Chapter 8

Measuring the Long-term Impact of Cash Transfers

Karen Macours

Paris School of Economics; INRAE

Tania Barham

University of Colorado, Boulder

John Maluccio

Middlebury College

Gelson Tembo

University of Zambia; Palm Associates, Lusaka, Zambia

1. Introduction

Cash transfer programs often have multiple objectives, from providing a safety net to breaking the intergenerational transmission of poverty. This combination of short-term poverty alleviation and long-term economic mobility lay at the heart of

the first-generation programs in Latin America, starting with Mexico's Progresa in 1997. Regular cash transfers were given to increase household consumption and provide a semi-permanent safety net. At the same time, monitored conditions were included to increase household investment in children's human capital, with the expectation that the eventual returns from those investments would allow beneficiaries to lead more economically productive adult lives. As conditional cash transfer (CCT) programs spread worldwide, design features proliferated and programs employed different theories of change, but all programs with conditions related to health or education investments had similar long-term objectives.

In addition to potential long-term returns to induced human capital investments, cash transfer programs can also lead to longer-term impacts through other channels. Conditions and social marketing, while constraining decisions, provide information (e.g., regarding dietary diversity, prenatal healthcare, vaccination or the value of education) thereby reducing information failures and affecting household decision making. Both CCTs and unconditional cash transfers (UCTs), by relaxing internal liquidity or insurance constraints, can also trigger increased productive investments in livestock, agricultural and business assets, or migration, leading to income and consumption multipliers. In some cases exposure to cash transfers can change preferences or aspirations, or affect mental health, all of which can influence decision making. And where the targeting of transfers shifts intrahousehold bargaining positions, different intrahousehold equilibria might emerge. For

example, transfers made to women could lead to semi-permanent changes in household composition through marriage, divorce or migration.

In sum, there are many channels, as well as possible pathways through those channels, by which cash transfer programs can have long-term impacts. This complexity can make long-term returns uncertain, for example if other constraints or offsetting effects influence the pathways. A growing literature examines whether, under what conditions, for whom and how impacts materialize. In addition to challenges related to the multitude of potential pathways, there are also significant methodological challenges to assessing long-term impacts. Assessment can be difficult in part because most early evaluations were not originally designed to examine absolute long-term returns, and because mobility and selection becoming increasingly important over longer timeframes. Complications due to mobility are most obviously present when studying the long-term effects for individuals who typically change households (and geographic location) to continue schooling, marry or seek employment. Given that many programs target poor populations in remote or marginalized areas with limited economic opportunities, program-induced shifts in migration patterns are expected as beneficiaries take advantage of returns to increases in human capital. As such, robust approaches to addressing selection and attrition are crucial to the study of long-term effects.

Finally, cash transfer programs often target households as well as specific individuals within them, and long-term returns could be important for both. Fully

understanding the long-term returns to program-induced human capital investment in individual children requires waiting for a sufficiently long time until they have become economically active adults. In contrast, long-term returns resulting from increases in household productive investments might be ascertained over shorter timeframes. But measuring productive investment returns while transfers are ongoing, although indicative of mechanisms leading to long-term impacts (e.g., asset accumulation), does not provide rigorous evidence of their persistence. To examine returns to household productive investments we focus in this chapter on evidence of multiplier effects for programs that have been running over many years, or on outcomes measured two or more years after eligibility for the program has ended.

The chapter is organized as follows. In section 2 we discuss methodological challenges for studying long-term returns, highlighting common pitfalls and potential solutions. We then lay out a conceptual framework for considering the different pathways towards long-term returns to identify the contextual variables and types of outcomes to measure. Following the framework, we draw lessons from several studies with well-identified causal estimates.[1] Section 4 highlights evidence

[1] A complementary approach, not considered in this chapter, is to use shorter-term experiments on cash transfers as inputs into the estimation of structural models, in turn used to predict longer-term outcomes (Todd and Wolpin 2023). Related, the chapter does not engage with the open question of whether surrogate estimates can replace direct estimation of long-term effects (Athey et al. 2019).

on long-term (after five or more years) returns to human capital investments, and section 5 covers the evidence on multiplier effects and productive investments two or more years after the end of transfers. Cash-for-work or graduation (ultra-poor) programs, which have different theories-of-change differ from cash transfer programs, are covered elsewhere in the handbook (Alik-Lagrange and Imbert 2025; Bandiera et al. 2025). Most of the methodological considerations we describe, however, apply to estimating long-term impacts of such programs. Section 6 concludes, highlighting several key open questions.

