

Published on **Development Impact**

Who benefits from the indirect effects of cash transfer programs?

BERK ÖZLER | OCTOBER 16, 2024

This page in: **English** 



Surprise! I have been thinking about cash transfer programs and spillover effects recently...

GiveWell is trying to update their recommendations, and they are specifically trying to update their priors on the cost-effectiveness of cash transfers, based on the relatively recent Egger et al. (2022) paper, which evaluated a large-scale, one-time cash transfer program by GiveDirectly (GD), providing close to PPP\$2,000 (\$1,000 nominal) per beneficiary household (HH), which accounted for approximately three quarters of total HH expenditure. At its peak, the program accounted for about 15%

of the GDP of the area it covered, which implies (by my crude calculations) that about one-fifth of the households in program areas were beneficiaries (GD targeted HHs that lived under thatched roofs). As that paper documented large positive spillovers to non-beneficiaries in target areas (with only a small amount of precisely measured price inflation); estimated (less precisely) a multiplier effect of 2.5; and differed from previous studies methodologically, GiveWell wanted to know how they should update their thinking on the basis of this new study with some striking headlines.

You will soon be able to read their assessment online. But, as I had to read this paper carefully and tried to reconcile it with previous studies (of GD cash transfers, as well as others that found spillovers, price effects, etc.), I thought that it might be worth sharing a few things that occurred to me with our readers.

Different spillovers for different people? Let's go back in time...

One of the things that is happening now is that people are questioning the findings (and relevance) of the earlier GD evaluations in light of the more recent one by Egger et al. (2022). This is natural and reassess, we should. Here is a list of blog posts (from which you can link to more) to jog your memory about the longer-term (three-year) findings from the original GD evaluation, from 2018:

- [***https://blogs.worldbank.org/en/impactevaluations/dear-governments-want-help-poor-and-transform-your-economy-hold-recalculating***](https://blogs.worldbank.org/en/impactevaluations/dear-governments-want-help-poor-and-transform-your-economy-hold-recalculating)
- [***https://blogs.worldbank.org/en/impactevaluations/givedirectly-three-year-impacts-explained***](https://blogs.worldbank.org/en/impactevaluations/givedirectly-three-year-impacts-explained)
- [***https://blogs.worldbank.org/en/impactevaluations/givedirectly-three-year-impacts-explained-authors***](https://blogs.worldbank.org/en/impactevaluations/givedirectly-three-year-impacts-explained-authors)
- [***https://blogs.worldbank.org/en/impactevaluations/evidence-based-or-interpretation-based***](https://blogs.worldbank.org/en/impactevaluations/evidence-based-or-interpretation-based)

If you were to skim these posts, especially the last one, you will note that a contentious for issue that original study, let's call it HS18 (for Haushofer and Shapiro, 2018), was the presence of negative spillovers. One objection, which I thought was misguided to say the least, came from a respected CGD blog post: it argued that if the spillover effects from treating 10% of a village's population were real, they must be the same for everyone. Using this logic, they extrapolated the estimated negative effects to the entire population in the study areas and concluded: *"Unless cash*

recipients literally spent the money on gasoline to set fire to their neighbors' farms, the scope of negative spillovers required to explain the Kenya results seems implausible."

Well...based on Egger et al. (2022), that post did not age well. Where did it go wrong? It's a little technical, but easy to understand: come with me...

The study design in HS18 was such that eligible HHs in treatment areas were randomly selected into (or out of) the cash transfer program. In other words, there were both eligible (about 10% of a treated village's population, on average) and ineligible HHs (about 80%), who did **not** receive any transfers. As I explained in excruciating detail in the blog posts above more than six years ago, HS18 did **not** collect any data on ineligible HHs. So, the spillovers they estimated were on eligible HHs only, i.e., on less than 10% of the population that looked identical to those who did receive cash. The CGD blogger assumed that whatever spillovers were there on this group must be identical for the remaining 80% who were ineligible.

It was this assumption that grated on me. It was not only facts **not** in evidence, but it also seemed unrealistic: why would spillovers be the same for the poor and the non-poor (using these words loosely to refer to the GD eligibility criteria, instead of eligible/ineligible)?

The study design in Egger et al. (2022) is different from HS18 (and HS16) in a number of ways. One of those ways, which is most welcome, is the fact that all eligible HHs within a village area treated if that village is chosen for treatment (T). So, there are no spillovers measured for eligibles in T (spillovers might be present for them, in addition to the direct benefits, but they cannot be measured with this study design. See [this paper](#) for details.) Therefore, **any spillovers reported in T are on ineligible HHs only**. There is a second difference between HS18 and Egger et al. (2022): the intensity of treatment in larger geographic areas (called sublocations) was manipulated by randomly varying the number (share) of villages assigned to T or C (control) in each sublocation. This means that, using the random distance between T & C, spillovers to nearby villages can also be estimated – this time both for eligible and ineligible HHs. It is these estimates that will prove most helpful in adjudicating the plausibility of the assumption that spillover effects of large cash transfers should be similar for everyone, regardless of eligibility status...

