2015; Bedoya et al., 2019). For each estimate, we collect impacts on annualized household consumption and standard errors, transfer size, and program cost in USD PPP.[2] To measure cost effectiveness, we focus our analysis on effects on consumption normalized either per dollar of transfer or program cost. We further leverage variation both within and across trials to estimate the effects of the size of transfers and the time since transfers occurred, and we compare estimates across UCT and TUP programs.

We begin by aggregating UCT estimates across contexts, and find the average UCT increased consumption by 0.35 per unit of transfer. We use tools from the metaänalysis

---

[2]We convert reported estimates, transfer size, and costs from each study in USD PPP indexed to 2010 dollars to enable comparison across studies, following Meager (2019).

4

literature to aggregate estimates that account for both the sampling variance in estimated effects and the variance of true effects across contexts and interventions (Baird et al., 2014; Burke et al., 2015; Vivalt, 2020). We estimate a tight standard deviation of true effects across contexts relative to estimates for other types of interventions (Vivalt, 2020), with a coefficient of variation of true effects of 0.34, in line with the relative homogeneity of UCT interventions. We find consistent results across methods.

We provide a theory of UCT size and persistence building on Carroll & Kimball (1996), to derive predictions under different assumptions on households' production technologies. We begin by showing that, in a benchmark model with constant returns to scale technology and no poverty traps, larger UCT should have relatively larger impacts as time since transfers grows. This is because, in a broad class of models of intertemporal optimization that includes ours, households' marginal propensity to consume is decreasing in income. We present two extensions that weaken this prediction. First, if households face decreasing returns to scale, larger transfers decrease households' marginal return on investment, reducing their optimal investment and potentially leading to smaller impacts as time since transfers grows. Second, if households face poverty traps, larger transfers may lead to smaller impacts as time since transfers grows if the density of the distance of households from a scarcity poverty threshold is decreasing in transfer size (Balboni et al., 2020).

We find that larger UCT have smaller impacts on consumption per unit of transfer, at both short and long time horizons. These results are robust to the inclusion of controls, and using either randomized within variation or both within and across study variation in transfer size and time since transfers. If anything, across specifications, point estimates suggest *more* persistent impacts of *smaller* UCT. These results contrast with the prediction from our benchmark constant returns to scale model, that larger UCT should have more persistent impacts. However, either decreasing returns to scale or poverty traps can explain these results in our theoretical model. In the case of poverty traps, explaining our results would require smaller cash transfers to push more households out of poverty per unit of

transfer – this is consistent with findings from Balboni et al. (2020), who estimate a poverty threshold of 504 USD PPP in Bangladesh, half the size of the average transfer in our sample.

To further motivate the geographic focus of this metaänalytic effort and interpret these magnitudes, we compare our estimates to comparable ones from richer countries (Auclert et al., 2018; Fagereng et al., 2019). UCT impacts are more persistent in developing countries – one year impacts of UCT on household consumption in developing countries are smaller, while impacts at three or more years are larger. In our model, this result is consistent with a larger marginal return on investment in developing countries, a well documented stylized fact (De Mel et al., 2008; McKenzie & Woodruff, 2008; Hussam et al., 2020). To provide additional context for our estimates, we note that cumulative impacts on consumption over the first three years are larger than the size of transfers. Following the approach of Banerjee et al. (2015), this provides strong evidence that UCT pass a cost-benefit analysis.

We next consider the impacts of adding complementary interventions in multifaceted graduation programs. Our estimates imply that TUP complementary interventions increase impacts on consumption. These complementary interventions are relatively expensive in our sample, and as a result the average TUP is 5%-43% less cost effective at increasing consumption than the average UCT. However, we also find evidence of greater heterogeneity in the impacts of TUP relative to UCT interventions, suggesting TUP may be more cost effective than UCT in a range of contexts. Lastly, the relative cost effectiveness of TUP increases over time, and our estimates imply TUP increases cumulative consumption per unit of program cost by more than UCT after 3.4-7.7 years; we note that only 4 of the 20 TUP estimates in our metaänalysis sample are more than 3 years since last transfer, highlighting the need for additional long run estimates.[3]

This study makes three central contributions. First, we contribute to a large literature on

---

[3]Viewed in isolation, this increasing cost effectiveness is potentially consistent with perfect markets for TUP complementary interventions through the lens of a standard buffer stock model, as the implied internal rates of return from these estimates are similar to estimates of marginal returns to investment in developing countries. However, contrasted with our estimates of the impacts of increasing transfer size, it instead suggests a role of failures in the markets for these complementary interventions, and potentially also for frictional poverty traps.

the impacts of temporary cash transfers on household consumption. In this literature, a large body of work has leveraged random and quasi-random income shocks to estimate short run impacts in richer countries (Jappelli & Pistaferri, 2010) and developing countries (Haushofer & Shapiro, 2016). More recent work in Norway has taken advantage of high frequency household balance sheet data to estimate dynamic impacts leveraging natural experimental variation in cash transfers (Fagereng et al., 2019). Aggregating across experimental UCT studies meets the data requirements to produce comparable dynamic estimates for developing countries. These impacts can also be interpreted as intertemporal marginal propensities to consume, key parameters for disciplining models of intertemporal optimization subject to constraints (Auclert et al., 2018). Second, we contribute to a deep literature that investigates the role of poverty traps and implements empirical tests (Carter & Barrett, 2006; Kraay & McKenzie, 2014; Balboni et al., 2020). Complementarily, recent work leveraging experimental variation has consistently found large UCT and TUP programs increase household consumption 4 or more years after initial transfers, suggesting these programs enable households to escape from poverty (Bandiera et al., 2017; Blattman et al., 2020). Through a metaänalytic lens, we provide a relatively powered test of the relationship between intervention size and the impacts of cash transfers on household consumption over time. In addition, we theoretically link this relationship to the existence of poverty traps. Third, we contribute to a growing literature in economics that uses metaänalysis to aggregate estimated program impacts across studies (Baird et al., 2014; Burke et al., 2015; Meager, 2019; Vivalt, 2020). We do so by providing a template for future "metabenchmarking" of the cost-effectiveness of one intervention (here, the cost-effectiveness of either large UCT or TUP at increasing household consumption) against UCT.

In summary, we present a nuanced view of the impacts of intervention size on cost effective sustained poverty reduction in the context of temporary cash transfers and their theoretical underpinnings. We produce evidence that increasing intervention intensity, through increased transfer size, reduces cost effectiveness in the short and medium run. In our model,

one plausible explanation is that smaller transfers push more households across the poverty threshold per unit of transfer — scarcity poverty traps may therefore provide a justification for lower, rather than higher, intensity interventions. On the other hand, we find increasing intervention scope, through complementary interventions, increases long-run cost effectiveness; the contrast between this result and the effect of increasing intervention intensity is consistent with the presence of frictional poverty traps. The long-run cost effectiveness of complementary interventions only manifests at relatively distant time horizons; while frictional poverty traps may justify increasing intervention scope, high costs of complementary interventions can undermine cost effectiveness. While targeted larger interventions may be desirable from a policy standpoint, our results highlight that poverty traps are neither necessary nor sufficient to justify larger interventions.

This paper is organized as follows. Section 2 describes the metaänalysis sample and data extraction. Section 3 estimates average effects. Section 4 estimates the impacts of increasing transfer size, Section 5 compares our estimates to those from richer countries, and Section 6 estimates the impacts of complementary interventions. Section 7 concludes.

2   Data and Context

In this section, we describe the construction of the tables of studies, interventions, and estimates, including impacts of UCT and TUP interventions on consumption and intervention characteristics.[4]

2.1   Study inclusion

Our metaänalysis focuses on randomized control trials of temporary unconditional cash transfers (UCT) and multifaceted graduation programs that include unconditional cash and/or

---

[4]The full extracted data from the studies, including both the reported point estimates and the impacts on consumption in 2010 USD PPP that we use for analysis, are made available at `https://docs.google.com/spreadsheets/d/1txIOHFWQcLjmHllUv2b-AzmyhcnvxtOALLXdbG5KDjs`.

asset transfers as a component (TUP) from lower and middle income countries. We classify programs as "UCT" when they include cash or mobile money transfers, and do not include any conditionalities stricter than attendance at meetings at the frequency of transfer payments. We allow for attendance requirements because these are in principle similar to distribution of cash at a centralized location.[5]

We restrict our analysis to studies that meet the following criteria:[6]

1. Consumption over a defined reference period is an outcome

2. Value of transfers and program cost are reported

3. The time of estimates is after transfers are completed

4. Includes a randomized control group that did not receive the intervention at the time of the estimate

5. Treatment is randomized across households, or authors argue that spillover effects are small

We require (1) and (2) to construct our primary outcomes of interest, impacts on household consumption per unit of transfer and per unit of cost. We require (3) because continuing and temporary transfers both empirically and theoretically have different types of impacts (Pennings, 2021). In addition, temporary transfers are the relevant benchmark for a wide range of temporary development interventions, including the TUP programs we consider in Section 6. We require (5) because our primary interest is in partial equilibrium effects. Empirically, a majority of candidate studies satisfied this requirement, while a much smaller number of studies argued their estimates reflect both partial and general equilibrium effects.

---

[5]We provide two examples of programs that provide unconditional cash transfers with meetings or trainings, one we classified as UCT and another as TUP, that clarify the distinction. Blattman et al. (2016), which we classify as TUP, study a program that provided a one time cash grant of 150 USD cash, eight days of training covering business and group dynamics, and five supervisory visits from program staff. Carneiro et al. (2020), which we classify as UCT, study a program that provided 25 months of 22 USD cash transfers to women during and following pregnancy, along with community level messaging on health information, and for a subset of households optional monthly meetings on child feeding support and options to request one-on-one counseling.

[6]To identify studies, we follow the approach of Croke et al. (2016). We began with the sample from GiveDirectly Cash Research Explorer (primarily UCT) and articles cited by or citing Banerjee et al. (2015) (primarily TUP). We then included any additional studies we could find that satisfied our inclusion criteria.

Theoretically, as we discuss in Section 4.1, partial equilibrium effects allow us to test standard models of household intertemporal optimization, and recover intertemporal marginal propensities to consume that are useful targets for calibration of macroeconomic models.

2.2 Data extraction

Total transfer size and total program cost   To calculate total transfer size, we calculated the average sum of the value of all transfers made to beneficiary households. Program costs were similarly calculated by adding the total value of all transfers to any additional reported costs of implementation per beneficiary household.

Years since last transfer   To capture the dynamics of the impacts of cash transfers, we use years since last transfer as our standardized measure of time for each estimate to enable comparisons across studies. We use years between the average time households received their last transfer and the average time of the survey wave. We use time since last transfer, rather than, for example, time since first transfer, to be conservative with respect to finding large persistence of impacts of cash transfers on consumption. This choice is also impactful when we compare across interventions with different transfer durations — we therefore include robustness to the inclusion of transfer duration as a control for specifications comparing across interventions.

Impacts on household consumption per unit of transfer and per unit of cost There is substantial variation in how papers report impacts on household consumption. To construct consistent measures of impacts on household consumption, some conversions were necessary. We applied the following set of rules (in order) to construct impacts on household consumption from each paper:

1. Impacts on household consumption were prioritized over impacts on log household consumption or per capita household consumption

    a. Impacts on log household consumption were multiplied by control mean household consumption (or baseline household consumption when control mean household consumption was not reported)

    b. Impacts on per capita household consumption were multiplied by baseline mean household size (or average household size from a representative survey from the same country when mean household size was not reported)

2. Impacts on household consumption were prioritized over impacts on non-durable household consumption

3. When impacts of multiple distinct aggregations of treatment arms were reported, estimates with more disaggregated treatment arms were prioritized, and estimates with disaggregation by transfer size were prioritized

4. Estimates from the authors' preferred specification were prioritized

5. Impacts on household consumption were annualized if reported over a different reference period

Intent-to-treat and treatment-on-the-treated Compliance for both UCT and TUP programs were almost always near universal, as the programs studied by papers in this analysis typically offered large transfers relative to household income and contamination was rare, but there is meaningful variation in compliance. To ensure appropriate comparison, in all cases we therefore use treatment-on-the-treated instead of intent-to-treat estimates, scaling intent-to-treat impacts on household consumption by the inverse of the impacts of assignment on participation. When transfer sizes and program costs were reported per household assigned to treatment (instead of per beneficiary), transfer sizes and program costs also needed to be scaled.

Standardization of monetary units As a final step, all monetary values were converted to 2010 USD PPP to facilitate comparison across contexts. In addition, this conversion was done before impacts on household consumption per unit of transfer or cost were constructed;

this was intended to avoid bias towards finding growth of impacts on household consumption per unit of transfer over time in the presence of inflation.

2.3   Descriptive statistics

Descriptive statistics on the UCT and TUP interventions, and on the estimated impacts on consumption of these interventions, are reported in Table 1. As much of our analysis focuses on impacts of UCT, we begin our discussion of these interventions. The average transfers in these studies are quite large, representing 0.62 years of consumption on average, but there is meaningful variation in transfer size across interventions. Transfers last 8 months on average, so any increase in consumption during those 8 months caused by the UCT will not typically be accounted for in our results. The transfers in UCT are relatively low cost to deliver, with the average UCT in our sample reporting 18% overhead. The 14 interventions from 7 RCTs comprising our sample represents a large base of evidence compared to many other categories of interventions in development economics — Vivalt (2020) identifies more estimates only for micronutrient supplementation, conditional cash transfers, and deworming, while using less strict inclusion criteria.

The average estimated effect of UCT is 0.58 per unit of transfer, with a standard deviation of estimates across studies of 0.57. This variation could be explained either by variation in true effects of UCT across contexts or by sampling error. The average standard error of estimates is 0.28, heuristically suggesting sampling error is responsible for a large share of the variation in estimated effects. We revisit this formally and interpret these magnitudes in Section 3.

As an alternative approach to understand variation in our estimates, we plot point estimates and 95% confidence intervals for UCT and TUP interventions in Figure 1.[7] The UCT estimates appear tightly clustered, with confidence intervals for only 6 of the 18 UCT estimates excluding 0.4. This is despite the heterogeneity across interventions in size and

---

[7]Figure 1 also contains posterior mean estimates and credible intervals for these estimates, which we discuss in Section 3.2.

Table 1: Descriptive statistics

	UCT			TUP		
	Mean (1)	SD (2)	# of obs. (3)	Mean (4)	SD (5)	# of obs. (6)
Panel A: Interventions						
Total transfer size (2010 USD PPP)	963	689	14	952	496	11
Total transfer size (Years of baseline consumption)	0.62	0.42	13	0.29	0.23	11
Baseline annualized consumption (2010 USD PPP)	1573	475	13	4188	2504	11
GDP per capita (2010 USD PPP)	2384	2132	14	3464	2340	11
Africa	0.93	0.27	14	0.36	0.50	11
Transfer duration (Years)	0.66	0.55	14	1.35	0.78	11
Year of last transfer	2015.4	3.3	14	2011.5	2.6	11
Total intervention cost (2010 USD PPP)	1102	751	14	3315	1778	11
Intervention cost per unit of transfer	1.18	0.11	14	3.93	2.04	11
Panel B: Estimates						
Effect on annualized consumption (2010 USD PPP)	416	251	18	509	515	20
Standard error (2010 USD PPP)	157	86	18	201	125	20
Effect on annualized consumption per unit of transfer	0.58	0.57	18	0.50	0.45	20
Standard error	0.28	0.30	18	0.21	0.12	20
Inverse variance weight	0.06	0.08	18	0.05	0.04	20
Years since last transfer	1.5	2.2	18	2.6	2.0	20
Sample size	1620	942	18	3127	5411	20

Notes: Descriptive statistics for interventions are presented in Panel A, and for their associated estimates in Panel B. Statistics for UCT interventions and associated estimates are in Columns 1 through 3, while statistics for TUP interventions and associated estimates are in Columns 4 through 6. For each variable, Columns 1 and 4 present sample means, Columns 2 and 5 present sample standard deviations, and Columns 3 and 6 present the number of non-missing observations.

duration, and heterogeneity across estimates in time — estimates are on average 1.5 years after the last transfer, with a standard deviation of 2.2 years, providing a range of estimates at both short and long time horizons. We also note that the estimates with tighter confidence intervals typically come from studies with relatively larger cash transfers; we revisit this in our discussion of power in Appendix D.1.