2. Methodological considerations

2.1 Selection due to attrition, incomplete administrative data, mortality and fertility

The first obvious challenge any long-term assessment needs to tackle is potential selection. Whether using primary or secondary data, information on long-term outcomes for a sample not fully representative of the population of treatment and control groups before the start of the program jeopardizes the ability to draw valid and policy relevant conclusions. Many long-term studies focus on relatively young cohorts who benefitted from cash transfers early in life (0–6 years) or at school-going ages (~7–17 years), or who were targeted with cash grants for productive investments as young adults. Since young people tend to be particularly mobile,

---

Evidence to date suggests the ability of those methods to approximate long-term returns accurately may be limited (Bernard 2023).

longitudinal survey attrition concerns are central and pose challenges for both internal and external validity (Molina-Millán and Macours 2025). Attrition cannot be ignored or assumed away, since individual mobility is precisely what we would expect to happen following investments in human capital, for example as seen in research on rural productivity gaps (Gollin et al. 2014). Cash transfers can also alter who migrates, leading to selective attrition where the types of persons researchers are able to interview at follow-up differ between treatment and control groups in observable as well as possibly unobservable ways. Several studies on long-term impacts examine samples with substantial attrition, severely limiting their scope for reaching valid conclusions.[2]

The problems related to potential attrition can be addressed in different ways. Studies based on primary data collection should have protocols and resources in place to intensively track migrants and ensure information is collected from a sample representative of the target population.[3] Protocol development should start

[2] For example, the six-year follow-up of Progresa had 40% attrition for individuals exposed at school-going ages, rising to 60% by the 10-year follow-up. In contrast, by implementing comprehensive tracking protocols and other complementary strategies (including phone surveys), attrition in the 20-year follow-up was 15%, and 7% for outcomes with proxy reports.

[3] Strong protocols typically involve starting the follow-up survey in the original localities and collecting detailed tracking information for respondents not found there (reason for absence, expected date of return, phone number, and address and other contacts in their destination locale).

from a diagnostic of the reasons for attrition and their relationship to the primary research questions. Studies focused on long-term productive interventions of household-level transfers in rural settings might find most beneficiaries still in their original households or nearby locations many years later (e.g., Handa et al. 2019). In contrast, studies focusing on youth in settings where migration is common should expect the opposite. In the 20-year follow-up of Progresa only one-third of former beneficiaries exposed at school-going ages still lived in their original communities. Tracking to all destinations can come with substantial financial costs, but not doing so risks introducing selection bias to such a large extent that it defeats the purpose of the long-term study. At a minimum, studies with high levels of attrition need to limit their conclusions to the selective subsamples, i.e., individuals who could more easily be found. Unfortunately, that group may not be the most relevant for understanding long-term returns. At worst, there will be differential selective attrition in treatment and control groups, so that even obtaining internally valid estimates is difficult.

After that, protocols can include repeat visits to localities to interview visiting or return migrants, continuously updating and acting on new information and tracking migrants to all new locations through in-person tracking or phone interviews. Carefully designing tracking protocols is paramount because if they are not transparent and understood by the population they can generate suspicion and mistrust and thus compromise survey efforts. Piloting enables field teams to understand those potential pitfalls and develop strategies to address them.

Two data collection strategies to complement intensive tracking include implementing phone interviews and collecting proxy information. Phone interviews can lower costs and also potentially help circumvent security and confidentiality concerns that arise when tracking (in some instances undocumented) migrants. Information collected by asking proxy respondents about individuals who cannot be interviewed directly can reduce attrition for some variables. But both strategies yield more basic measures than direct in-person interviews. Phone surveys also cannot be too long and are generally inadequate for direct assessment of learning or cognition. Proxy reports need to be restricted to easily observed outcomes; depending on the context, parents or family members may or may not be able to accurately report key outcomes such as education, location, occupation or fertility. Both strategies could introduce bias, so that validation on subsamples overlapping with direct in-person interviews should be part of the research design.

An alternative to primary data collection with intensive tracking is to use population census or administrative data that cover the universe of potentially eligible individuals from both treatment and control. In principle population census data (Molina-Millán et al. 2020; Parker and Vogl 2023) has information on all individuals in a country, regardless of where they currently reside so that attrition is not a concern. Similarly, many governmental administrative databases are designed to include all individuals in certain subpopulations. For example,

administrative school records are considered complete in some countries and can contain information on attendance, enrollment, grade progression and sometimes even test scores, with the latter enabling the evaluation of impacts on learning (Duque et al. 2018; Gazeaud and Ricard 2024). In addition, different types of administrative data, such as birth or crime records, offer opportunities to examine other outcomes (Attanasio et al. 2021).[4]

Population census or administrative data sources, however, are only useful for estimation of long-term impacts under two conditions: 1) the data must have information to infer initial treatment status and when exposure began; and 2) the data must cover the universe of the potentially eligible population.