Lastly, we compare characteristics of TUP and UCT interventions and estimates. TUP and UCT interventions provide similar size transfers on average, but differ in a number of important characteristics. TUP interventions are typically implemented with higher consumption households in higher income countries, and are much less likely to be in Africa. Transfers last 16 months on average, twice as long as for UCT interventions. In addition, TUP interventions are typically much more expensive, with additional components complementary to cash transfers — the average TUP intervention includes 2.7 units of costs per

Figure 1: Estimates and posterior estimates
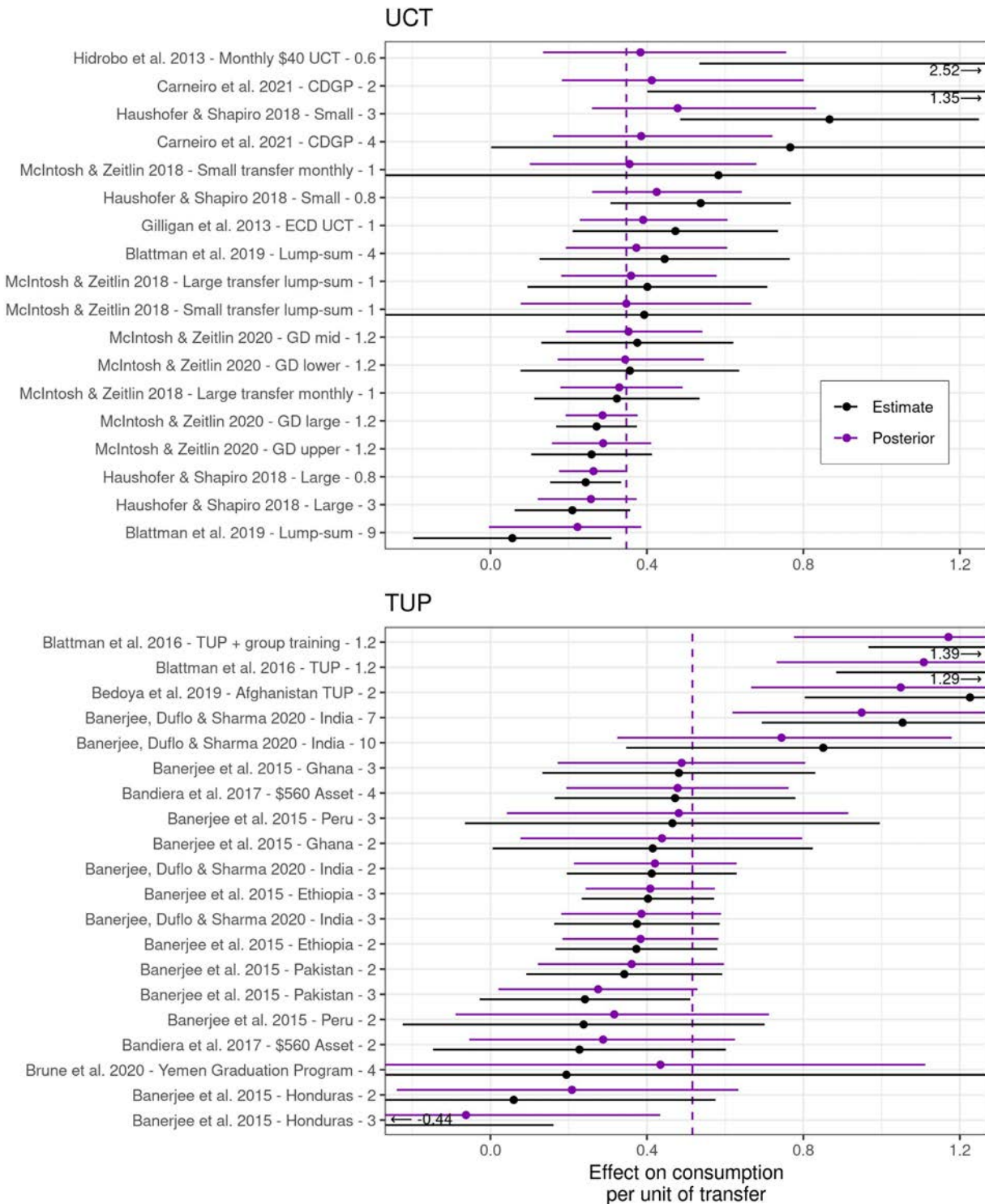
Notes: Estimates of the impact of UCT and TUP interventions on annualized consumption per unit of transfer are presented in this figure. Raw estimates and 95% confidence intervals are presented in black. Bayesian posterior means and 95% credible intervals from the model in Column 5 of Table 2 are presented in purple, with dotted lines to mark the average effect.

14

unit of transfer more than the average UCT intervention. In addition, the TUP interventions in our sample preceded the UCT interventions by 4 years on average, and as a result estimates are on average 1.1 years since last transfer later. In Section 6, we conduct analysis benchmarking the impacts of TUP against UCT interventions, and test robustness to adjusting for key differences between the TUP and UCT interventions in our sample.

3 Metaänalysis

3.1 Estimation strategy

Our objective is to aggregate estimates of the impacts of UCT from a set of randomized control trials (RCTs), each of which may have multiple UCT arms and multiple survey waves. Before testing for evidence of nonlinearities, we begin with a traditional metaänalysis, estimating both the average impact of UCT across estimates and heterogeneity in impacts across contexts. We let r index RCTs, a index experimental arms within RCT, and t be years since the last transfer for each estimate. We observe a set of estimates $\hat{\tau}_{rat}$ of the impact of UCT on household consumption per unit of transfer for arm a of RCT r, t years since the last transfer, in addition to the standard error on each estimate $\hat{\text{se}}_{rat}$.

We model these estimates as reflecting a combination of both sampling error and true variation in impacts across contexts. Following recent metaänalyses in economics (Baird et al., 2014; Burke et al., 2015; Meager, 2019; Vivalt, 2020), we specify

$$\hat{\tau}_{rat}|\tau_{rat} \sim N(\tau_{rat}, \hat{\text{se}}_{rat}^2)$$
$$\tau_{rat} \sim N(\beta, \sigma_\tau^2)$$

(1)

where τ_{rat} is the true impact of arm a of RCT r, t years since the last transfer. The first line of Equation 1 states that the estimated effect is normally distributed, is unbiased for the true effect, and has variance equal to the reported standard error squared.[8] The normality

---

[8]We note that this implicitly imposes no within RCT correlation between estimated effects, while within

assumption is motivated by a central limit theorem, and each estimate coming from a large sample. The second line of Equation 1 states that the true effect is normally distributed with mean β and variance σ_τ^2. While the true effects are unlikely to be normally distributed, this model in its limit nests both assuming full external validity, with an identical impact across all contexts ($\sigma_\tau = 0$), and assuming no external validity, with the average impact being completely uninformative about the impact in any context ($\sigma_\tau \to \infty$). In this sense, this approach is agnostic about the source of variation in heterogeneity of impacts across estimates, which could be driven by characteristics of interventions, experimental samples, contexts, aggregate shocks (Rosenzweig & Udry, 2019), or their interaction.

We take three approaches to estimating parameters of the above model, with a focus on estimating β, the average impact of UCT on consumption per unit of transfer.

First, we begin by estimating β by taking the inverse variance weighted average of estimated impacts $\hat{\tau}_{rat}$, which we estimate using weighted least squares.[9] This is equivalent to a pooled regression across estimates under homoskedasticity. It is also equivalent to maximum likelihood estimation of β when fixing $\sigma_\tau = 0$, and is therefore efficient when the true impact is identical across contexts.

Second, we estimate (β, σ_τ) jointly by maximum likelihood. The estimated β can still be interpreted as a weighted average of the estimated impacts $\hat{\tau}_{rat}$, but the weights are now the inverse of unconditional variance of the estimated effects, which is the sum of both the sampling variance ($\hat{se}_{rat}^2$) and the estimated variance of true effects (σ_τ^2). The estimated variance of true effects, in turn, reflects the difference between the empirical variance of estimated effects, and what one would expect the variance of estimated effects to be only accounting for sampling error.

Third, our preferred approach, we closely follow Burke et al. (2015) and Vivalt (2020) and estimate Equation 1 using Bayesian methods. Specifically, we impose a uniform prior

---

RCT estimated effects are correlated both through the control group (across treatment arms) and due to serial correlation (across time). Later in this section, we discuss how these correlations may affect inference.

[9]For inference with this approach, we calculate robust standard errors clustered at the RCT-level.

on σ_τ with wide support[10] and a uniform prior on $\beta|\sigma_\tau$, while Equation 1 provides the prior for $\tau|\beta, \sigma_\tau$. Together, these provide a joint prior on the unknown parameters $(\beta, \sigma_\tau, \tau)$. We then estimate our posterior over $(\beta, \sigma_\tau, \tau)$ given the observed $(\hat{\tau}_{rat}, \hat{se}_{rat})$. To facilitate comparison between Bayesian posteriors of (β, σ_τ) and weighted least squares and maximum likelihood point estimates and standard errors, we report posterior means, posterior standard deviations, and 95% credible intervals for (β, τ) in place of point estimates, standard errors, and p-values, respectively.

Our maximum likelihood and Bayesian estimates can then be used to produce more efficient estimates of each true treatment effect, τ_{rat}. With maximum likelihood, with point estimates for (β, σ_τ), this involves shrinking each estimate $\hat{\tau}_{rat}$ towards the average effect β. Specifically, the conditional expectation of the true effect given the estimated effect is $\mathbf{E}[\tau_{rat}|\beta, \sigma_\tau, \hat{\tau}_{rat}, \hat{se}_{rat}^2] = \frac{\hat{se}_{rat}^2}{\sigma_\tau^2 + \hat{se}_{rat}^2}\beta + \frac{\sigma_\tau^2}{\sigma_\tau^2 + \hat{se}_{rat}^2}\hat{\tau}_{rat}$, a weighted average of the average effect β and the estimated effect $\hat{\tau}_{rat}$, with more weight on the estimated effect when the variance of true effects is large and when the estimated effect is precisely estimated. We can similarly calculate the asymptotic variance of the error of this shrunken estimate to construct tighter confidence intervals on the true effect. Alternatively, the Bayesian approach described above yields a posterior for each τ_{rat}, which can be used to construct posterior means and credible intervals.

Regarding inference, we note that Equation 1 leaves unmodeled correlation in estimated effects within RCT, both across arms because of a common control group and over time because of serial correlation. These correlations are not typically reported by studies, although they could be estimated with microdata from each RCT. To test robustness of our inference, when estimating Equation 1 by weighted least squares or by maximum likelihood, we report both analytical standard errors and standard errors estimated by block bootstrap at the RCT-level for estimates of (β, σ_τ).

---

[10] We follow Vivalt (2020) and specify the support from 0 to 10 times the standard deviation of $\hat{\tau}_{rat}$.

3.2   Results

We present our estimates of average effects on consumption per unit of transfer from Equation 1 in Table 2. While we focus our discussion in this section on estimates for UCT in Panel A, we also report estimates for multifaceted graduation programs (TUP) in Panel B to support our comparison between UCT and TUP in Section 6. We focus the discussion primarily on our Bayesian estimates in Column 5, except when we discuss the robustness of our results to alternative approaches to estimation and inference.

First, our preferred estimate of the average impact of UCT on consumption per unit of transfer in Column 5 implies that for every unit of transfer, annualized consumption increases by 0.35. We make three observations regarding this estimate. First, our estimate is smaller than the average impact of ongoing cash transfers from safety nets on consumption per unit of transfer (0.74, from Ralston et al. (2017)). However, our estimates are for completed transfers, and are therefore likely to be smaller, as households may save more out of a temporary income shock than a more permanent income shock. Second, our estimate is smaller than estimates of the one year impacts of cash transfers in richer countries (0.5, from Auclert et al. (2020)). We rationalize this result through the lens of a model of intertemporal substitution and discuss it further in Section 4.1. Third, our estimates are likely to be heterogeneous, as they may vary by time horizon, transfer size, and context.

Second, our preferred estimate of the standard deviation of impacts of UCT on consumption per unit of transfer in Column 5 is 0.12. Taking seriously the assumption of normality of the distribution of effects, it implies that 95% of the impacts of cash transfers on household consumption per unit of transfer are between 0.12 and 0.58. We note that this estimate implies a coefficient of variation for impacts of cash transfers of 0.34 — this is smaller than all of the estimates for a range of interventions from Vivalt (2020), suggesting that impacts of unconditional cash transfers on consumption are relatively generalizable, consistent with the homogeneity of the intervention.[11]

---

[11]In Appendix B.1, we show that our results on the mean and standard deviation of effects of UCT on

Table 2: Average effects and standard deviation of effects

| | Effect on consumption per unit of transfer | | | | |
| | WLS | | MLE | | Bayes |
	(1)	(2)	(3)	(4)	(5)
Panel A: UCT					
β: Intercept (Mean effect)	0.298	0.298	0.335	0.335	0.347
	(0.014)	(0.029)	(0.048)	(0.074)	(0.054)
	[0.000]	[0.000]	[0.000]	[0.000]	{0.259, 0.471}
σ_τ: Std. dev. of effects			0.083	0.083	0.117
			(0.060)	(0.083)	(0.067)
					{0.011, 0.269}
# of RCTs	7	7	7	7	7
# of observations	18	18	18	18	18
Panel B: TUP					
β: Intercept (Mean effect)	0.469	0.469	0.517	0.517	0.516
	(0.071)	(0.072)	(0.091)	(0.118)	(0.102)
	[0.000]	[0.000]	[0.000]	[0.000]	{0.314, 0.719}
σ_τ: Std. dev. of effects			0.350	0.350	0.395
			(0.080)	(0.134)	(0.096)
					{0.240, 0.614}
# of RCTs	10	10	10	10	10
# of observations	20	20	20	20	20
Bootstrapped standard errors		X		X	

Notes: Estimates of the model of impacts of interventions on household consumption per unit of transfer in Equation 1 are presented in this table. Estimates for UCT interventions are in Panel A, and estimates for TUP interventions are in Panel B. Columns 1 through 4 report estimates of each parameter, with standard errors in parentheses and p-values in square brackets, while Column 5 reports Bayesian posterior means of each parameter, with posterior standard deviations in parentheses and 95% credible intervals in curly brackets. Column 1 uses robust standard errors clustered at the RCT-level, while Columns 2 and 4 use standard errors block bootstrapped at the RCT-level.

Our estimates of the average effect of unconditional cash transfers are robust to the choice of methods. We estimate an average impact of 0.30 under weighted least squares and 0.34 using maximum likelihood, in comparison to a posterior mean estimate of 0.35 using Bayesian methods. As these approaches appear to yield comparable estimates in our context, we report our preferred estimates using Bayesian posterior means for the remainder of this paper. For inference, Bayesian posterior standard deviations are comparable to block

consumption per unit of transfer are robust to the exclusion of any individual RCT in our sample.

bootstrapped standard errors from maximum likelihood – they are 27% smaller for UCT in Panel A, and 14% smaller for TUP in Panel B. We interpret this as evidence that the bias in inference from unmodeled within-RCT correlation between estimates mentioned in Section 3.1 is relatively small, and focus our discussion in the remainder of the paper on our Bayesian estimates.

The estimated average impacts on consumption per unit of transfer is meaningfully smaller in UCT than in TUP studies; we explore this dimension of heterogeneity further in Section 6.

Leveraging our estimates of the average impact and the standard deviation of impacts of UCT and TUP per unit of transfer, we plot the full set of posterior estimates and 95% credible intervals for each of the estimates in our metaänalaysis in Figure 1. The posterior estimates are equal to the reported estimates, but shrunken towards our estimates of average impacts, as described in Section 3.1. As we estimate a relatively tight standard deviation of impacts of UCT, most estimates are shrunken quite close to the posterior mean (and particularly those estimates with wide standard errors). As more precise estimates receive higher weight when estimating the average impacts of UCT and TUP, Figure 1 visualizes which estimates drive our result.

With the average impacts of cash transfers on household consumption per unit of transfer analyzed, we now return to our motivating question: how do the effects of large and small cash transfers differ over time? In particular, are the impacts of large cash transfers more persistent per dollar transferred? In the next section, we model consumption responses to cash transfers to identify a theory consistent test of the relative persistence of larger cash transfers and implement this test.