First, for programs with geographical targeting, information on place of birth (or residential location during childhood) can determine treatment assignment. Alternatively, national administrative data may include individual identifiers that enable matching with cash transfer program administrative data to determine eligibility. Such matching, while in theory powerful, can be imperfect. In particular,

[4] Population census or administrative data are particularly valuable for studying outcomes that may be uncommon such as teen pregnancy, crime, mortality, tertiary education or international migration. Analysis of rarer phenomena needs to be based on large samples or risk being underpowered. A trade-off is that these data types typically contain only basic information, hindering exploration of full pathways to impact.

selection bias concerns arise if large fractions of individuals cannot be matched and the underlying reasons for unsuccessful matches are unknown.

Second, many types of administrative data are only partial by construction (for example because not all children are enrolled in school or take exams, or individuals migrate out of the region covered), leading to other potential selection concerns. National level census or administrative data also typically do not contain international migrants, and because cash transfers can affect international migration directly, this has implications for internal and external validity. Whether and to what extent international migration is relevant for the study population will be context dependent. A separate concern stems from possible differential mortality. When program exposure occurs in early life, infant and child mortality can be directly affected (Barham 2011), so that the population of children surviving 5, 10 or 20 years later is selected. Even among older children selective mortality is a possible concern, in particular in settings where mortality of young adults may be related to risk-taking behaviors that are themselves affected by the cash transfers.

Whereas international migration and mortality lead to attrition, differential fertility responses to cash transfers introduce different complications. Cash transfers can affect childbearing so that the composition of children exposed in treatment and control groups is affected by the treatment itself. In some cases it may be possible to restrict analysis to children who were already born (or conceived) when the cash

transfers were first announced. Alternatively, fertility responses can be directly incorporated as part of the analysis (Baird et al. 2019).

In nearly all studies, whether based on intensively tracked survey or census data there will be some remaining selection concerns. Therefore, thorough documentation and analysis of attrition, including testing for differential selective attrition and whether baseline variables interacted with treatment assignment predict attrition should be done. A general limitation for such analyses in long-term studies, however, is that there may be few individual-level baseline variables available. For example, for young children baseline measures of many of the long-term outcome variables of interest are typically not available or relevant, like measures related to the labor market (Ghanem et al. 2023). Therefore, in addition to the diagnostics described above attrition correction methods including bounding or inverse probability weighting should be considered. These methods can rely on strong assumptions so that consideration of their validity for the context at hand is essential. When assumptions are hard to substantiate, credibility can be enhanced by showing results are robust to different correction methods.

2.2 Power and absolute versus differential impacts

Another methodological challenge is that many impact evaluations, even though they were experimental, were not originally designed to measure long-term impacts. The objective of most first generation CCT evaluations was to provide rigorous evidence on short-term impacts as inputs into decisions about program

continuation or expansion; therefore, many had phase-in designs in which an initial control group later received the program (Colombia, Mexico, Nicaragua). The CCT in Honduras is an exception with municipal-level randomization and a control group that never received the program, allowing Molina-Millán et al. (2020) to estimate and compare absolute impacts across a wide range of ages using national census data. Phase-in designs, on the other hand, only permit estimation of differential impacts using the initial short-term treatment-control comparison and limit the statistical power of longer-term comparisons because after the control group is phased in identification is based only on differences in duration of program exposure. Interpreting the sign of such differential estimates as reflecting the sign of the absolute effects requires the assumption that the absolute impacts in the original treatment and the phased-in control groups go in the same direction. To address those limitations and increase power, long-term follow-ups of phase-in designs typically focus on cohorts for whom the experimental difference in exposure occurred at particularly sensitive ages. Even for this approach, however, interpreting estimates as informative about the magnitudes of absolute program impacts requires assumptions regarding sensitive ages that should be substantiated. Not all evaluations of important programs are experimental, and for some programs that were initially randomized, their later expansions were not. Non-experimental methods can be used to analyze these, but with pitfalls similar to those just described. Difference-in-difference (two-way fixed effects) studies relying on

geospatial and time variation in national rollouts of programs also only allow identification of differential impacts in the long term if all regions are eventually incorporated. Regression discontinuity designs (RDD) based on eligibility thresholds that can be used to estimate short-term impacts face related concerns in the long term since eligibility criteria can change over time and the same or similar criteria be used for eligibility for other social programs. For both these non-experimental approaches, the identifying assumptions can be hard to substantiate in the long term and should be discussed in detail.

2.3 Further measurement and inference challenges

A first additional challenge is that returns to human capital investments may only fully materialize at a later lifecycle stage. Therefore, for many outcomes of interest it is arguably just a matter of waiting long enough, but that may not always be feasible. When measurement happens during transitions across lifecycle stages it can make results harder to interpret. Many CCTs target educational investment in primary or secondary school. Even a follow-up 10 years after initial exposure could mean assessing outcomes while individuals are still transitioning to adulthood. Interpreting impacts during this transition can be complicated precisely because changes in human capital can lead to different trajectories. For example, children who remained in school longer because of the CCT might still be in school (but will ultimately attain higher education), might have only recently entered the labor market, or might even be combining schooling and work. Individuals in the control

might have transitioned to working earlier, possibly allowing them to currently earn higher wages because of longer labor market experience at that point, while at the same time being on a lower long-term wage trajectory. Such differential trajectories are hard to predict, however, because the opposite could also occur. The CCT could enable children to progress through schooling more rapidly, attaining higher education but not necessarily delaying labor market entry. Under either scenario, labor market outcomes are censored until both treatment and control have fully transitioned.