4 Transfer size ("Intensity")

In this section, we model and estimate the differential persistence of the impacts of larger UCT on consumption per unit of transfer. In Section 4.1, we show that a decreasing marginal propensity to consume implies that larger UCT should have relatively larger impacts on consumption per unit of transfer as time since last transfer increases, and link this to models of intertemporal optimization. In Section 4.2, 4.3, 4.4, and 4.5, we test this prediction using variation in transfer size and time since last transfer within and across randomized control trials of UCT.

4.1 Model

4.1.1 Environment

We consider a discounted expected utility maximizing household deciding, in each time period, how much of its available resources to consume and how much to invest. To simplify exposition, we assume the household lives for two periods, indexed 0 and 1. In period 0, the household begins with cash-on-hand $w_0 + h$, where w_0 is the household's own resources and h is a UCT the household receives. The household consumes c_0, and invests the remainder; the household cannot borrow, so investment $w_0 + h - c_0 \geq 0$. The household has a stochastic constant returns to scale investment technology — in period 1, the household has total resources $R(w_0 + h - c_0) + y_1$, where R is the stochastic return on investment with expectation ρ, and y_1 is the household's stochastic income from other sources.[12] The household then consumes its full resources in period 1.

We assume the household solves the following optimization problem:

$$c_0(h, \rho) \equiv \arg\max_{c_0} u(c_0) + \beta \mathbf{E}\left[u(R(w_0 + h - c_0) + y_1)\right] \tag{2}$$

---

[12]We make the technical assumption, as in Carroll & Kimball (1996), that R is not perfectly correlated with y_1. Implicitly, we use R in place of $R(\rho) = \rho + \tilde{R}$, where $\tilde{R}$ is a mean 0 stochastic shock to the return on investment.

We refer to the household's choice of consumption as a function of the size of the UCT it receives and its expected return on investment, $c_0(h, \rho)$, as the household's initial period consumption function. We equivalently define the household's future period consumption function $c_1(h, \rho) \equiv R(w_0 + h - c_0(h, \rho)) + y_1$, which is stochastic through its dependence on R and y_1. As is common practice, we refer to $dc_0(h, \rho)/dh$ as the marginal propensity to consume, and following Auclert et al. (2018) we refer to $dc_t(h, \rho)/dh$ as the period t intertemporal marginal propensity to consume.

In analyzing this optimization problem, we leverage the result from Carroll & Kimball (1996) that households will have a strictly decreasing marginal propensity to consume with a utility function exhibiting hyperbolic absolute risk aversion, a class that nests common functional forms including constant relative risk aversion with a subsistence constraint. That is, $d^2c_0(h, \rho)/dh^2 < 0$. One can draw the intuition underlying this result from the buffer stock model in Deaton (1991), which excludes uncertainty in R. In Deaton (1991), poorer households consume hand-to-mouth (so their marginal propensity to consume is 1), while richer households hold precautionary savings (so their marginal propensity to consume is less than 1).

We next consider the dynamic treatment effects of UCT on consumption in this framework. In Section 3, we observed estimates of the period t impacts of cash transfers on consumption per unit of transfer, which we call $\tau_t(h, \rho)$. In this framework, $\tau_t(h, \rho) \equiv (c_t(h, \rho) - c_t(0, \rho))/h$.

4.1.2 Comparative statics

Next, we generate 3 key predictions on comparative statics of the impacts of UCT on consumption per unit of transfer with respect to transfer size, the time since the transfer, and the return on investment. Formal derivations for all results are in Appendix A.

Proposition 1. *The initial period impacts of UCT on consumption per unit of transfer are decreasing in transfer size.*

That is, $d\tau_0(h, \rho)/dh < 0$. This follows immediately from concavity of the household's initial period consumption function. The economic intuition follows from the household's decreasing marginal propensity to consume – the initial period impacts of UCT on consumption per unit of transfer are equal to the "average propensity to consume" over the transfer, which will decrease in transfer size whenever the marginal propensity to consume is decreasing in transfer size.[13]

Proposition 2. *The future period impacts of UCT on consumption per unit of transfer are increasing in transfer size.*

That is, $d\tau_1(h, \rho)/dh > 0$. To produce this result, we rewrite the household's intertemporal budget constraint in terms of the impacts of UCT on consumption per unit of transfer: $\tau_0(h, \rho) + \frac{1}{R}\tau_1(h, \rho) = 1$. As the return on investment R is not affected by h, when initial period impacts per unit of transfer are decreasing, future period impacts per unit of transfer must be increasing. Alternatively phrased, when household consumption is concave in transfer size, household savings must be convex in transfer size, and this causes the impacts of cash transfers on consumption per unit of transfer to be increasing in transfer size in future periods.[14]

These two predictions formalize the claim that the impacts of larger interventions should be more **persistent** — the impacts of larger interventions, when expressed per unit of transfer, should increase over time relative to the impacts of smaller interventions. This definition suggests a natural test of the greater persistence of the impacts of larger interventions predicted by this model. Specifically, it motivates a regression of the dynamic impacts of UCT on consumption per unit of transfer, $\tau_t(h, \rho)$, on transfer size h, time since transfer t, and their interaction $h \times t$. Our two predictions suggest that the coefficient on transfer size should be negative (as impacts should be decreasing in transfer size in initial periods), but

---

[13]This consequence of a decreasing marginal propensity to consume, along with models that generate it, is discussed in Fagereng et al. (2019).

[14]In an extension with T periods, the same result holds both for the net present value of future consumption and for consumption in the last period.

the coefficient on the interaction should be positive.

However, our predictions hinge on the decreasing marginal propensity to consume implied by the model, and the constant returns to scale investment technology we assumed. We consider two extensions to our model, through alternative investment technologies that are particularly relevant in developing countries, that weaken our predictions – decreasing returns to scale, and non-convexities that generate a poverty trap.[15]

Proposition 2a. *Introducing decreasing returns to scale decreases the effect of transfer size on the future period impacts of UCT on consumption per unit of transfer.*

With decreasing returns to scale, as transfer size grows, households increase their investment, but their marginal return on investment also falls. The fall in the marginal return on investment causes households to reduce their investment, which causes the future period intertemporal marginal propensity to consume to decrease in transfer size. As a result, with decreasing returns to scale, the sign of the effect of transfer size on future period impacts of UCT on consumption per unit of transfer becomes ambiguous. While constant or decreasing future period impacts on consumption per unit of transfer with respect to transfer size (that is, $d\tau_1(h, \rho)/dh \leq 0$) is not consistent with our base model with constant returns to scale, it can be rationalized with decreasing returns to scale.

Proposition 2b. *Introducing non-convexities that generate a poverty trap increases the effect of transfer size on impacts of UCT on initial (future) period consumption per unit of transfer when the density of households at a poverty threshold is decreasing (increasing) in transfer size.*

To introduce a simple form of non-convexities, we consider households deciding whether or not to make a lumpy investment. When the household does not make this investment, it has less resources in the future period, and as a result is less able to make this investment in the future period. This generates a poverty trap, often used to explain persistent poverty in

---

[15]For parsimony, we consider these extensions in a simplified deterministic version of our model.

developing countries (Carter & Barrett, 2006; Banerjee et al., 2019; Balboni et al., 2020). As initial period resources are required to make this investment, each household has a "poverty threshold" of initial resources, above which they make the investment and below which they do not. When transfer size increases, a household just below the poverty threshold will now make the lumpy investment, and as a result discontinuously *decrease* their initial period consumption (while discontinuously increasing their future period consumption). As a result, larger transfers will have relatively larger impacts on consumption per unit of transfer in future periods when the density of households at the poverty threshold is increasing in transfer size, and relatively smaller impacts when the density is decreasing in transfer size. Therefore, alternative densities of distances of households from the poverty threshold can rationalize any pattern of effects of increasing transfer size on impacts of UCT on consumption, including a decreasing impact of UCT on future consumption per unit of transfer with respect to transfer size.

Lastly, we consider the role of the return on investment in determining the impacts of transfers on consumption.[16] A large body of research has found high marginal returns to investment in developing countries relative to richer countries (De Mel et al., 2008; McKenzie & Woodruff, 2008; Hussam et al., 2020). Our results from Section 3 are focused on developing countries, in contrast to existing work estimating intertemporal marginal propensities to consume in richer countries. Differences between our results and existing work are therefore likely to be partially driven by differences in the marginal return on investment between developing countries and richer countries.

Proposition 3. *The future period impacts of UCT on consumption per unit of transfer are increasing in the return on investment.*

That is, $d\tau_1(h, \rho)/d\rho > 0$. To provide the intuition underlying this result, we can revisit the household's intertemporal budget constraint written in terms of impacts of UCT on consumption per unit of transfer. Rewriting with future impacts as the numeraire, and

_____

[16]For parsimony, we consider this comparative static in a simplified deterministic version of our model.

abstracting from uncertainty in the return on investment (i.e., assuming $R = \rho$), $\rho\tau_0(h, \rho) + \tau_1(h, \rho) = \rho$. We can then interpret a change in the return on investment ρ as having two effects – a price effect, and an income effect.[17] The price effect causes future consumption to be cheaper, so the impact on future consumption increases while the impact on present consumption decreases. The income effect increases the net present value of the transfer in units of future consumption, so the impacts on both present and future consumption increase.

We summarize these predictions in Table 3. Our base model in Equation 2 generates the prediction that initial period impacts per unit of transfer should decrease in transfer size (Proposition 1), while future period impacts per unit of transfer should increase in transfer size (Proposition 2). Introducing decreasing returns to scale weakens the prediction that future period impacts should increase in transfer size (Proposition 2a), while introducing non-convexities that generate poverty traps weakens both predictions (Proposition 2b). While any impacts are potentially consistent with non-convexities that generate poverty traps, other models provide testable predictions. Lastly, increases in the return on investment should increase future period impacts per unit of transfer (Proposition 3).

Table 3: Summary of predictions on impacts on consumption per unit of transfer τ_t

	Transfer size h		Return on investment ρ	
	Period 0 $\dfrac{d\tau_0(h,\rho)}{dh}$	Period 1 - Period 0 $\dfrac{d^2\tau_t(h,\rho)}{dhdt}$	Period 0 $\dfrac{d\tau_0(h,\rho)}{d\rho}$	Period 1 - Period 0 $\dfrac{d^2\tau_t(h,\rho)}{d\rho dt}$
Base model	$-$	$+$	?	$+$
with DRTS	$-$	?		
with Poverty traps	?	?		

4.2 Descriptive evidence

To begin this analysis, we plot posterior estimates of the impacts of small UCT, large UCT, and TUP on consumption per unit of transfer from Section 3 against years since last transfer in Figure 2. The area of each point in the figure is proportional to the inverse posterior variance of the associated estimate. We see that there is meaningful variation in relatively high weight estimates in time since last transfer for both large and small UCT up to 2.5 years since last transfer; this variation is the source of our identification of the effect of log transfer size and time since last transfer on the impacts of UCT on consumption.

Figure 2: Variation in posterior estimates with respect to transfer size and years since last transfer
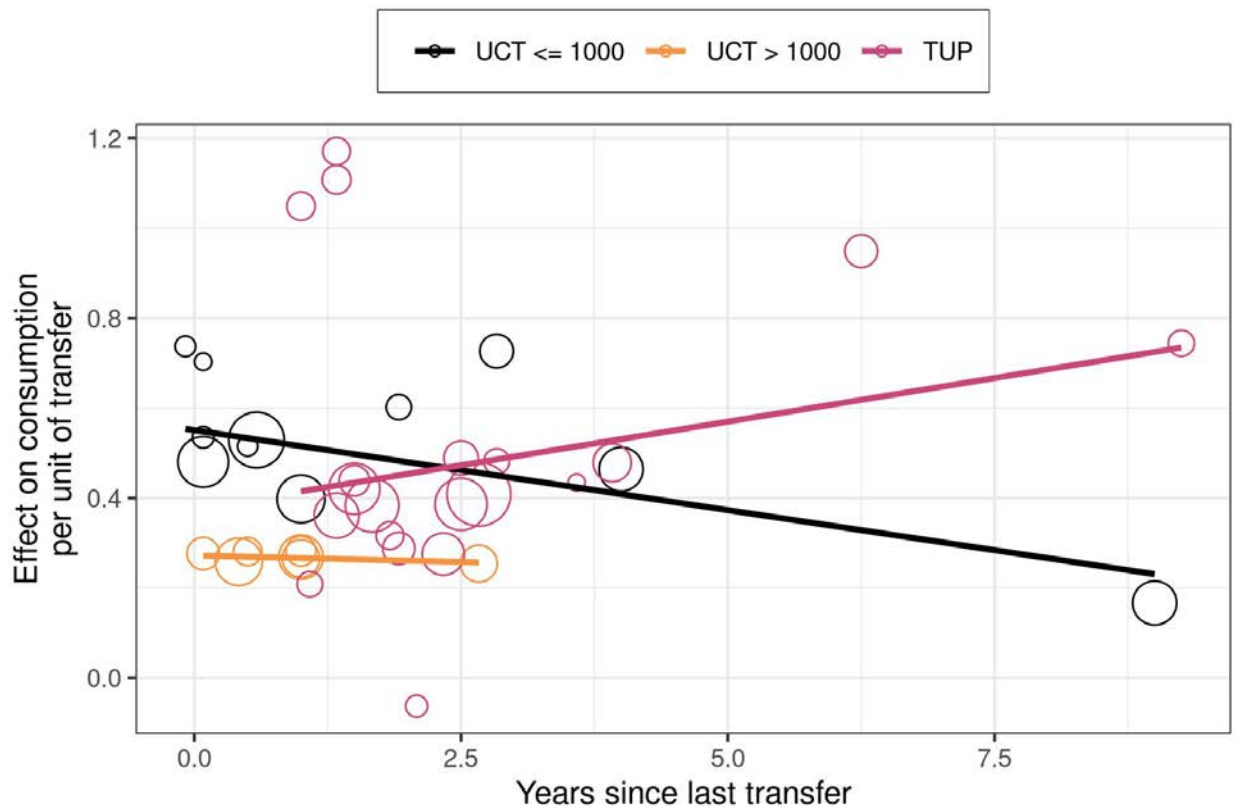
Notes: This figure presents Bayesian posterior means for each estimate of the impact of an intervention on consumption per unit of transfer, and plots these means against years since last transfer. Areas of circles corresponding to each estimate are proportional, within each group of interventions, to inverse posterior variances. Estimates from UCT interventions with transfer size smaller than 1000 USD PPP are plotted in black, estimates from UCT interventions with transfer size larger than 1000 USD PPP are plotted in orange, and estimates from TUP interventions are plotted in pink.

We make four observations. First, we show in Figure 2 that the effects of UCT on

consumption per unit of transfer appear relatively persistent over time, with some evidence of small decreases over time. Second, there is limited evidence that the effect of smaller UCT on consumption per unit of transfer decreases over time relative to the effect of larger UCT. In Section 4.1, we showed that larger UCT should have larger impacts on consumption per unit of transfer at longer time horizons, which we do not observe in the data. Third, we note that these estimates may be relatively short term, as we only have common support for small and large UCT out to two and a half years. However, in Section 5, we show this time frame is sufficient to see differences between the persistence of our estimates and estimates from richer countries. Fourth, smaller UCT appear to have larger effects on consumption per unit of transfer than larger UCT. We formalize this analysis in Section 4.5, building on this descriptive evidence and leveraging variation both within and across RCTs, and discuss our results.

4.3 Empirical strategy

To test how the impacts of cash transfers on consumption vary with transfer size and time since last transfer, we estimate Equation 1 but with the introduction of controls. Specifically, we estimate the following model using the Bayesian approach described in Section 3.1.

$$\hat{\tau}_{rat}|\tau_{rat} \sim N(\tau_{rat}, \hat{se}_{rat}^2)$$
$$\tau_{rat} \sim N(X'_{rat}\beta, \sigma_\tau^2)$$

$$(3)$$

We consider two primary specifications of the observable characteristics of each estimate X_{rat}, both of which include a constant. First, we use years since last transfer, to test whether impacts of UCT vary over time, and log transfer size, to test whether larger UCT have larger impacts per unit of transfer. Second, we use years since last transfer, log transfer size, and their interaction. As discussed in Section 4.1, in this second specification, buffer stock models generate the prediction that the coefficient on log transfer size should be negative while the coefficient on the interaction should be positive. For robustness, we also consider

28

specifications where we include additional control variables, including RCT fixed effects. The inclusion of RCT fixed effects isolates within RCT variation in years since last transfer (from multiple survey waves) and in log transfer size (from variation in transfer size across arms).

In estimating β in Equation 3, our parameter of interest is a weighted average causal effect of shifting transfer size at different time horizons. We note that this interpretation of β rests on an exogeneity assumption; we use balance tests to test this assumption in Section 4.4. However, we also note that this is distinct from the parameter estimated by any individual estimate of the differential effect of larger cash transfers, which estimates the effect of increasing transfer size in a single context. Estimation of the average effect requires a large number of studies, and this is reflected in reduced precision of our estimates in Section 4.5.

4.4 Balance

In Equation 3, which leverages variation in years since last transfer and log transfer size across studies, one might be concerned that years since last transfer and log transfer size are correlated with other characteristics of the RCT or the intervention. However, we note that if most of the variation in years since last transfer and in log transfer size is within RCT, we should expect to find balance with respect to these variables.

To test balance, we estimate by OLS

$$Y_{rat} = X'_{rat}\eta + \epsilon_{rat} \tag{4}$$

for important characteristics of RCTs or interventions Y_{rat}, with the same specification of the right hand side as in Equation 3. We report the results of these balance tests in Table 4, with robust standard errors clustered at the RCT-level.

Across our two specifications, we fail to reject the null of balance for 14 of 18 coefficients. We now consider the 4 rejections, and note that 2 of them are relatively mechanical. First,

Table 4: Balance: Transfer size and persistence

	log baseline consumption	log GDP per capita	Africa	Transfer duration (years)	Year of last transfer	Intervention cost per unit of transfer
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: No interaction						
Years since last transfer	0.093	0.010	0.014	-0.084	-1.03	-0.003
	(0.051)	(0.038)	(0.016)	(0.025)	(0.19)	(0.008)
	[0.089]	[0.795]	[0.403]	[0.004]	[0.000]	[0.659]
log transfer size	0.153	-0.141	0.045	-0.290	0.60	-0.074
	(0.168)	(0.178)	(0.059)	(0.161)	(1.07)	(0.030)
	[0.377]	[0.442]	[0.456]	[0.093]	[0.583]	[0.025]
Panel B: Interaction						
Years since last transfer *	-0.126	-0.012	-0.028	-0.126	-0.98	-0.008
log transfer size	(0.131)	(0.130)	(0.038)	(0.269)	(0.985)	(0.036)
	[0.357]	[0.928]	[0.470]	[0.646]	[0.336]	[0.822]
Controls						
Years since last transfer	X	X	X	X	X	X
log transfer size	X	X	X	X	X	X
Mean dep. var.	7.358	7.644	0.944	0.676	2014.83	1.179
# of RCTs	6	7	7	7	7	7
# of observations	16	18	18	18	18	18

Notes: Columns 1 through 6 present regression coefficients, with robust standard errors clustered at the RCT-level in parentheses and p-values in brackets.

a one year increase in years since last transfer is associated with a 1.05 year decrease in the year of the intervention. This is because longer run estimates (with larger years since last transfer) must come from interventions that occurred more years ago. Second, a 10% increase in transfer size is associated with a 0.009 decrease in cost per unit of transfer. This is because many of the costs of UCT are fixed with respect to transfer size, such as targeting and delivery, so larger transfers are cheaper to deliver per unit of transfer. Third, a one year increase in years since last transfer is associated with approximately one month shorter duration of transfers. This correlation is potentially because UCT with longer duration often have the objective of increasing consumption at shorter time horizons, and so often include relatively short run follow up surveys. For robustness, we therefore also estimate specifications that include transfer duration as a control. Fourth, a one year increase in

years since last transfer is associated with a 9% increase in baseline household consumption; this magnitude is relatively small and unlikely to bias our estimates.

4.5 Results

Table 5: Transfer size and persistence

	(1)	(2)	(3)	(4)	(5)	(6)
			Effect on consumption per unit of transfer			
β_1: Years since last transfer	-0.029	0.191	-0.024	0.235	-0.036	-0.003
	(0.017)	(0.395)	(0.019)	(0.425)	(0.038)	(0.599)
	{-0.062, 0.004}	{-0.502, 1.063}	{-0.061, 0.014}	{-0.518, 1.179}	{-0.114, 0.036}	{-1.106, 1.280}
β_2: log transfer size	-0.151	-0.128	-0.136	-0.107	-0.246	-0.241
	(0.051)	(0.070)	(0.058)	(0.080)	(0.093)	(0.142)
	{-0.254, -0.052}	{-0.263, 0.015}	{-0.252, -0.022}	{-0.259, 0.061}	{-0.433, -0.065}	{-0.514, 0.057}
β_3: Years since last transfer * log transfer size		-0.032		-0.038		-0.005
		(0.058)		(0.062)		(0.085)
		{-0.159, 0.069}		{-0.176, 0.072}		{-0.190, 0.151}
σ_τ: Std. dev. of effects	0.053	0.069	0.061	0.080	0.076	0.098
	(0.047)	(0.058)	(0.054)	(0.066)	(0.065)	(0.082)
	{0.003, 0.175}	{0.003, 0.215}	{0.003, 0.201}	{0.006, 0.246}	{0.003, 0.242}	{0.005, 0.307}
Controls						
Transfer duration			X	X		
RCT fixed effects					X	X
# of RCTs	7	7	7	7	5	5
# of observations	18	18	18	18	16	16

Notes: Estimates of the model of impacts of UCT interventions on household consumption per unit of transfer in Equation 3 are presented in this table. Columns 1 through 6 report Bayesian posterior means of each parameter, with posterior standard deviations in parentheses and 95% credible intervals in curly brackets.

Bayesian estimates of Equation 3 are presented in Table 5. In Columns 1 and 2, we do not include any additional controls; Columns 3 and 4 include transfer duration as a control, while Columns 5 and 6 include RCT fixed effects.

First, in Column 1, we regress the effect of cash transfers on consumption per unit of transfer on years since last transfer and log transfer size. This corresponds closely with the descriptive analysis presented in Figure 2. The point estimate on years since last transfer implies strong persistence of the effects of cash transfers on consumption: a one year increase in years since last transfer corresponds with a 0.029 decrease in the effect of cash transfers on consumption per unit of transfer, or a 8% annual decrease relative to the average effect. This estimate is precise, as the 95% credible interval excludes fade out of effects any faster than 18% of the average effect per year. The point estimate on log transfer size implies a doubling

of transfer size causes a 0.105 decrease in effects on consumption per unit of transfer, 30%
of the average effect. While this magnitude is large, our theory in Section 4.1 suggests that
this could be a consequence of our average estimates capturing relatively short run impacts
of UCT on consumption, as larger UCT are predicted to have relatively larger impacts per
unit of transfer at longer time horizons.[18]

Second, in Column 2, we regress the effect of cash transfers on consumption per unit of
transfer on years since last transfer, log transfer size, and their interaction. Our base model
now makes two predictions on these coefficients. First, the coefficient on log transfer size
should be negative, as when years since last transfer is 0, increasing transfer size should
decrease the effect of cash transfers on consumption per unit of transfer. Second, the coeffi-
cient on the interaction of years since last transfer and log transfer size should be positive,
because the effects of larger cash transfers on consumption per unit of transfer should grow
relative to the effects of smaller cash transfers as time since last transfer increases. Instead,
while we continue to find a negative coefficient on log transfer size, we also find a negative
coefficient on the interaction, inconsistent with our base model.[19]

We note that the 95% credible intervals for the coefficients on both log transfer size and
its interaction with years since last transfer do not exclude 0 in the interacted model. To
clarify our statistical precision, we plot posterior mean predicted impacts on the effect on
consumption per unit of transfer of log transfer size as a function of years since last transfer
along with 95% credible intervals in Figure 3. When years since last transfer is 0, this
corresponds to the coefficient on log transfer size. In our base specification, our 95% credible
intervals for the effect of increasing transfer size at any time from 0.2 to 2.5 years since last

---

[18]While our analysis in this section focuses on effects on consumption per unit of transfer to link closely
to our theory, it is instead the effects on consumption per unit of cost that may be most relevant concerning
optimal policy. In addition, as noted in Section 4.4, larger transfers have lower costs per unit of transfer
because of fixed overhead costs. We therefore replicate our results in Table 5 using effect on consumption per
unit of cost as our outcome of interest in Appendix B.2. While coefficients on log transfer size shrink, our
results and precision remain qualitatively similar, as even the smallest transfers in our sample have relatively
low overhead.

[19]In Appendix B.1, we show that our results on the effect of increasing log transfer size, and its interaction
with years since last transfer, are robust to the exclusion of any individual RCT in our sample.

transfer exclude any positive effect; we note that only 4 of our 18 point estimates for UCT are more than 2.5 years since last transfer. This highlights that relatively larger impacts of larger UCT would need to manifest at relatively long time horizons for our results to remain consistent with our base model.

Figure 3: Transfer size and persistence
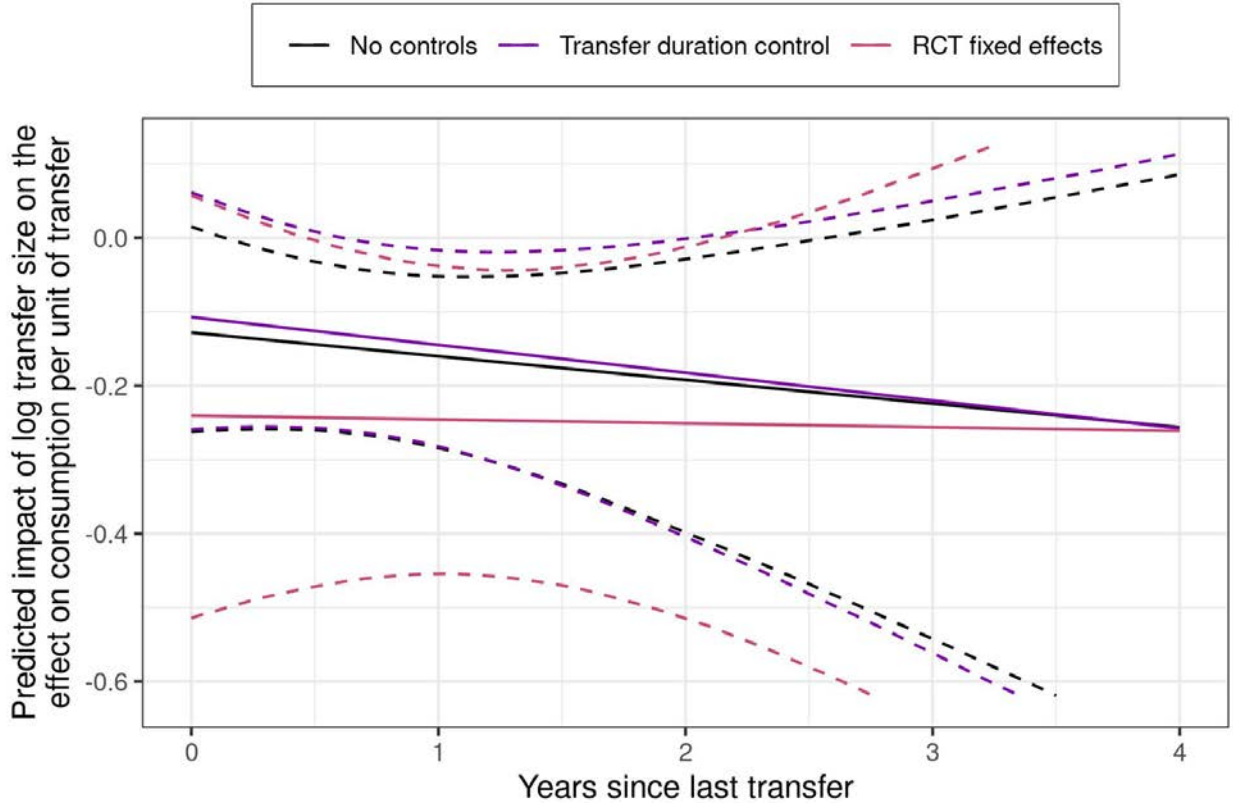
Notes: This figure presents Bayesian posterior means for predicted impacts of log transfer size on effects on consumption per unit of transfer from Table 5, and plots these predicted impacts as a function of years since last transfer. Solid lines are posterior means, while dotted lines are 95% credible intervals. Estimates from Column 2 ("No controls") are plotted in black, estimates from Column 4 ("Transfer duration control") are plotted in purple, and estimates from Column 6 ("RCT fixed effects") are plotted in pink.

Comparing our result — increasing transfer size decreases impacts on consumption per unit of transfer at both short and medium time horizons — to our predictions in Table 3 suggests our results are potentially consistent with either a decreasing marginal return on investment, or non-convexities in the investment technology that generate a poverty trap. That the results are consistent with a decreasing marginal return on investment is intuitive — decreasing returns to scale in the household's investment technology might reasonably

33

generate decreasing returns to scale in consumption impacts with respect to transfer size. That the results are consistent with poverty traps may be less expected, as poverty traps are commonly used to motivate larger interventions. However, as noted in Section 4.1, introducing a simple form of a poverty trap yields the prediction that medium run impacts of UCT on consumption per unit of transfer will decrease with respect to transfer size when the density of the distance of households to the poverty threshold is decreasing, consistent with our point estimate. Adding additional complexity to our model, such as multiple poverty thresholds or heterogeneous technologies, could further weaken this prediction. In all cases, the fundamental conclusion is that the presence of poverty traps may either increase or decrease the impact of transfer size on cost effectiveness, and therefore the relationship between intervention size and consumption impacts remains an empirical question even given knowledge of the presence of poverty traps. Our estimates provide an answer to this question on average across a range of developing countries.

Lastly, in Columns 3 through 6, we test the robustness of these results to the inclusion of either transfer duration or RCT fixed effects as controls. The qualitative patterns we describe above are in general unchanged. Controlling for transfer duration does not meaningfully affect our estimates, while controlling for RCT fixed effects increases magnitudes of point estimates but also increases the width of credible intervals.

We interpret our estimates in Table 5 as reflecting average impacts of increasing transfer size across contexts and interventions, and therefore these estimates do not speak to the possibility that increasing transfer size may have heterogeneous effects across contexts. If we had a sufficient number of RCTs where transfer size was randomly assigned and multiple waves of post-baseline consumption were collected, we could also estimate heterogeneity in dynamics and the effect of increasing transfer size on dynamics. Our sample contains 2 RCTs which both randomly vary transfer size and collect multiples waves of post-baseline consumption; additional such RCTs would allow estimation of heterogeneity in the relative impacts of larger cash transfers at different time horizons.

Lastly, we note that the credible intervals on many of these estimates are economically wide. As discussed in Abadie et al. (2020), these intervals reflect sampling error across RCTs: in other words, they are credible intervals for the conditional average effect of a randomly sampled RCT from the population of potential RCTs. However, we note that we would gain significant precision if we instead provided credible intervals on the conditional average effect of a randomly sampled RCT from our sample of RCTs. The former parameter is subject to both sampling and measurement error, while the latter parameter is only subject to measurement error. When confidence intervals are provided for estimates from a single RCT, these confidence intervals only account for measurement error. However, the credible intervals we report are for the parameter of interest: the conditional average effect of a random RCT.

5 Persistence relative to richer countries

We now turn to the relative persistence of UCT in developing countries and richer economies. In Section 4.1, we predicted that cash transfers should have larger effects at longer time horizons in developing countries relative to richer economies, because marginal returns to investment are higher in developing countries. In Section 5.1, we describe our empirical strategy to compare our estimates in developing countries to estimates from Auclert et al. (2018) and Fagereng et al. (2019) in richer countries, and implement this comparison in Section 5.2.

5.1 Empirical strategy

We compare our results to estimates of the impacts of UCT per unit of transfer on consumption over time from richer countries. We primarily use estimates from Fagereng et al. (2019), who estimate impacts on annual consumption up to 5 years since last transfer using lottery winnings in Norway. Auclert et al. (2018) notes that these estimates can be interpreted

as intertemporal marginal propensities to consume, and that these are key parameters for disciplining models of intertemporal optimization.