Related censoring is likely for other outcomes as well. Increased human capital can influence the marriage market, with differential educational and labor market trajectories affecting both the timing of marriage and the quality of the match. Therefore analysis on young adults before marriage decisions are completed in both treatment and control groups also has to consider possible censoring of outcomes. Parallel arguments apply to fertility decisions.

One work-around for these types of censoring is to collect data that allow defining variables in function of achieving an outcome by a certain age. If follow-up is done when individuals in both treatment and control groups have all reached or surpassed 18 years, for instance, by asking women whether and at what age they became pregnant researchers can unambiguously determine pregnancy by age 18 and estimate impacts on teen pregnancy. This is probably more informative than using age-at-first-pregnancy as an outcome variable, since the latter is undefined for

women who have not started childbearing and thus censored. In parallel fashion researchers can define completion of certain education levels by a given age (say high school by age 18), rather than completed grades which may not yet be final for everyone.

Of course when it is possible, another way around the challenge of censoring during transition into adulthood is to wait until all target individuals have reasonably completed the transition. For individuals exposed to cash transfers during school-going ages, this may mean measuring impacts after 20, rather than 10 years. For individuals who benefited from transfers during the earliest years of life, it means waiting even longer.

A second challenge is that cash transfers can also affect the speed at which children move through the different lifecycle phases, making it harder to measure some outcomes for comparable treatment and control groups. Continuing with the school progression example mentioned above, in many low- and middle-income countries (LMIC), late school entry, grade repetition and temporary school abandonment are common. Cash transfers can affect all of them. Consequently, the samples of children for whom there was administrative test data for a specific grade level would not be comparable between treatment and control groups. If exam difficulty varies by grade level, comparing outcomes across different grades (though possibly on more comparable treatment and control samples) likewise would be problematic. When the full educational trajectory is available for children the

importance of these concerns can be evaluated and accounted for, possibly by using bounds. Typically that would require administrative panel data at the child level followed through every grade.

A third challenge--specific to estimating the long-term returns from productive investments--is that they are more readily measured at the household- rather than the individual-level, but households change over time. Cash transfers can benefit both agricultural and non-agricultural household businesses, with households allocating the additional resources among various possible investments (Daidone et al. 2019). Determining long-term returns at the household-level is complicated when household composition changes over time, and depends on how the household panel is defined (i.e., which members from the original household are included), since surveys do not typically collect information about every original household member. Even for assessment of household-level returns, then, the challenge of individual mobility can affect long-term estimates.

A fourth challenge is related to understanding the sustainability of impacts after the program has ended. In many programs households or individuals remain eligible for many years but have to be recertified as eligible. Because recertification is usually a function of decisions taken by the household itself (e.g., related to school progression, fertility, migration or asset accumulation) there is typically no valid counterfactual for households that have graduated. Where eligibility is strictly

related to age, or program duration is fixed, however, impacts after graduation can more readily be assessed.

3. Conceptual framework: Pathways to long-term effects

Cash transfer programs designed to promote human capital accumulation typically target two life cycle stages, childhood (~0–6 years) and school-going ages (~7–17). For both age cohorts, transfers have been shown to lift constraints to human capital investment during the intervention period. Even without any other program-induced sustained shifts in investment, it is possible initial gains in human capital can be maintained and translate into better outcomes later in life. If in addition there is self-productivity or dynamic complementarities, the gains are potentially amplified (Heckman 2007). Further, if programs change preferences, aspirations or long-term prospects in ways that result in sustained shifts in investment, effects on human capital can be even larger. Impacts can differ for programs that affect educational attainment mainly by keeping children in school longer versus those that also influence learning, cognition or socio-emotional skills, or that influence nutritional status and health.

The literature on dynamic human capital production functions also demonstrates that the specific timing of interventions matters, with larger impacts when exposure occurs at more sensitive ages. Impacts can also fade out, for example in cohorts with high vulnerability who experience subsequent shocks or other external

constraints that limit the ability of their households to make investments exploiting dynamic complementarities.

Negative shocks or other external constraints can also limit the long-term returns to human capital in the labor market. One such external constraint is labor market competition. In many settings increased schooling for the poorest segments of the population resulting from cash transfers come at the same time as secular increases in schooling among the non-poor. As a result, the education increases for the poorest may not necessarily make them (relatively) more competitive in the overall labor market. On the other hand, human capital accumulation for large groups at the same time can have important social multipliers.