To facilitate comparisons, we estimate Equation 3, using indicators for years since last transfer being less than 1, between 1 and 2, and greater than 2 as our observable characteristics of estimates. The coefficients on these variables estimate weighted averages of estimated effects of cash transfers on consumption per unit of transfer that are less than one year, between one and two years, and greater than two years since last transfer, respectively. We then compare these estimates to the estimated impacts of cash transfers on consumption per unit of transfer at less than 1 year, between 1 and 2 years, and between 2 and 3 years from richer countries, respectively. We note that some of the estimates that contribute to estimating the impacts of UCT greater than 2 years since last transfer are more than 3 years after last transfer; this yields a conservative estimate of the impacts of cash transfers at longer time horizons in developing countries if the impacts of cash transfers on consumption are decreasing over time.

As a last step, we use these estimates to implement a simple cost-benefit analysis of cash transfers following the methodology in Banerjee et al. (2015). Banerjee et al. (2015) conduct a cost-benefit analysis of multifaceted graduation programs by comparing the net present value of their impacts on consumption to their cost. We similarly calculate the net present value of three year impacts of cash transfers on consumption per unit of transfer as $\sum_{t=1}^{3} \delta^t \beta_t$. We use $\delta = 1$ for ease of interpretation, as modestly smaller discount factors will not meaningfully affect our point estimate over just 3 years.

5.2 Results

In Table 6, we present our estimates of the impacts of cash transfers on consumption per unit of transfer at one, two, and three or more years, with estimates from Fagereng et al. (2019) for comparison. First, all of our estimates are close to our estimates of the average effect of cash transfers from Section 3, consistent with our finding that impacts of UCT are

persistent over time in Section 4. Second, our 95% credible intervals exclude estimates from Fagereng et al. (2019) at both year one and year three — our year one estimate in developing countries is smaller than in Norway, while our year three estimate is larger. As discussed in Section 4.1, this larger persistence of impacts is consistent with a larger marginal return on investment in developing countries than in Norway.

Table 6: Intertemporal marginal propensities to consume

	Effect on consumption per unit of transfer	Fagereng et al. (2020)
	(1)	(2)
β_1: Years since last transfer ≤ 1	0.358 (0.065) { 0.250, 0.505}	0.52
β_2: $1 <$ Years since last transfer ≤ 2	0.768 (0.416) {-0.050, 1.585}	0.17
β_3: $2 <$ Years since last transfer	0.316 (0.103) { 0.132, 0.541}	0.10
σ_τ: Std. dev. of effects	0.129 (0.074) {0.014, 0.300}	
$\beta_1 + \beta_2 + \beta_3$	1.442 (0.435) {0.596, 2.304}	
# of RCTs	7	
# of observations	18	

Notes: Estimates of the model of impacts of UCT interventions on household consumption per unit of transfer in Equation 3 are presented in Column 1 of this table. Column 1 reports Bayesian posterior means of each parameter, with posterior standard deviations in parentheses and 95% credible intervals in curly brackets. Column 2 reports comparable point estimates from Fagereng et al. (2019), as reported by Auclert et al. (2018).

Lastly, we use these results to conduct a cost-benefit analysis of UCT over the first three years. We estimate the net present value of the consumption impacts of UCT over the first three years is 1.44 per unit of transfer, statistically indistinguishable from the average cost of UCT in our sample. We interpret this result as evidence that UCT pass a cost-benefit analysis following Banerjee et al. (2015), even without any assumptions on the persistence

of effects beyond three years.

6 Multifaceted graduation programs ("Scope")

6.1 Descriptive evidence

We begin by making some simple comparisons between UCT and TUP estimates based on our initial metaänalysis results in Table 2 and Figure 1, descriptive statistics of interventions in Table 1, and comparisons of UCT and TUP estimates over time in Figure 2. First, as seen in Table 2, TUP interventions on average increase consumption by more per unit of transfer than UCT interventions. Our preferred estimate is 0.52 per unit of transfer for TUP, compared to 0.35 per unit of transfer for UCT. This result is consistent with findings of positive impacts on consumption of add-on interventions in TUP programs (Banerjee et al., 2018; Bossuroy et al., 2021). Second, these results are likely to reverse per unit of cost. While TUP interventions increase consumption on average by 50% more per unit of transfer than UCT interventions, Table 1 highlights that they cost on average more than three times as much per unit of transfer. Third, these averages are likely to mask important heterogeneity. While the coefficient of variation of impacts of TUP on consumption of 0.77 is small relative to almost all interventions in Vivalt (2020), this is sufficiently large to suggest some TUP programs may have much larger impacts than the average UCT intervention per unit of transfer, while others may have no impact at all. This heterogeneity is visible in Figure 1, where posterior estimates for TUP estimates are substantially more dispersed than for UCT — the largest posterior mean estimate of TUP impacts on consumption per unit of transfer is more than three times as large as for the average UCT, while the smallest posterior mean estimate is negative. Lastly, in Figure 2, the impacts of TUP and small UCT appear comparable per unit of transfer up to 2.5 years singe last transfer. However, TUP estimates at longer time horizons are larger than at short time horizons, suggesting the effects of TUP may grow over time relative to UCT. In Section 6.4, we implement these

tests using specifications defined in Section 6.2.

6.2 Empirical strategy

To test the differences in impacts of UCT and TUP on household consumption, we now estimate Equation 3, but with the inclusion of an indicator that a particular estimate is of the impacts of a TUP program. Specifically, we estimate

$$
\begin{aligned}
\hat{\tau}_{rat}|\tau_{rat} &\sim N(\tau_{rat}, \hat{\text{se}}^2_{rat}) \\
\tau_{rat} &\sim N(\beta \text{TUP}_{ra} + X'_{rat}\gamma, \sigma^2_\tau)
\end{aligned}
\tag{5}
$$

where X_{rat} includes a constant and, in some specifications, other important characteristics, and the coefficient β recovers the difference between UCT and TUP.

In estimating β in Equation 5, our parameter of interest is a weighted average causal effect of shifting a given program from UCT to TUP by adding complementary interventions. We note that this interpretation of β rests on an exogeneity assumption; we use balance tests to test this assumption in Section 6.3. However, we also note that this is distinct from the parameter estimated by any individual study comparing UCT to TUP, which estimates the effect of shifting a single program from UCT to TUP; in light of the heterogeneity we estimate in Table 2, these effects are likely to vary meaningfully across contexts. Estimation of this weighted average effect requires a large number of studies, and this is reflected in the precision of our estimates in Section 6.4. We also note that with sufficient RCTs comparing the impacts of UCT to TUP, we could also estimate the variance across contexts of the effect of shifting from UCT to TUP.

Lastly, we test for heterogeneity in differences between UCT and TUP with respect to two important characteristics. First, it is often proposed that non-cash components of TUP programs enable more persistent impacts on household consumption (Banerjee et al., 2015). Given this, we would expect that the impacts of TUP are relatively larger over longer time horizons. In some specifications, we therefore interact the TUP indicator with years since

last transfer, to test the null hypothesis that there is no difference in the changes in the impacts of TUP and UCT over time. Second, the standard deviation of cost per unit of transfer of TUP programs is roughly 20 times larger than for UCT programs, suggesting cost may be an important determinant of differences between TUP and UCT in impacts on consumption per unit of cost. In some specifications, we therefore interact the TUP indicator with an indicator that the TUP intervention had costs per unit of transfer greater than 4, roughly the average in our sample.

6.3 Balance

One possible concern when estimating Equation 5 is that differences in impacts of UCT and TUP per unit of transfer are caused by differences in characteristics between the UCT and TUP interventions we study that are not intrinsic to the distinction between UCT and TUP, such as beneficiary characteristics or the size of transfers. To alleviate this concern, we test for balance of important characteristics between UCT and TUP programs, and include a specification that controls for important characteristics for which we fail to find balance. To test balance, we estimate by OLS

$$Y_{rat} = \mu \text{TUP}_{rat} + X'_{rat}\eta + \epsilon_{rat} \tag{6}$$

This is similar to our balance tests in Section 4.4, but with the same specification of the right hand side as in Equation 5. We report the results of these balance tests in Table 7, with robust standard errors clustered at the RCT-level.

In an initial specification that does not include any controls, balance across UCT and TUP interventions is quite poor. TUP interventions target households with 69% higher consumption and in higher income countries, they are 64pp less likely to be in Africa, transfers last 9 months longer, and the programs took place 4 years earlier. However, TUP and UCT estimates are not significantly different in years since last transfer, and transfer sizes are

Table 7: Balance: Multifaceted graduation programs

	log baseline consumption	log GDP per capita	Africa	Transfer duration (years)	Year of last transfer	Years since last transfer	log transfer size
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: No controls							
TUP	0.692	0.421	-0.644	0.774	-3.88	1.139	0.179
	(0.253)	(0.271)	(0.167)	(0.335)	(1.57)	(0.851)	(0.268)
	[0.010]	[0.128]	[0.000]	[0.027]	[0.018]	[0.189]	[0.509]
Panel B: Controls							
TUP	0.498	-0.301			-4.31	1.560	0.158
	(0.262)	(0.343)			(1.14)	(0.993)	(0.455)
	[0.067]	[0.387]			[0.001]	[0.125]	[0.731]
Controls							
Africa	X	X			X	X	X
Transfer duration	X	X			X	X	X
Mean dep. var.	7.743	7.866	0.605	1.083	2012.79	2.081	6.732
# of RCTs	16	17	17	17	17	17	17
# of observations	36	38	38	38	38	38	38

Notes: Columns 1 through 6 present regression coefficients, with robust standard errors clustered at the RCT-level in parentheses and p-values in brackets.

similar.

We therefore also consider a specification that includes controls for an Africa indicator and transfer duration. Conditional on these controls, TUP interventions continue to target households that have significantly higher consumption, but the TUP coefficient decreases by one third. However, the TUP coefficient for year of last transfer remains at 4 years earlier. These differences highlight the challenges in comparing across interventions, and we interpret differences in results with and without these controls included as a test of the robustness of our results.

6.4 Results

Estimates of Equation 5 are presented in Table 8. In Panel A, we estimate the impacts of TUP programs on consumption per unit of transfer relative to UCT programs. In Panel B, we estimate the impacts of TUP programs on consumption per unit of cost relative to

UCT programs. Columns 1 and 3 estimate differences between UCT and TUP programs, Columns 2 and 4 allow for heterogeneity with respect to years since last transfer, while Column 5 allows for heterogeneity with respect to TUP costs.

Table 8: Multifaceted graduation programs and persistence

	(1)	(2)	(3)	(4)	(5)
Panel A: Effect on consumption per unit of transfer					
β_1: TUP	0.111	-0.056	0.464	0.324	0.099
	(0.114)	(0.170)	(0.170)	(0.193)	(0.129)
	{-0.122, 0.329}	{-0.401, 0.273}	{ 0.116, 0.789}	{-0.070, 0.693}	{-0.160, 0.350}
β_2: TUP *		0.078		0.090	
Years since last transfer		(0.056)		(0.050)	
		{-0.032, 0.190}		{-0.007, 0.190}	
β_3: TUP * (Cost per					0.031
unit of transfer ≥ 4)					(0.165)
					{-0.302, 0.349}
σ_τ: Std. dev. of effects	0.277	0.280	0.235	0.223	0.283
	(0.060)	(0.064)	(0.057)	(0.059)	(0.063)
	{0.171, 0.409}	{0.168, 0.419}	{0.136, 0.362}	{0.123, 0.354}	{0.173, 0.419}
Panel B: Effect on consumption per unit of cost					
β_1: TUP	-0.150	-0.317	-0.018	-0.130	-0.089
	(0.054)	(0.077)	(0.062)	(0.068)	(0.059)
	{-0.258, -0.045}	{-0.473, -0.170}	{-0.152, 0.094}	{-0.279, -0.008}	{-0.205, 0.028}
β_2: TUP *		0.082		0.076	
Years since last transfer		(0.028)		(0.022)	
		{ 0.027, 0.138}		{ 0.033, 0.122}	
β_3: TUP * (Cost per					-0.130
unit of transfer ≥ 4)					(0.062)
					{-0.258, -0.013}
σ_τ: Std. dev. of effects	0.123	0.114	0.070	0.056	0.116
	(0.027)	(0.024)	(0.030)	(0.025)	(0.026)
	{0.077, 0.183}	{0.075, 0.166}	{0.015, 0.135}	{0.010, 0.110}	{0.073, 0.173}
Controls					
Years since last transfer		X		X	
Africa			X	X	
Transfer duration			X	X	
# of RCTs	17	17	17	17	17
# of observations	38	38	38	38	38

Notes: Estimates of the model of impacts of UCT and TUP interventions on household consumption per unit of transfer or per unit of cost in Equation 5 are presented in this table. Columns 1 through 5 report Bayesian posterior means of each parameter, with posterior standard deviations in parentheses and 95% credible intervals in curly brackets.

First, in Column 1 of Panel A, we compare the effects of TUP and UCT on consumption per unit of transfer across our sample. Consistent with our estimates in Table 2, TUP have a 0.11 larger impact on consumption per unit of transfer than UCT. However, 95% credible

interval for the difference includes zero. In Column 3 of Panel A, with controls included, the estimate increases to 0.46, and the 95% credible interval now excludes zero. We interpret these results as evidence that TUP interventions generate larger increases in consumption per unit of transfer than UCT, suggesting that the impact of TUP on consumption cannot be fully explained by the value of the included cash and asset transfers.

Second, in Columns 1 and 3 of Panel B, we compare the effects of TUP and UCT on consumption per unit of cost. With and without controls, TUP programs decrease consumption by 0.02 and 0.15 per unit of cost, respectively, relative to UCT, with the latter 95% credible interval excluding 0. In contrast to our results in the previous paragraph, these results suggest complementary TUP interventions increase consumption less cost effectively than basic transfers.

Third, in Columns 2 and 4, we interact the TUP indicator with years since last transfer, to estimate how the effects of TUP change over time. Posterior mean estimates imply effects of TUP on consumption per unit of transfer grow by 0.08 to 0.09 per year relative to UCT, while effects on consumption per unit of cost grow by 0.08 per year relative to UCT. These magnitudes are large — as reference, the 95% credible interval in our base specification in Table 5 for the annual change in impacts on consumption per unit of transfer for a 10% increase in UCT size was from -0.016 to 0.007. To interpret the implication of these magnitudes for cost effectiveness, we ask how many years it would take the cumulative effect of TUP on consumption per unit of cost to catch up with UCT. Despite these large magnitudes, our estimates with and without controls imply TUP programs take 3.4 years and 7.7 years, respectively, to catch up to the cumulative impacts of UCT on consumption per unit of transfer.[20] In contrast, the average TUP estimate in our sample is 2.6 years since last transfer, highlighting the importance of additional long run estimates of the impacts of

---

[20]Alternatively, in Appendix C, we use these results to calculate the internal rate of return on shifting the average UCT to TUP as a function of the number of years we are willing to extrapolate our estimates to, following Banerjee et al. (2015). Specifications with and without controls imply a 5 year IRR of -57% and 50%. In contrast, specifications with and without controls imply a 10-year IRR of 15% and 78%, although these estimates extrapolate beyond the support of the majority of our sample.

TUP on consumption in order to inform these comparisons.[21]

Fourth, in Column 5, we allow for heterogeneity in the impacts of high and low cost TUP interventions. Our point estimates imply high and low cost TUP interventions have approximately identical impacts on consumption per unit of transfer, while high cost TUP interventions have much lower impacts on consumption per unit of cost. Restricting to low cost TUP interventions, the gap in impacts on consumption per unit of cost between UCT and TUP decreases by 41%. To visualize these relative magnitudes, we plot our estimates from Column 5 in Figure 4. This Figure highlights that the reversal in sign of the relative effect of TUP on impacts on consumption per unit of transfer and per unit of cost that we observed in Columns 1 and 3 are driven by the much higher costs per unit of transfer of TUP relative to UCT, and that differences in costs of implementation across TUP programs are just as important as differences in costs between UCT and TUP.

Finally, we note that consumption is one of many outcomes a social planner may place weight on when deciding whether to implement a given intervention instead of UCT. As discussed in Section 1, we focus on consumption because of the tight link between theory and the impacts of UCT, and because consumption is a common metric for success for many development programs, including TUP. However, this does not disqualify extending this exercise to benchmark the cost effectiveness of TUP in shifting other outcomes of interest, including weighted averages of consumption and other outcomes. Our benchmarking exercise for consumption impacts of TUP against UCT could then be interpreted as estimating the effective "price", in units of consumption impacts, of outcomes that TUP increases more cost effectively than UCT.

---

[21]In Appendix B.1, we show that our results on the effect of TUP and its interaction with years since last transfer, relative to TUP, are robust to the exclusion of any individual RCT in our sample, with the exception of Banerjee et al. (2020). In particular, the coefficient on the interaction between TUP and years since last transfer shrinks to 0 when Banerjee et al. (2020) is excluded, further highlighting the importance of additional long run estimates.

Figure 4: Multifaceted graduation programs and program costs
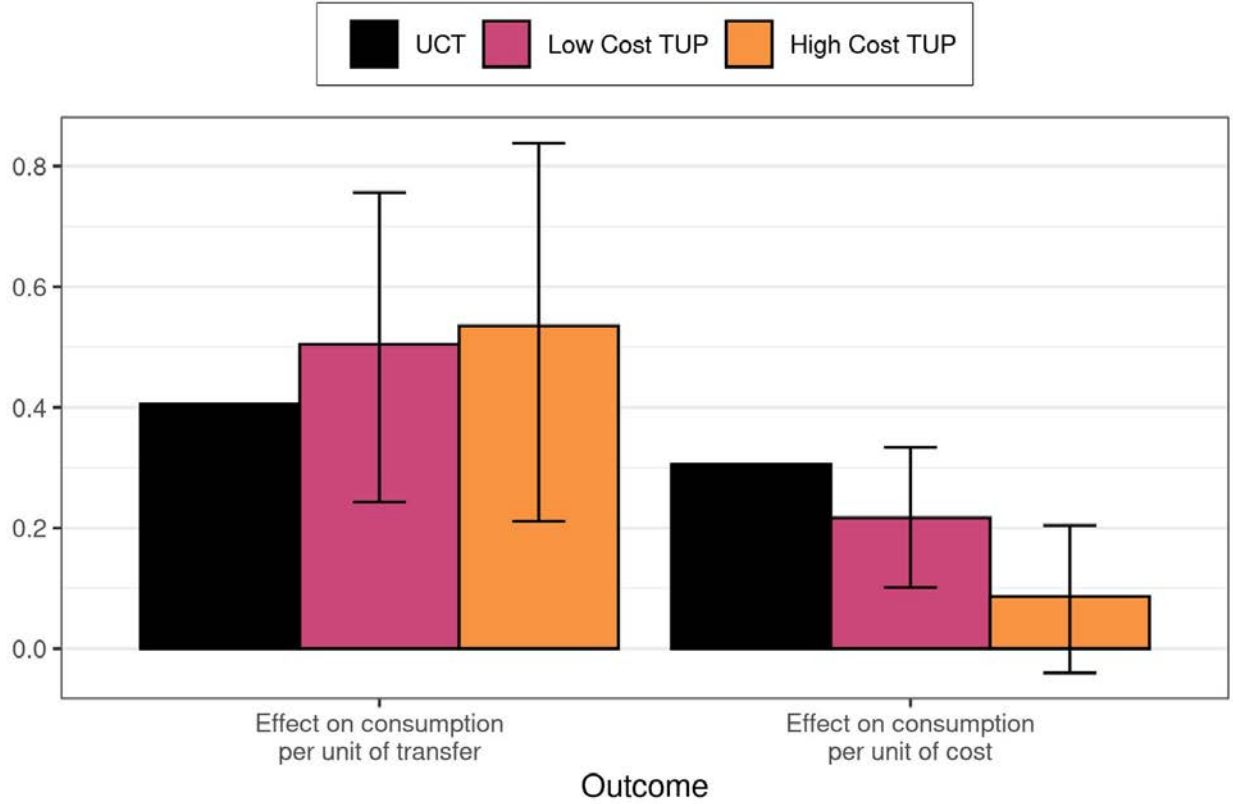
7   Conclusion

In this paper, we innovate by providing theory and formal tests of the impacts of increasing intervention size on cost effectiveness, in the context of temporary unconditional cash transfers (UCT). To implement these tests, we synthesize evidence from a deep literature on the consumption impacts of UCT and TUP programs to provide point estimates for cost effectiveness, external validity, and heterogeneity across contexts. This fills an important gap in a large literature in development economics that has studied the potential of targeted interventions to push households out of poverty. Existing work has documented the impacts of both large and small interventions, increasing intervention scale, intervention type, and complementarities across interventions, while focus on the dynamic impacts of increasing

intervention size through increasing intensity or scope.

We produce three key results. First, we estimate that UCT increase household consumption by 0.35 per unit of transfer. Consistent with the homogeneity of the intervention itself, these estimates are remarkably consistent across contexts compared to other development interventions, and therefore offer a useful benchmark for a broader set of interventions in developing countries. In addition, these impacts are meaningfully more persistent than comparable estimates from richer countries. Second, we find that increasing UCT intensity (through transfer size) reduces cost effectiveness and does not affect persistence. We argue these results are not consistent with predictions from a standard buffer stock model, but can be explained by poverty traps or decreasing returns to scale. Third, we find that increasing UCT scope (through adding complementary graduation program interventions in TUP) reduces cost effectiveness at commonly evaluated time horizons and increases persistence and heterogeneity. This highlights the importance of context-specific estimates of the long run impacts of TUP to inform policy.

These theoretical and empirical results on the impacts of increasing intervention size on cost effectiveness discipline a debate on the need for "big push" interventions. In particular, the presence of poverty traps alone does not justify increasing intervention size. Instead, the distribution of poverty thresholds conditional on targeting is crucial. We note that there are justifications for increasing intervention size beyond cost effectiveness — increasing transfer size or providing complementary programs to the poorest households, for example, can be a powerful tool for poverty reduction.

References

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2020). Sampling-based versus design-based uncertainty in regression analysis. *Econometrica*, 88(1), 265–296.

Auclert, A., Bardóczy, B., & Rognlie, M. (2020). Mpcs, mpes and multipliers: A trilemma

for new keynesian models. *NBER Working Paper*, (w27486).

Auclert, A., Rognlie, M., & Straub, L. (2018). The intertemporal keynesian cross. (25020).

Baird, S., Ferreira, F. H., Özler, B., & Woolcock, M. (2014). Conditional, unconditional and everything in between: a systematic review of the effects of cash transfer programmes on schooling outcomes. *Journal of Development Effectiveness*, 6(1), 1–43.

Balboni, C., Bandiera, O., Burgess, R., Ghatak, M., & Heil, A. (2020). Why do people stay poor?

Bandiera, O., Burgess, R., Das, N., Gulesci, S., Rasul, I., & Sulaiman, M. (2017). Labor markets and poverty in village economies. *The Quarterly Journal of Economics*, 132(2), 811–870.

Banerjee, A., Breza, E., Duflo, E., & Kinnan, C. (2019). Can microfinance unlock a poverty trap for some entrepreneurs? (26346).

Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., Parienté, W., Shapiro, J., Thuysbaert, B., & Udry, C. (2015). A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science*, 348(6236).

Banerjee, A., Duflo, E., & Sharma, G. (2020). Long-term effects of the targeting the ultra poor program. (28074).

Banerjee, A., Karlan, D., Osei, R. D., Trachtman, H., & Udry, C. (2018). Unpacking a multi-faceted program to build sustainable income for the very poor.

Bedoya, G., Coville, A., Haushofer, J., Isaqzadeh, M. R., & Shapiro, J. (2019). No household left behind: Afghanistan targeting the ultra poor impact evaluation.

Blattman, C., Fiala, N., & Martinez, S. (2020). The long-term impacts of grants on poverty: Nine-year evidence from uganda's youth opportunities program. *American Economic Review: Insights*, 2(3), 287–304.

Blattman, C., Green, E. P., Jamison, J., Lehmann, M. C., & Annan, J. (2016). The returns to microenterprise support among the ultrapoor: A field experiment in postwar uganda. *American economic journal: Applied economics*, 8(2), 35–64.

Bossuroy, T., Goldstein, M., Karlan, D., Kazianga, H., Pariente, W., Premand, P., Thomas, C., Udry, C., Vaillant, J., & Wright, K. (2021). Pathways out of extreme poverty: Tackling psychosocial and capital constraints with a multi-faceted social protection program in niger.

Burke, M., Hsiang, S. M., & Miguel, E. (2015). Climate and conflict. *Annual Review of Economics*, 7(1), 577–617.

Carneiro, P. M., Kraftman, L., Mason, G., Moore, L., Rasul, I., & Scott, M. (2020). The impacts of a multifaceted pre-natal intervention on human capital accumulation in early life.

Carroll, C. (2019). Theoretical foundations of buffer stock saving.

Carroll, C. D. & Kimball, M. S. (1996). On the concavity of the consumption function. *Econometrica*, 64(4), 981–992.

Carter, M. R. & Barrett, C. B. (2006). The economics of poverty traps and persistent poverty: An asset-based approach. *The Journal of Development Studies*, 42(2), 178–199.

Croke, K., Hicks, J. H., Hsu, E., Kremer, M., & Miguel, E. (2016). Does mass deworming affect child nutrition? meta-analysis, cost-effectiveness, and statistical power. (22382).

De Mel, S., McKenzie, D., & Woodruff, C. (2008). Returns to capital in microenterprises: evidence from a field experiment. *The quarterly journal of Economics*, 123(4), 1329–1372.

Deaton, A. (1991). Saving and liquidity constraints. *Econometrica*, 59(5), 1221–1248.

Fagereng, A., Holm, M. B., & Natvik, G. J. J. (2019). Mpc heterogeneity and household balance sheets.

Fedorov, V. V. & Leonov, S. L. (2013). *Optimal design for nonlinear response models*. CRC Press.

Ghatak, M. (2015). Theories of poverty traps and anti-poverty policies. *The World Bank Economic Review*, 29(suppl_1), S77–S105.

Haushofer, J. & Shapiro, J. (2016). The short-term impact of unconditional cash transfers to the poor: Experimental evidence from kenya. *The Quarterly Journal of Economics*, 131(4), 1973–2042.

Haushofer, J. & Shapiro, J. (2018). The long-term impact of unconditional cash transfers: Experimental evidence from kenya.

Hussam, R., Rigol, N., & Roth, B. (2020). Targeting high ability entrepreneurs using community information: Mechanism design in the field.

Jappelli, T. & Pistaferri, L. (2010). The consumption response to income changes. *Annual review of economics*, 2, 479–506.

Kaplan, G. & Violante, G. L. (2014). A model of the consumption response to fiscal stimulus payments. *Econometrica*, 82(4), 1199–1239.

Kraay, A. & McKenzie, D. (2014). Do poverty traps exist? assessing the evidence. *Journal of Economic Perspectives*, 28(3), 127–48.

McIntosh, C. & Zeitlin, A. (2018). Benchmarking a child nutrition program against cash: experimental evidence from rwanda.

McKenzie, D. & Woodruff, C. (2008). Experimental evidence on returns to capital and access to finance in mexico. *The World Bank Economic Review*, 22(3), 457–482.

Meager, R. (2019). Understanding the average impact of microcredit expansions: A bayesian hierarchical analysis of seven randomized experiments. *American Economic Journal: Applied Economics*, 11(1), 57–91.

Muralidharan, K. & Niehaus, P. (2017). Experimentation at scale. *Journal of Economic Perspectives*, 31(4), 103–24.

Muralidharan, K., Romero, M., & Wüthrich, K. (2019). Factorial designs, model selection, and (incorrect) inference in randomized experiments. (26562).

Pennings, S. (2021). Cross-region transfer multipliers in a monetary union: Evidence from social security and stimulus payments. *American Economic Review*, 111(5), 1689–1719.

Ralston, L., Andrews, C., & Hsiao, A. (2017). The impacts of safety nets in africa: what are we learning?

Rosenzweig, M. R. & Udry, C. (2019). External Validity in a Stochastic World: Evidence from Low-Income Countries. *The Review of Economic Studies*, 87(1), 343–381.

Vivalt, E. (2020). How Much Can We Generalize From Impact Evaluations? *Journal of the European Economic Association.* jvaa019.

Appendix A   Model appendix

Proof of Proposition 1   First, we rewrite $\tau_0(h) = \frac{c_0(h) - c_0(0)}{h}$, omitting the argument ρ. Taking derivatives yields $\tau_0'(h) = \frac{1}{h}\left[c_0'(h) - \frac{c_0(h) - c_0(0)}{h}\right]$. As h is positive, the term inside the brackets is strictly negative because $c_0'(h)$ is strictly decreasing by strict concavity of $c_0(h)$. It follows that $\tau_0'(h) < 0$.

Proof of Proposition 2   First, rewriting our definition of $c_1(h)$, omitting the argument ρ, we have $c_0(h) + \frac{1}{R}c_1(h) = w_0 + h + \frac{1}{R}y_1$. Substituting $h = 0$, we can write $c_0(h) - c_0(0) + \frac{1}{R}(c_1(h) - c_1(0)) = h$. Dividing by h yields $\tau_0(h) + \frac{1}{R}\tau_1(h) = 1$. Differentiating yields $\tau_1'(h) = -R\tau_0'(h)$, and therefore $\tau_1'(h) > 0$.

Proof of Proposition 2a   We consider the following simplified and deterministic version of Equation 2, with the introduction of decreasing returns to scale.

$$c_0(h) \equiv \arg\max_{c_0} \log c_0 + \beta \log \rho(w_0 + h - c_0)^\alpha$$

where $\alpha < 1$ introduces decreasing returns to scale. Differentiating the objective function with respect to c_0 yields a closed form solution for the household's consumption functions.

$$(c_0(h), c_1(h)) = \left(\frac{1}{1 + \beta\alpha}(w_0 + h), \rho\left(\frac{\beta\alpha}{1 + \beta\alpha}\right)^\alpha (w_0 + h)^\alpha\right)$$

Taking second derivatives of the consumption functions yields

$$(c_0''(h), c_1''(h)) = \left(0, -\rho\alpha(1 - \alpha)\left(\frac{\beta\alpha}{1 + \beta\alpha}\right)^\alpha (w_0 + h)^{-(2-\alpha)}\right)$$

Note that $c_1''(h) = 0$ when $\alpha = 1$, and $c_1''(h) < 0$ when $0 < \alpha < 1$. Noting as above that $\tau_t'(h) = \frac{1}{h^2}\int_0^h \int_0^m c_t''(n)\,dn\,dm$, this implies that $\tau_0'(h) = \tau_1'(h) = 0$ when $\alpha = 1$, and $\tau_0'(h) = 0$ and $\tau_1'(h) < 0$ when $0 < \alpha < 1$. Therefore, introducing decreasing returns to scale decreases

the effect of transfer size on the future period impacts of UCT on consumption per unit of transfer.

Proof of Proposition 2b   We consider the following simplified and deterministic version of Equation 2, where the household has a lumpy investment technology.

$$c_0(h) \equiv \arg\max_{c_0} \log c_0 + \beta \log\left(y_1 + \mathbf{1}\{w_0 + h - c_0 \geq s^*\}\rho s^*\right)$$

If the household saves anything less than s^*, they receive 0 in period 1, but if the household saves at least s^*, they receive ρs^* in period 1. It is therefore always optimal for the household to either save 0 (since they'll receive 0 even if they save more, but less than s^*) or s^* (since they'll receive ρs^* even if they save more). As a result, the household's behavior can be summarized by a threshold h^* – the household saves s^* if and only if h is above h^*. This yields the consumption functions

$$(c_0(h), c_1(h)) = (w_0 + h - \mathbf{1}\{h \geq h^*\}s^*, y_1 + \mathbf{1}\{h \geq h^*\}\rho s^*)$$

The threshold h^* is determined by the household's indifference between saving 0 and saving s^*, that is

$$\log(w_0 + h^*) + \beta \log y_1 = \log(w_0 + h^* - s^*) + \beta \log(y_1 + \rho s^*)$$

Solving for the threshold yields

$$h^* = -w_0 + \frac{\left(1 + \frac{\rho s^*}{y_1}\right)^\beta}{\left(1 + \frac{\rho s^*}{y_1}\right)^\beta - 1} s^*$$

Next, let the expectation operator $\mathbb{E}$ denote the expectation across households, which vary in their resources (w_0, y_1). The consumption functions c_0 and c_1 are now random variables, as

is the threshold h^*. Assume the resources (w_0, y_1) are continuously distributed with convex support, yielding a continuous distribution with convex support for h^*, and let f and F denote the density and cumulative distribution function, respectively, of h^*. We can now write

$$(\mathbb{E}\left[c_0(h)\right], \mathbb{E}\left[c_1(h)\right]) = (\mathbb{E}\left[w_0\right] + h - F(h)s^*, \mathbb{E}\left[y_1\right] + F(h)\rho s^*)$$

Taking second derivatives yields $\frac{d^2}{dh^2}\mathbb{E}c_0(h) = -f'(h)s^*$ and $\frac{d^2}{dh^2}\mathbb{E}c_1(h) = f'(h)\rho s^*$. Substituting these expressions into $\mathbb{E}\tau_0(h)$ and $\mathbb{E}\tau_1(h)$ completes the proof.