Cash transfers can affect human capital when households receive them during the early years of life improving several crucial inputs for early childhood development including nutrition, preventive health care, stimulation and the home environment. The transfers do so by relaxing liquidity constraints, as well as by providing information (through social marketing or soft conditions), targeting mothers and imposing targeted investments via (hard) conditions. The literature on early childhood development, however, demonstrates that although there are clear biological reasons why early improvements affect outcomes much later in life, including empirical proof-of-concept effects on adult economic well-being (Gertler et al. 2021), the later gains might not automatically materialize if additional complementary investments are needed. Similarly, even though the time from

investment during school-going ages to adult outcomes is shorter than for investment in early childhood, the substantial literature on returns-to-schooling demonstrates that large positive returns are not guaranteed. Higher education can affect economic and mobility decisions, but also fertility and household formation decisions, and all of these can be intertwined making causal pathways complex, particularly for women.

Another important consideration for disentangling the various possible pathways and drawing policy relevant conclusions is that while some constraints might remain constant over the 20+ years a researcher aims to evaluate long-term returns, others can change over time or space. For example, when individuals relocate, their external constraints change. Therefore, studying whether mobility is affected by exposure is essential. Where there is high mobility and migration, causal pathways can be even more complex.

In addition to operating through increased human capital, cash transfers can have long-term effects by temporarily lifting other constraints and enabling households to make productive investments that have long-term economic returns, particularly in settings with incomplete markets, and thus permanently relaxing prior constraints. Most directly, cash transfers can lift liquidity or insurance constraints. When transfers are regular, a predictable stream can also improve management of household consumption risk, and in doing so liberate resources for productive investment. Transfers can also relax information constraints if they allow

households to experiment and learn about different economic activities, possibly through changes in beliefs or aspirations, including those resulting from changes in mental health. Moreover, temporary exposure to transfers can lead to changes in social equilibria, for instance by shifting intrahousehold bargaining, or resolving collective action problems within extended families or communities.

Finally, whether and to what extent cash transfers operate through the different channels likely differs between interventions and contexts, as well as even between households exposed to the same intervention. Studying longer-term impacts, therefore, also means recognizing that more time has passed during which effect heterogeneity can amplify or diminish. Consequently, consideration of a wide variety of pathways probably needs to be the starting point. Where data on possible channels does not exist, there is a risk they will be overlooked. A detailed understanding of the context in which households make decisions—both at the start of the intervention and as it changes over time and space—can help unpack the possible pathways.

4. Evidence on human capital investments and adult returns

A growing literature examines the effects of cash transfers during early childhood (0–6) and school-going (7–17) ages on human capital and labor market outcomes 5–10 years later. Molina-Millán et al. review the long-term evidence for CCTs in 2019; since then, several additional studies have examined long-term effects of CCTs including Cahyadi et al. (2020), Parker and Vogl (2023), Barham et al. (2024)

and Gazeaud and Ricard (2024) and of UCTs including Handa et al. (2019). Consequently, evidence on effects into early adulthood is accumulating, particularly for individuals initially exposed to transfers at school ages. Because the first programs with rigorous evaluation designs only began in the late 1990s or later, however, understanding to date of the effects after individuals have fully transitioned into adulthood is more limited. Only one study—a 20-year follow-up of Progresa—documents effects beyond early adulthood (Araujo and Macours 2021). But even that study is unable to directly measure adult long-term effects of exposure that began in the earliest years of childhood. Figure 1 presents a summary of countries for which evidence is available based on age of intervention and time to follow-up, with references to the relevant studies.

Figure 1: Age at intervention and follow-up: Evidence available by country

	Lifecyle stage at intervention start		
	Early childhood (~0-6)	**School age (~7-17)**	**Young adulthood (~18-25)**
Time to follow-up			
5-10 years	*School age* -Colombia [34] -Ecuador [34] -El Salvador [34]	*School age & young adult* -Cambodia [34] -Colombia [4, 34] -Ecuador [34]	*Adult* -Mexico [37]

	-Honduras [36] -Indonesia [15] -Malawi [6] -Mexico [17,34] -Morocco [19] -Nicaragua [34]	-Honduras [36] -Indonesia [15] -Mexico [17,34] -Malawi [6] -Morocco [19] -Nicaragua [10] -Pakistan [34]	
11-15 years	*School age & young adult* -Honduras [36] -Mexico [37]	*Young adult & adult* -Honduras [36] -Mexico [37]	*Adult* -Mexico [37]
16-20 years	*Young adult* -Mexico [2]	*Adult* -Mexico [2]	

Note: The figure cites the Molina-Millán et al. (2019) review [34] for studies they cover and adds references to work published since.