Proof of Proposition 3 We consider the following simplified and deterministic version of Equation 2.

$$c_0(h, \rho) \equiv \arg\max_{c_0} \frac{c_0^{1-\eta}}{1-\eta} + \beta \frac{(\rho(w_0 + h - c_0) + y_1)^{1-\eta}}{1-\eta}$$

This problem is equivalent to CES preferences over consumption in period 0 and consumption in period 1, with the budget constraint $c_0 + \frac{1}{\rho}c_1 = w_0 + h + \frac{1}{\rho}y_1$, which has a well known solution.

$$(c_0(h, \rho), c_1(h, \rho)) = \left(\frac{1}{1 + \rho^{(1/\eta)-1}\beta^{1/\eta}}\left(w_0 + h + \frac{1}{\rho}y_1\right), \frac{\rho^{1/\eta}\beta^{1/\eta}}{1 + \rho^{(1/\eta)-1}\beta^{1/\eta}}\left(w_0 + h + \frac{1}{\rho}y_1\right)\right)$$

Differentiating with respect to h and ρ yields $\text{Sign}(\frac{d^2 c_0}{d\rho dh}) = \text{Sign}(1-\eta)$ and $\frac{d^2 c_1}{d\rho dh} > 0$.

Appendix B   Results appendix

## Appendix B.1   Leave one RCT out estimates

We consider the robustness of our results to the set of included studies by replicating our analysis excluding one RCT. We implement this robustness check for 6 estimates that we discuss at length: 1) the mean effect of UCT on consumption per unit of transfer, from Column 5 of Panel A of Table 2, 2) the standard deviation of effects of UCT on consumption

per unit of transfer, from Column 5 of Panel A of Table 2, 3) the effect of increasing log transfer size on UCT impacts on consumption per unit of transfer, from Column 1 of Table 5, 4) the interaction effect of increasing log transfer size and years since last transfer on UCT impacts on consumption per unit of transfer, from Column 2 of Table 5, 5) the effect of TUP on consumption per unit of cost relative to UCT, from Column 1 of Panel B of Table 8, and 6) the interaction effect of TUP relative to UCT and increasing years since last transfer, from Column 2 of Panel B of Table 8. In Figure A1, we present posterior means and 95% credible intervals for each of these parameters leaving out each RCT, and compare to posterior means and credible intervals with all RCTs included.

Figure A1: Leave one RCT out estimates
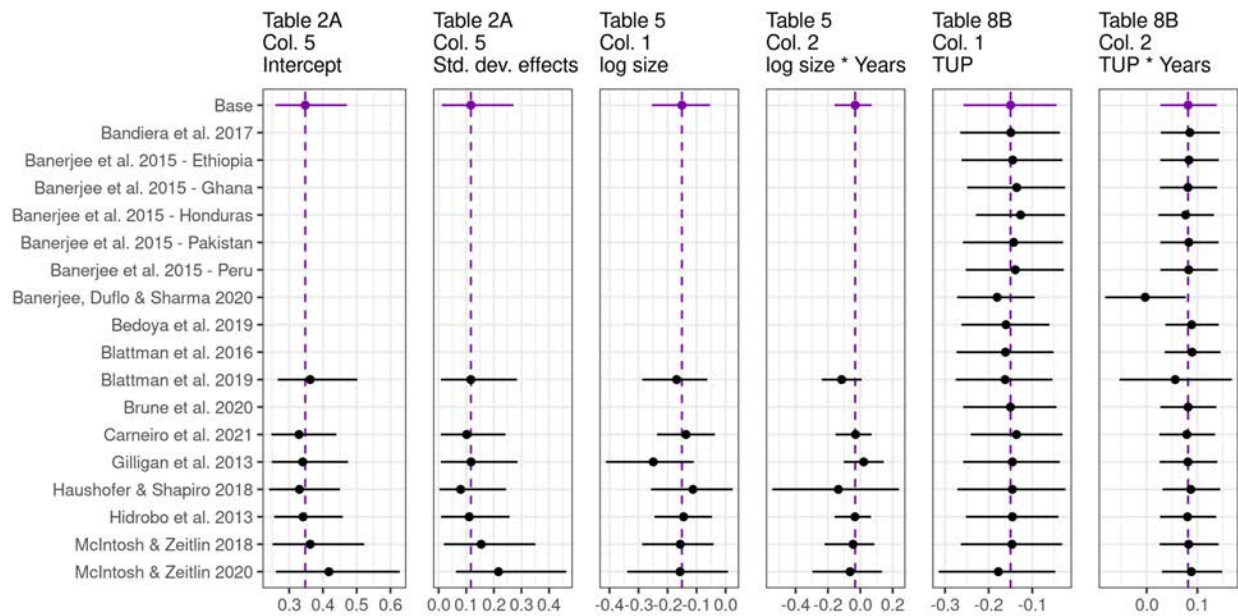
Notes: Posterior means and 95% credible intervals for pararmeters are presented in this figure. "Base" estimates including the full sample of RCTs are in purple, while estimates excluding individual RCTs are in black (with each row of the figure corresponding to a particular excluded RCT). Each of the columns of the figure corresponds to a particular parameter, and headers contain a description of the parameter (along with the Table and Column in which the estimates of the parameter are reported).

We note that our results in general (and in particular the qualitative patterns we describe in the paper) change very little with the exclusion of any individual study, but we note two exceptions to this. First, we lose substantial precision for our estimate of the interaction effect of increasing log transfer size and years since last transfer on UCT impacts on consumption

per unit of transfer when we exclude Haushofer & Shapiro (2016, 2018) — this is because this RCT contains both relatively precise estimates, and contributes a large share of variation in the interaction. However, excluding it makes the posterior mean estimate somewhat more negative, while we show in Section 4.1 that a conventional buffer stock model generates the prediction that this coefficient should be positive. Second, excluding Banerjee et al. (2020) causes our estimate of the interaction effect of TUP relative to UCT and increasing years since last transfer to drop substantially, with the posterior mean estimate becoming negative. This is the only case where the 95% credible interval when we exclude a single RCT no longer includes the posterior mean estimate with the full sample. This is because, as seen in Figure 1, Banerjee et al. (2020) precisely estimate very long run impacts of TUP in India, and these estimates are meaningfully larger than their shorter run estimates. This supports our claim in Section 6.4 that there is a particularly high value to additional long run estimates of UCT and TUP across multiple contexts in determining the cost effectiveness of TUP at increasing consumption relative to UCT.

Appendix B.2   Transfer size and persistence per unit of cost

Our theory and empirics on the role of transfer size in determining the dynamic impacts of UCT on consumption in Section 4 put limited focus on the distinction between transfer size and cost. In practice, this is because transfers represent the dominant share of costs of UCT; as noted in Table 1, the average UCT intervention in our sample delivers 1 unit of transfer for every 1.18 units of cost. However, as shown in Table 4, larger transfers have systematically lower costs per unit of transfer, as much of the cost of delivering UCT is fixed overhead. Therefore, although our estimates in Table 5 show that effects on consumption per unit of transfer are decreasing in transfer size, this does not guarantee that effects on consumption per unit of cost should also decrease in transfer size. In Table A1, we therefore replicate our analysis in Table 5, but use effects on consumption per unit of cost as our outcome of interest instead of effects on consumption per unit of transfer. Although coefficients on log

transfer size decrease in magnitude, we still find that effects on consumption per unit of cost are significantly decreasing in transfer size.

Table A1: Transfer size and persistence per unit of cost

	Effect on consumption per unit of cost					
	(1)	(2)	(3)	(4)	(5)	(6)
β_1: Years since last transfer	-0.021	0.200	-0.015	0.250	-0.030	0.032
	(0.015)	(0.318)	(0.017)	(0.339)	(0.033)	(0.493)
	{-0.051, 0.007}	{-0.380, 0.881}	{-0.047, 0.017}	{-0.369, 0.977}	{-0.097, 0.032}	{-0.907, 1.067}
β_2: log transfer size	-0.100	-0.078	-0.082	-0.051	-0.153	-0.142
	(0.042)	(0.055)	(0.047)	(0.064)	(0.074)	(0.114)
	{-0.183, -0.019}	{-0.184, 0.034}	{-0.174, 0.012}	{-0.171, 0.080}	{-0.299, -0.008}	{-0.362, 0.095}
β_3: Years since last transfer * log transfer size		-0.032		-0.039		-0.009
		(0.046)		(0.049)		(0.070)
		{-0.132, 0.052}		{-0.145, 0.051}		{-0.158, 0.124}
σ_τ: Std. dev. of effects	0.043	0.053	0.049	0.060	0.063	0.080
	(0.037)	(0.045)	(0.043)	(0.049)	(0.054)	(0.067)
	{0.003, 0.140}	{0.003, 0.167}	{0.003, 0.159}	{0.003, 0.185}	{0.003, 0.199}	{0.004, 0.250}
Controls						
Transfer duration			X	X		
RCT fixed effects					X	X
# of RCTs	7	7	7	7	5	5
# of observations	18	18	18	18	16	16

Notes: Estimates of the model of impacts of UCT interventions on household consumption per unit of cost in Equation 3 are presented in this table. Columns 1 through 6 report Bayesian posterior means of each parameter, with posterior standard deviations in parentheses and 95% credible intervals in curly brackets.

Appendix C   Internal rate of return

To summarize the relative cost effectiveness of TUP, we calculate the internal rate of return for shifting a UCT intervention to TUP using our estimates in Columns 2 and 4 of Panel B in Table 8. Specifically, we solve

$$\int_0^T \left(\frac{1}{1 + \text{IRR}/100} \right)^t (\beta_1 + \beta_2 t) dt = 0$$

for the internal rate of return (IRR), where β_1 and β_2 are the impacts of TUP and TUP * Years since last transfer, respectively, on the intervention effect on consumption per unit of cost. The parameter T governs the number of years since last transfer to which we extrapolate. We plot internal rates of return using our estimates from Column 2 (with no controls) and Column 4 (with controls) at a range of time horizons in Figure A2.

Figure A2: Internal rate of return of shifting UCT to TUP
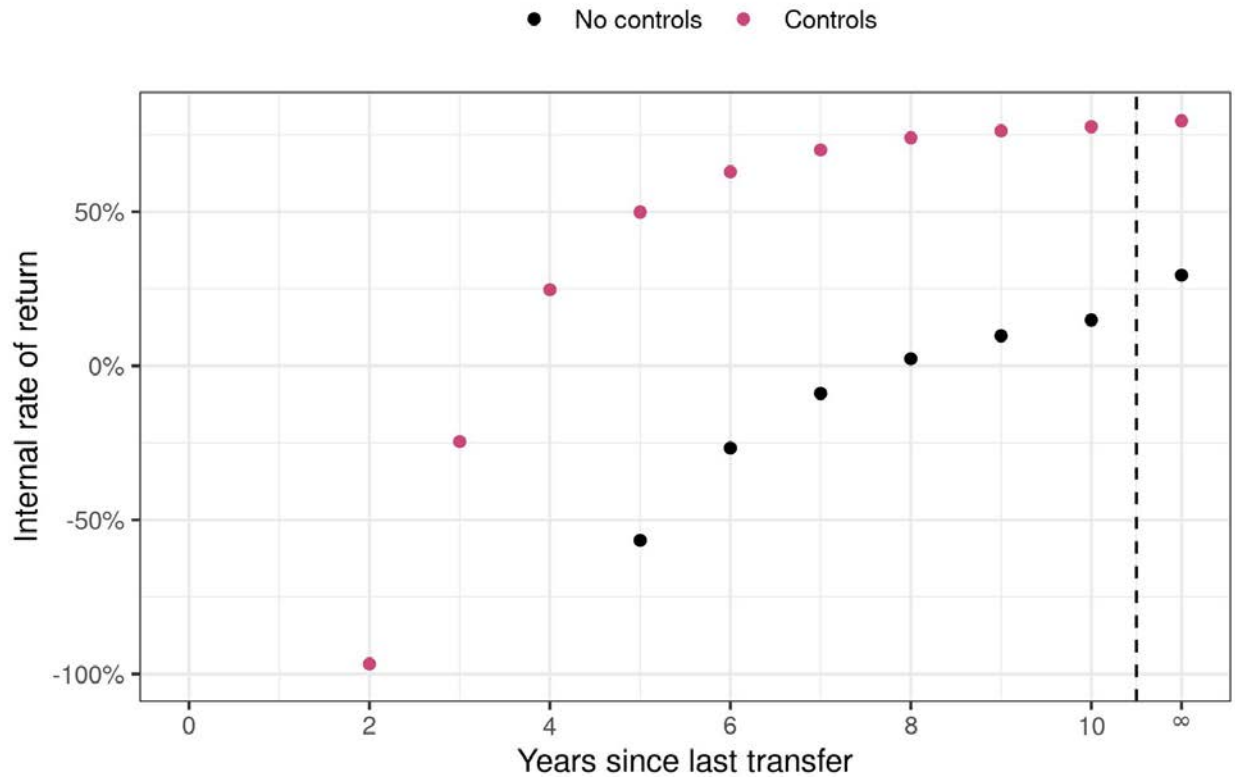
Appendix D Power and optimal experimental designs

Two of our primary findings highlight the importance of increasing the statistical power of future work — for transfer size, we fail to reject that larger UCT do not have differentially persistent impacts on consumption, and for complementary interventions, we find evidence of significant heterogeneity in the impacts of TUP relative to UCT across contexts. To improve precision in future work, we derive optimal experimental designs for estimating the cost effectiveness of increasing intervention size through either transfer size or adding complementary interventions. In addition, we produce a dashboard for implementing power calculations for researchers implementing experiments to compare cost effectiveness across interventions of different sizes.[22]

---

[22]The dashboard is available at `https://datanalytics.worldbank.org/connect/#/apps/759`.

Appendix D.1 Intervention intensity

We now consider the problem of a researcher who is designing an experiment to estimate heterogeneity of impacts of cash transfers per unit of transfer with respect to transfer size. This general problem of optimal design when the researcher is interested in estimating a nonlinear model has been studied in the statistics literature (Fedorov & Leonov, 2013), and our problem is a special case of estimating a nonlinear dose-response function. We consider a highly tractable version of the problem where the researcher has a linear null hypothesis and a nesting alternative hypothesis that includes a nonlinear function of cash transfer size in a linear model, and we restrict the researcher to only use two different cash transfer sizes in their experimental design to test the null. This version of the problem is empirically relevant, as the studies in this metaänalysis with multiple cash transfer sizes all used between 2 and 4 different transfer sizes, and any published tests for nonlinearities in these studies took this form.

We begin with a version of the model we estimate in Section 4.3, where we allow the impact of cash transfers per unit of transfer to vary by transfer size, although to simplify the analysis we assume the researcher has access to only a single cross section. For concreteness, we suppose throughout this section that the researcher is interested in consumption as an outcome. Letting τ_a be the average treatment effect on consumption of cash transfer a per unit of transfer, and Size_a be the size of the cash transfer, we assume

$$\tau_a = \beta_1 + \beta_2 f(\text{Size}_a) \tag{A1}$$

for known non-constant f. Multiplying by Size_a implies the following model of household consumption, where Y_i is the consumption of household i

$$Y_i = \beta_0 + \beta_1 \text{Size}_{a(i)} + \beta_2 \text{Size}_{a(i)} f(\text{Size}_{a(i)}) + \epsilon_i \tag{A2}$$

where $a(i)$ is the cash transfer received by household i. When $\text{Size}_{a(i)}$ is randomly assigned across households and Equation A1 holds, then ϵ_i is a mean 0 error conditional on $\text{Size}_{a(i)}$, with conditional variance $\sigma_\epsilon^2(\text{Size}_{a(i)})$.[23]

As the average treatment effect of a cash transfer of size Size_a, τ_a, is not directly observed, the researcher runs an experiment in which N_0 households are randomly assigned to a control group, N_1 households are randomly assigned to receive a cash transfer of size Size_1, and N_2 households are randomly assigned to receive a cash transfer of size Size_2; assume without loss of generality that $\text{Size}_2 > \text{Size}_1$. The researcher is primarily interested in estimating β_2; rejecting the null that $\beta_2 = 0$ would correspond to a rejection of a linear model of the impacts of cash transfers. To estimate β_2, the researcher estimates Equation A2 by least squares, yielding the least squares estimator $\hat{\beta}_2$ of β_2.

The researcher is interested in maximizing their precision to test for the presence of nonlinearities, so they select the number of observations assigned to each arm and the size of cash transfers, $(N_0, N_1, N_2, \text{Size}_1, \text{Size}_2)$, in order to minimize the variance of their estimator of nonlinearities, $\text{Var}[\hat{\beta}_2]$. We solve this problem in three steps. First, we calculate the variance of this estimator. Second, we solve for the optimal number of households assigned to each treatment arm conditional on the choice of cash transfer sizes and the total number of observations. Third, we provide results on the optimal cash transfer sizes.

To begin, we derive the variance of the least squares estimator of nonlinearities in the impacts of cash transfers on consumption, $\hat{\beta}_2$. First, we write $\hat{\beta}_2$ as a function of the average outcome in each of the three treatment arms. Letting $\overline{Y}_a$ be the average consumption of

---

[23]To interpret this conditional variance, let $Y_i(s)$ denote individual i's outcome when assigned a cash transfer of size s. Then $\sigma_\epsilon^2(s) - \sigma_\epsilon^2(0) = \text{Var}\left[Y_i(s) - Y_i(0)\right] + \text{Cov}(Y_i(s) - Y_i(0), Y_i(0))$, where variances are taken across individuals. Therefore, larger variances of treatment effects and larger covariances of treatment effects with baseline outcomes are associated with increasing conditional variance of the average outcome with respect to transfer size.

households assigned to treatment arm a,

$$\hat{\beta}_2 = \frac{1}{\text{Size}_2\left(f(\text{Size}_2) - f(\text{Size}_1)\right)}\overline{Y}_2 - \frac{1}{\text{Size}_1\left(f(\text{Size}_2) - f(\text{Size}_1)\right)}\overline{Y}_1$$
$$+ \frac{\text{Size}_2 - \text{Size}_1}{\text{Size}_1\text{Size}_2\left(f(\text{Size}_2) - f(\text{Size}_1)\right)}\overline{Y}_0 \quad \text{(A3)}$$

Taking the variance of both sides, we derive the variance of $\hat{\beta}_2$.

$$\text{Var}[\hat{\beta}_2] = \frac{\sigma_\epsilon^2(\text{Size}_2)}{N_2\text{Size}_2^2\left(f(\text{Size}_2) - f(\text{Size}_1)\right)^2} + \frac{\sigma_\epsilon^2(\text{Size}_1)}{N_1\text{Size}_1^2\left(f(\text{Size}_2) - f(\text{Size}_1)\right)^2}$$
$$+ \frac{\sigma_\epsilon^2(0)\left(\text{Size}_2 - \text{Size}_1\right)^2}{N_0\text{Size}_1^2\text{Size}_2^2\left(f(\text{Size}_2) - f(\text{Size}_1)\right)^2} \quad \text{(A4)}$$

Next, we consider the optimal choice of (N_0, N_1, N_2) conditional on $(\text{Size}_1, \text{Size}_2, N_0 + N_1 + N_2)$. This corresponds to a common evaluation problem, where the researcher does not have control over the sizes of cash transfers, and faces a budget constraint limiting the total number of observations. Taking the first order conditions of the minimization problem subject to the constraint that $N_0 + N_1 + N_2 \leq N$ yields the following optimal choices of assignments

$$(N_0^*, N_1^*, N_2^*) \propto (\sigma_\epsilon(0)(\text{Size}_2 - \text{Size}_1), \sigma_\epsilon(\text{Size}_1)\text{Size}_2, \sigma_\epsilon(\text{Size}_2)\text{Size}_1) \quad \text{(A5)}$$

We make a few observations about these optimal choices. First, the assumed f does not matter for the optimal design — this is because, conditional on f, Size_1, and Size_2, β_2 is linear in the difference between estimated impacts per unit of transfer of the small and the large cash transfer, so the optimal design minimizes the variance of this difference. Second, when errors are homoskedastic (so $\sigma_\epsilon(s) = \sigma_\epsilon$), and the smaller transfer is half the size of the larger transfer (so $\text{Size}_2 = 2\text{Size}_1$), these yield observations assigned to control, the smaller cash transfer, and the larger cash transfer in the ratio 1:2:1. Third, more observations are assigned to cash transfer arms when the variance of errors is larger for individuals assigned

to that arm. For example, if the variance of consumption were larger across households who receive a large cash transfer, this would suggest assigning relatively more individuals to larger cash transfers. Fourth, when the larger cash transfer is much larger than the smaller cash transfer, the rule assigns very few individuals to the large cash transfer. This highlights that, when a much larger cash transfer is an option, relatively few individuals may need to be added to the larger cash transfer arm to precisely estimate nonlinearities. Fifth, the average cash transfer distributed, that is, $\frac{N_2^* \text{Size}_2 + N_1^* \text{Size}_1}{N}$, is approximately Size_1 when errors are homoskedastic and Size_2 is large. This is because the optimal rule allocates proportionately fewer observations to the larger cash transfer as the larger cash transfer increases in size. Therefore, increasing the size of the larger cash transfer may not have large budget implications when the implementer is constrained.

Lastly, we consider properties of the variance of the estimator of nonlinearities with respect to transfer size under the optimal assignment. Substituting (N_0^*, N_1^*, N_2^*) into Equation A4 yields the following expression for the variance of this estimator under the optimal assignment, which we denote $\text{Var}^*[\hat{\beta}_2]$.

$$N\text{Var}^*[\hat{\beta}_2] = \frac{\left[(\sigma_\epsilon(0) + \sigma_\epsilon(\text{Size}_1))\frac{\text{Size}_2}{\text{Size}_1} + (\sigma_\epsilon(\text{Size}_2) - \sigma_\epsilon(0))\right]^2}{\text{Size}_2^2 \left(f(\text{Size}_2) - f(\text{Size}_1)\right)^2} \tag{A6}$$

We make two observations about this variance, considering the case when the variance of errors is homoskedastic and not affected by transfer size (so $\sigma_\epsilon(s) = \sigma_\epsilon$), as similar results hold with sufficient restrictions on heteroskedasticity. In addition, we assume $f(s) = s^\alpha$, so α is the elasticity of the impacts of cash transfers on consumption per unit of transfer to transfer size relative to the impacts of a small transfer — although this is not a realistic model, it is useful to build intuition.[24] First, power to detect nonlinearities is increasing in transfer size (holding fixed the relative size of the larger cash transfer) if and only if $\alpha > -1$. For negative α, Equation A1 imposes that for sufficiently large transfer sizes, the impacts of cash transfers

---

[24]Under these assumptions, the above expression simplifies to $N\text{Var}^*[\hat{\beta}_2] = \sigma_\epsilon^2 \text{Size}_1^{-2-2\alpha} \left(\frac{\text{Size}_2^\alpha}{\text{Size}_1^\alpha} - 1\right)^{-2}$.

per unit of consumption approach a constant β_1. When this happens sufficiently quickly, comparing two relatively large cash transfers will not be informative about nonlinearities, as the differences in impacts of these two cash transfers will be very similar. Second, power to detect nonlinearities is increasing in the relative size of the larger cash transfer. This is driven by two factors. First, by assumption, the difference in the impact of the larger cash transfer and the smaller cash transfer, expressed per unit of transfer, is increasing in the size of the larger cash transfer, holding fixed the size of the smaller cash transfer. Second, increasing the size of the larger cash transfer also increases the precision with which the impact of the larger cash transfer per unit of transfer can be estimated. Both of these forces cause power to be increasing in the relative size of the larger cash transfer. As the optimal budget distributed through cash transfers, expressed as a function of the relative size of the larger cash transfer, is bounded, these results suggest that increasing the relative size of the larger cash transfer may be a relatively low cost approach to increasing the power to detect nonlinearities in the impacts of cash transfers on household consumption.

Appendix D.2 Intervention scope

We next consider the problem of a researcher who is designing an experiment with two sets of objectives: to estimate the impacts of two given interventions on household consumption, and to estimate the differences in the cost effectiveness of these interventions at increasing household consumption. We refer to these two interventions as UCT and TUP, but we consider this a placeholder for the interventions of interest to the researcher.

Similar to Appendix D.1, the researcher runs an experiment in which N_0 households are randomly assigned to a control group, N_{UCT} households are randomly assigned to receive a cash transfer, and N_{TUP} households are randomly assigned to receive TUP. We denote, for notational consistency with Appendix D.1, the cost of the cash transfer arm with Size_{UCT}, and the cost of TUP with Size_{TUP}. The researcher then estimates the following model of

household consumption, where Y_i is the consumption of household i

$$Y_i = \beta_0 + \beta_{UCT}\text{UCT}_i + \beta_{TUP}\text{TUP}_i + \epsilon_i \tag{A7}$$

where UCT_i and TUP_i are indicators that the household was assigned to receive a cash transfer and assigned to receive TUP, respectively. We assume ϵ_i is a mean 0 error conditional on assignment, and because assignment is random we can interpret β_{UCT} and β_{TUP} as the average treatment effects of the cash transfer and of TUP, respectively. To simplify discussion, we assume the variance of ϵ_i conditional on assignment is σ_ϵ^2, but as in Appendix D.1 this is straightforward to generalize. The researcher estimates Equation A7 by least squares, yielding the least squares estimator $\hat{\beta}$ of β.

As stated above, the researcher is interested in three estimands of interest. The first and second are the average treatment effects of UCT and TUP per unit of cost, which are estimated by $\hat{\beta}_{UCT}/\text{Size}_{UCT}$ and $\hat{\beta}_{TUP}/\text{Size}_{TUP}$, respectively. The third is the difference in the cost effectiveness of UCT and TUP at increasing household consumption, which is estimated by $\hat{\beta}_{TUP}/\text{Size}_{TUP} - \hat{\beta}_{UCT}/\text{Size}_{UCT}$. Letting $\overline{Y}_a$ be the average consumption of households assigned to treatment arm a, for $a \in \{0, UCT, TUP\}$, note that $\hat{\beta}_{UCT} = \overline{Y}_{UCT} - \overline{Y}_0$ and $\hat{\beta}_{TUP} = \overline{Y}_{TUP} - \overline{Y}_0$. Substituting and applying variance formulas yields the following formula for the variance of the two estimators of interest.

$$\text{Var}\left[\frac{\hat{\beta}_{UCT}}{\text{Size}_{UCT}}\right] = \frac{1}{\text{Size}_{UCT}^2}\left(\frac{\sigma_\epsilon^2}{N_{UCT}} + \frac{\sigma_\epsilon^2}{N_0}\right) \tag{A8}$$

$$\text{Var}\left[\frac{\hat{\beta}_{TUP}}{\text{Size}_{TUP}}\right] = \frac{1}{\text{Size}_{TUP}^2}\left(\frac{\sigma_\epsilon^2}{N_{TUP}} + \frac{\sigma_\epsilon^2}{N_0}\right) \tag{A9}$$

$$\text{Var}\left[\frac{\hat{\beta}_{TUP}}{\text{Size}_{TUP}} - \frac{\hat{\beta}_{UCT}}{\text{Size}_{UCT}}\right] = \frac{1}{\text{Size}_{TUP}^2}\frac{\sigma_\epsilon^2}{N_{TUP}} + \frac{1}{\text{Size}_{UCT}^2}\frac{\sigma_\epsilon^2}{N_{UCT}} + \frac{(\text{Size}_{UCT} - \text{Size}_{TUP})^2}{\text{Size}_{UCT}^2\text{Size}_{TUP}^2}\frac{\sigma_\epsilon^2}{N_0}$$

$$\tag{A10}$$

As the researcher is interested in the average treatment effect of UCT, the average treat-

ment effect of TUP, and also the difference in the cost effectiveness of UCT and TUP, the researcher wishes to minimize the weighted sum of these variances. The researcher faces a constraint on the number of observations, $N_0 + N_{UCT} + N_{TUP} \leq N$. The researcher therefore solves

$$\min_{N_0, N_{UCT}, N_{TUP}} \alpha_{UCT} \mathrm{Var}\left[\frac{\hat{\beta}_{UCT}}{\mathrm{Size}_{UCT}}\right] + \alpha_{TUP} \mathrm{Var}\left[\frac{\hat{\beta}_{TUP}}{\mathrm{Size}_{TUP}}\right]$$
$$+ (1 - \alpha_{UCT} - \alpha_{TUP}) \mathrm{Var}\left[\hat{\beta}_{TUP} - \frac{\mathrm{Size}_{TUP}}{\mathrm{Size}_{UCT}}\hat{\beta}_{UCT}\right] \quad \text{(A11)}$$
$$N_0 + N_{UCT} + N_{TUP} \leq N$$

where α_{UCT} and α_{TUP} is the weight the researcher places on the variance of the estimated average treatment effect of UCT and TUP, respectively, and $1 - \alpha_{UCT} - \alpha_{TUP}$ is the weight the researcher places on the variance of the cost effectiveness of TUP relative to UCT.

We then take the first order conditions of this minimization problem subject to the constraint that $N_0 + N_{UCT} + N_{TUP} \leq N$. Letting $A \equiv \mathrm{Size}_{TUP}/\mathrm{Size}_{UCT}$, yields the following optimal choices of assignments

$$(N_0^*, N_{UCT}^*, N_{TUP}^*) \propto$$
$$\left(\sqrt{\alpha_{UCT}A + \alpha_{TUP}A^{-1} + (1 - \alpha_{UCT} - \alpha_{TUP})\left(\sqrt{A} - \sqrt{A^{-1}}\right)^2},\right.$$
$$\left.\sqrt{(1 - \alpha_{TUP})A}, \sqrt{(1 - \alpha_{UCT})A^{-1}}\right) \quad \text{(A12)}$$

We make a few observations about these optimal choices. First, when the researcher only cares about the cost effectiveness of TUP relative to UCT (e.g., the researcher is interested in which intervention is optimal to scale up), and the UCT and TUP are the same cost (so $A = 1$), then the optimal design assigns no observations to control, as the estimator of the difference in cost effectiveness is proportional to the difference in means of the groups assigned to TUP and UCT. However, when UCT and TUP are difference costs, a control group is

necessary so the effects of UCT and TUP can each be estimated and then standardized by their costs. Second, when UCT and TUP are equal size, the common design of assigning $1/3$ of observations to each arm is justified by weights of $\alpha_{UCT} = \alpha_{TUP} = 1 - \alpha_{UCT} - \alpha_{TUP} = 1/3$ — that is, when the researcher puts equal weight on minimizing the variance of each of the three estimators. However, in many cases, we would anticipate the average treatment effect of either UCT or TUP to be larger than the difference in their effects, meaning additional weight is needed on the variance of the difference in cost effectiveness — in general, this motivates a design with fewer observations assigned to control. In contrast, when the researcher does not care about the cost effectiveness of TUP relative to UCT but only about the effects of each individual intervention, this motivates a design with additional observations assigned to control.

Lastly, we consider properties of the minimized weighted sum of the variances of the estimators of the cost effectiveness of UCT, the cost effectiveness of TUP, and the difference in cost effectiveness of UCT and TUP evaluated at the optimal design, which we denote V^*. Specifically, we substitute $(N_0^*, N_{UCT}^*, N_{TUP}^*)$ into the minimand Equation A11. This yields

$$
NV^* = \frac{1}{\text{Size}_{UCT}\text{Size}_{TUP}} \left(\sqrt{\alpha_{UCT}A + \alpha_{TUP}A^{-1} + (1 - \alpha_{UCT} - \alpha_{TUP})\left(\sqrt{A} - \sqrt{A^{-1}}\right)^2} \right.
$$
$$
\left. + \sqrt{(1 - \alpha_{TUP})A} + \sqrt{(1 - \alpha_{UCT})A^{-1}} \right)^2 \quad \text{(A13)}
$$

Here, we note that, an increase in either Size_{UCT} or Size_{TUP} will always decrease V^* — intuitively, cost effectiveness is more precisely estimated with larger interventions. Note that this guidance does not necessarily imply that these estimates will be higher powered for larger interventions, as changing the size of interventions may also change their impacts.