Across the board, the long-term evidence that cash transfers lead to higher educational attainment (typically measured by years of completed schooling) is strong. There is also some evidence that children reach higher levels of schooling, including secondary or even tertiary (Colombia, Honduras, Indonesia, Mexico). Some long-term studies find persistent increases in enrollment, suggesting even larger educational gains will follow (Indonesia, Morocco, Nicaragua, Zambia). Unsurprisingly, magnitudes of estimated effects differ, because programs and

contexts are different and also because some research designs only allow identification of differential (rather than absolute) effects, for example between longer versus shorter exposure periods or between exposure at more versus less sensitive ages. Notably, final education achieved by CCT beneficiaries often surpasses the grade level directly targeted by eligibility criteria, suggesting the effects result in part from shifts in investment behavior beyond those imposed by the conditions alone.

Long-term educational gains are observed for cohorts exposed during school-going ages for whom education was directly targeted as described above, but also for cohorts exposed earlier in childhood. For individuals exposed during early childhood a plausible pathway is that nutrition and health investments improved early-life nutritional status, health and cognition, all of which can improve later schooling. A few CCT programs had positive long-term (5–10-year) effects of early childhood exposure on cognitive outcomes (Honduras, Nicaragua) and on anthropometrics (Indonesia). In contrast, findings for other CCTs point to fadeout of earlier positive short-term effects and/or catchup of control groups that received similar benefits a few years later (cognitive outcomes in Ecuador and Mexico; anthropometrics in Mexico and Nicaragua). In these cases, it is unclear whether differences could re-emerge at older ages. There is no positive long-term evidence on these channels for UCTs, however, probably because their short-term effects on investments in nutrition and health in early childhood are small (Fernald et al.

2024). Finally, the 20-year Progresa follow-up shows that there are clear and substantial gains for individuals for whom early childhood exposure began 18 months prior (but was equal during school ages), pointing to the potential importance of early life investments.

Long-term impacts from early-life exposure like those seen for Progresa could stem from investments in the earliest years enabling earlier school enrollment. Where programs are ongoing and differences in eligibility persist when children reach the age of school entry, it is not possible to fully disentangle whether effects are driven by the cumulative exposure to the CCT since early childhood, the start of the schooling conditionality when children reach school age, or a combination of both. Examination of programs that ended and are no longer operating by the time children reached school age, however, allow isolation of the impact of early childhood exposure and have demonstrated long-term gains (Honduras, Nicaragua).

Evidence on whether increased educational attainment translates into improved learning (as measured for example by achievement test scores) is more mixed. Long-term learning gains are documented for Colombia and Nicaragua, but not found for Cambodia, Malawi, Mexico or Morocco. The lack of effects could reflect limited potential in many contexts for increased schooling to improve learning, in line with broader evidence on the learning crisis in LMIC. At the same time, no improvement in average learning could reflect offsetting effects resulting from

increased class sizes related to positive enrollment effects or from possible negative peer effects related to changing classroom composition, for example if academically weaker children remain in school longer. A better understanding of the mixed long-term results for learning requires better understanding of the long-term effects on age-appropriate grade progression to characterize the overall effectiveness of the school system and, for studies using administrative test data, to help evaluate selection into grade-level exams.

There is also growing evidence on economic outcomes in adulthood, a key return to investments in human capital. Most studies investigate labor market measures during young adulthood when a relatively large share of youth is still in school (introducing censoring) or has only recently entered the labor market. Although this is an important transition period to examine, as described earlier drawing firm conclusions about ultimate labor market impacts based on examination of young adults is fraught. Araujo and Macours (2021) sidestep this concern by waiting until all have completed schooling, estimating positive effects of exposure during school-going ages on income for both adult men and women after 20 years. This finding is in line with estimates of labor market gains after 13 years in Parker an Vogl (2023). Both these results for Mexico suggest that prior estimates for effects of the program on young adults after six years may have underestimated labor market returns. In practice, many young adults in these contexts pursue schooling long after they turn 18 years old, in part because of earlier disruptions in their

schooling. In addition, as discussed earlier if the program increased grades attained, beneficiaries mechanically have less work experience than otherwise similar cohorts. This can reduce the net effect on returns when measured during young adulthood if returns to work experience are diminishing (i.e., if returns to the first few years of experience, which only those who have left school have, are relatively high).

For programs where the 5–10-year evidence suggests no clear gains in the labor market (Cambodia, Ecuador, Honduras, Malawi, Pakistan), longer-term follow-up may be needed to settle the debate. Similarly, for programs for which there are 10-year positive effects on young adults (Nicaragua), determining whether those accurately reflect labor market returns in later adulthood also requires further follow-up. The serial long-term studies of a different well-known intervention, the Jamaica early childhood stimulation intervention shows that labor market returns became larger as individuals went from their early 20s into their 30s (Gertler et al. 2014, 2021), raising the possibility of that returns for young adults are underestimates. In the meantime, for studies conducted during the transition to adulthood, collecting information on future income expectations, as Araujo and Macours (2021) do for those exposed during the earliest years of life, can provide important clues, since those expectations should reflect the existing constraints that individuals can predict.

Finally, a lack of long-term labor market impacts may also point to other binding constraints. Exposure to transfers during school ages is unlikely to help all beneficiaries circumvent the various constraints they face. Smaller program effects for the indigenous population in Honduras, for instance, suggests that constraints related to schooling may have been larger for them than for the non-indigenous population.

Two other channels by which cash transfers, through their effects on human capital, might influence constraints individuals face that affect long-term labor market returns include effects on 1) mobility and migration; and 2) fertility and family formation.

First, evidence on migration for individuals exposed during childhood is limited, even though in many settings domestic migration is common and often related to economic opportunities. In Nicaragua, temporary domestic migration related to labor market participation activities increased, but in Honduras there was no consistent pattern. Findings from both Honduras and Mexico, however, show positive and relatively large increases in international migration, especially for young men. More broadly for Mexico the 20-year follow-up shows school-age exposure increased migration for both men and women, including changes in both domestic and international migration. Moreover, their long-term gains in income can be linked directly to increased income in the US following international migration there.

Second, higher educational attainment can translate into different fertility and household formation patterns. Exposure at critical school-going ages in Mexico postponed childbearing and new household formation. In contrast, in Indonesia there were no effects on childbearing and in Malawi control group catch-up erased demonstrated short-term program effects. Cash transfers can also affect these outcomes in the long run by changing marriage market equilibria (Malawi), or because improved nutrition affects the age of menarche (Nicaragua).

5. Evidence on productive investments and returns

In addition to stimulating human capital investments, there is also a smaller literature examining whether regular cash transfer payments can trigger productive investments, consistent with the inseparability of consumption and production decisions in contexts with market imperfections. Under certain conditions, such productive investments can lead to important multiplier effects, or long-term increases in consumption beyond the value of the initial transfers.

For Progresa, Gertler et al. (2012) find that beneficiaries invested one-quarter of the transfers in productive assets, raising short-term agricultural incomes and increasing consumption after six years by 5% over and above the direct contribution of the transfers.[5] Handa et al. (2018) reach similar conclusions estimating the three-

[5] In line with those results, Parker and Vogl (2023) document improvements in household-level asset indices after 13 years.

year effects of two national programs providing regular and sizeable UCTs targeting vulnerable households with children in Zambia. Beneficiary households spent part of their transfers on asset accumulation, agricultural and livestock investments, resulting in livelihood diversification and a consumption multiplier of 1.67. A national CCT program with modestly sized transfers in Indonesia, on the other hand, did not affect asset accumulation nor did it have transformative effects on consumption (Cahyadi et al. 2020).

The long-term positive effects on consumption multipliers for Mexico and Zambia come from large scale national level programs that were ongoing at the time of follow-up. Therefore, the findings suggest the possibility of widespread and continued long-term impacts given the underlying asset accumulation and potential positive returns to those assets, but do not guarantee them. Recent work using machine learning techniques to detect heterogeneity of treatment effects on asset accumulation and consumption for four national UCTs in Africa including Zambia shows that the gains do not necessarily materialize for all program beneficiaries (Handa et al. 2023). Moreover despite evidence of substantial initial asset accumulation due to the program, estimation after the transfers ended suggests effects can fade out rapidly (Handa et al. 2019).

Most other evidence on long-term effects of regular cash transfers is based on examining fixed-period programs after they ended. To gauge potential for persistent longer-term effects, we focus on assessments two or more years after the

final transfer, since effects shortly after cash receipt ends could still reflect the immediate impacts of the transfers themselves, rather than returns to productive investments made with those transfers. Baird et al. (2019) find that positive productive effects of UCTs had largely dissipated two years after a program in Malawi ended. Haushofer and Shapiro (2018) show that regular transfers over a 9-month period, and separately a one-time lump sum cash transfer, led to more assets after three years in Kenya but find no conclusive corresponding effects on consumption.

Finally, a growing number of studies compare impacts of lump sum cash grants (often targeting young adults) with those of more complex intervention packages for benchmarking. A lump sum cash grant to 17–20-year-old women in Nairobi led to a sustained shift into self-employment six years after the transfer, but no long-term impact on income or economic well-being (Brudevold-Newman et al. 2023). A similar transfer, targeted to criminally involved young men in their mid-twenties in Liberia, had no significant impact on antisocial behavior or economic outcomes after 10 years, nor on forward looking behavior, time preferences or self-control (Blattman et al. 2023). And in Ethiopia, a short-term impact on occupation and income of a lump sum cash grant to (mostly female) job applicants in their early twenties had dissipated five years later, also with no impact on health or mental health (Blattman et al. 2022). These results suggest that one-off modest lump-sum

transfers may not be sufficient to put young adults on a fundamentally different trajectory.

Comparing the above results with the stronger long-term impacts of regular cash transfers raises the question as to whether differences in results are due to differences in design or in the targeted populations. An ongoing study by Banerjee et al. (2023) experimentally varying lump sum transfers, regular transfers of the same total amount for a two-year period, and regular transfers for a 12-year period (essentially universal basic income), will allow assessment of the relevance of the regularity and of long and predictable streams of cash transfers. Related, Macours et al. (2022) show that a one-year program combining regular cash transfers conditional on human capital investments with a lump sum transfer for business investment, compared to the program without the lump sum transfer, enabled households to protect against shocks through income diversification and led to sustained consumption increases two years later.

6. Discussion and open questions

Taken together, the literature provides rigorous evidence that cash transfers often lead to long-term human and physical capital accumulation, which under the right conditions can lead to increases in income, consumption and other indicators of well-being. Apart from their direct effect on human and physical capital, cash transfers can also lead to sustained changes in investment behavior through shifts in internal constraints or preferences, or in intrahousehold bargaining. In addition,

where transfers are conditional, the conditions contain information value that households may continue to act on in the longer term. Molina-Millán et al. (2020), for instance, document that there were substantial long-term effects on cohorts born after the program ended, pointing to shifts in investments that lasted well beyond the duration of the transfers. More evidence is needed to help understand when and to what extent such sustained shifts in investment occur.

Many open questions remain on labor market and life trajectories as the exposed cohorts age. Dynamic information on job quality, industry and location could be important for understanding possible tradeoffs between pay and quality of entry jobs, and their relationship with longer-term employment outcomes. More information on savings, loans and investments in productive assets is needed to better understand how earnings are being invested for future income generation activities. Information on non-labor market outcomes, including mental health, stress, violence, social connections and networks could help trace the key pathways and trajectories.

Some studies examine intermediate or long-term effects on aspirations, risk aversion, time preferences and mental health (e.g., Handa et al. 2020), but more work is needed to disentangle when and how designs can be adapted to trigger changes in those outcomes that then translate into longer-term gains. Macours and Vakis (2020) show that exogenous variation in the exposure to female leaders that randomly received the largest benefit package increased aspirations of the other

beneficiaries and led to higher human capital and productive investments two years after transfers ended. Two meta-analyses on the impact of cash transfers on mental health and happiness, however, show only modest initial effects that subsequently fade out (Romero et al. 2021; McGuire et al. 2022), indicating longer-term effects do not automatically occur.

There is a larger set of open questions regarding effects of UCTs on longer-term well-being. This includes understanding impacts on children's human capital and skill development, but also impacts on social and behavioral outcomes that affect the communities in which beneficiary children grow up, including community engagement, criminal activity and changes in family structure.

A further area for research is how cash transfers affect the children of the individuals exposed during childhood. Evidence on this 2nd generation question is thin (Baird et al. 2019; Barham et al. 2023) in part because the first generation is still fairly young in most countries. If dynamic complementarities exist across generations, parents who benefited from programs should invest more in their children's human capital, potentially leading to multiplier effects on intergenerational human capital. However, if the programs led parents to migrate for work or to work more demanding jobs, parents may also spend less time with their children or be absent for extended periods, possibly negatively affecting human capital outcomes. As the 2nd generation pathways are arguably even more complex, involving both behavioral and biological mechanisms, this research

agenda would benefit greatly from multidisciplinary collaborations and from systematic efforts to collect data at regular intervals throughout childhood, and on large enough samples to disaggregate pathways by gender or other relevant dimensions.

Investigating the above questions will provide a fuller accounting of the long-term effects of cash transfers, but a final set of open questions relates to program design. For example, do short-term differences in impact from alternative program designs (such as targeting mothers versus fathers or having hard versus soft conditions) translate into different long-term impacts? Also, do programs lose their effectiveness over time? This could occur because initial staff motivation or institutional support wanes or because beneficiary responses to the intervention, even if it is implemented in the same fashion, can change over time, for instance because the economic context for which a program was initially designed changes for reasons unrelated to the intervention. More generally, in rapidly changing contexts, long-term returns from a program implemented 20 years ago may have more limited relevance for informing policy design today. With only a few studies tackling such questions (notably Cahyadi et al. 2020), this is an obvious area that deserves more research. Evidence on these questions can provide crucial feedback for the design of new programs.

Author Acknowledgements: We thank Arthur Alik-Lagrange, Rema Hanna, Benjamin A. Olken, an anonymous reviewer and participants at the Social Protection Handbook Conference for comments. All remaining errors and omissions are our own.

A full set of references for all Handbook chapters, including this one, can be found at this link: https://www.dropbox.com/scl/fi/9lqs2mdrawkjdrv4m648e/References-Social-Protection-Handbook.pdf?rlkey=jt0f8kute31mhdke77aoiw99d&st=kd7l8ff1&dl=0