What spillover effects did Egger et al. (2022) find?

Egger et al. don't make it easy to find the effects on non-recipient HHs disaggregated by their eligibility status (who knows, it could have been an editor who made them move it out of the text). Table 1 in the main paper shows large and statistically

significant increases in HH expenditure (specifically, non-durable expenditures). These are pooled across all non-recipients, i.e., ineligible in T, as well as eligibles and ineligible in C. For those, who are curious enough to download the [Online Appendix](#) and scroll down to Table B8, we get the disaggregated view: spillover effect estimates, coefficient estimate (standard error), on eligibles in C are \$21 (84) for annualized HH expenditures, i.e., practically zero. The same effects are \$412 (148) on ineligible HHs in T & C combined! *[If I am not mistaken, Givewell has used the raw data obtained from the authors to disaggregate the spillovers to ineligible HHs further between T & C and found that they are similar to each other, albeit with an expected negative gradient as distance increases, meaning that spillover effects on ineligibles are slightly higher within-villages than between them].* **Bottom line:** it is clear that the spillovers of large cash transfers to approximately a fifth of all HHs in T accrued almost entirely to ineligible HHs in nearby areas. If you were extrapolating the effects to ineligibles using what happened to eligibles, the multiplier would have ended up close to 1, instead of 2.5...

There is a lesson or three in here: first, if you're consuming news about spillovers (especially, if you're using it to make important financial or policy decisions), you should ask whether these effects are heterogeneous and try to figure out what they are for subgroups you care about. Second, we should avoid speculating unnecessarily and confidently when we don't have the data. Third, the spillover effects estimated in HS18 do not look so out of place in light of these new findings, once we disaggregate spillover effects to compare like with like: it seems possible to have zero or even small negative spillover effects on eligible HHs, while having large positive ones on ineligible ones.

That last one of these is a good puzzle, perhaps worthy of a new research project on its own: what is it about eligibility (especially one that is based on a seemingly blunt targeting tool, like thatched roofs) that causes this massive heterogeneity? One explanation is that ineligible HHs are much more likely to own small businesses, which seemed to benefit from the demand shock caused by the one-time transfers. They seem to have done so by reducing the slack in their inputs. Eligible HHs may have had small wage gains through increased (albeit perhaps temporary, in the absence of large increases in savings and investment) demand, but clearly not enough to make a dent in their HH expenditures and vastly smaller than that for ineligible HHs. Since ineligible HHs, which make up 80% of the population, may themselves be a heterogeneous group, likely hiding some households that are more similar to eligible HHs, figuring out the mechanisms underlying the heterogeneity of

spillover effects seems like a question of primary importance – to equitably maximize overall program impacts. Please let me know if you design a clever study to get at this (probably best to start with the data from Egger et al. to see what can be gleaned from it before starting a brand-new study).

We make progress by generating new evidence, updating our priors, and sometimes completely changing our minds about an important issue. Egger et al. (2022) provides an important new data point and, as such, obliges us to update our priors about spillover effects. But it also differs in a number of important ways from previous studies (including evaluations of GD transfers and other CT programs that examined spillover effects). Hence, my view is that it should rightly take its place among them and be recognized as important, rather than displacing everything that came before it. There will be some, including researchers/experts in the field, who disagree with this stance and want to more or less dismiss the earlier HS16 and 18 results as now being debunked. To this group, the randomized saturation design employed in Egger et al. (2022) exposed the methodological shortcomings of the earlier HS work, specifically that it did not account for spillovers across villages. Reasonable people can disagree on how much we should discount previous studies and how we should update our priors on the overall impacts of cash transfers (and you can see GiveWell's take, which incorporates the opinions of a number of academics, when it comes out in the next few weeks). But we should be careful to remain even keeled, examine all the evidence, and do our hardest to not be influenced in our interpretations of the evidence by motivated thinking, hype, or other similarly irrelevant factors: we can't cheer for and parade the evidence when the findings are good and ignore or, worse, dismiss them when they're null or negative.

A thought experiment

Let's finish this post with a thought experiment about a donor, perhaps very much like yourself, who is trying to decide between two hypothetical charities A and B. Charity A increases the wellbeing of its beneficiaries more than B does theirs (say, by X per person, X being whatever metric you're using to rank charities, such as DALYs or income or else). For simplicity, let's say the target populations of A and B are identical in each case and, without loss of generality, the share of eligible and ineligible HHs is equal at 50%. However, B benefits non-beneficiaries more than A (say, by Y per person). There is no information on heterogeneity of effects. How should the donor allocate their charity dollars between A and B?

If you were thinking along Rawlsian lines, worrying only about the poorest people among us, you might allocate little to no funds to B. With a convex function underlying your welfare judgments, the weight you'd put on non-beneficiaries would also decline rapidly, causing you to give more of your money to A, as long as X is not $\ll Y$. With equal weights on everyone, you'd give it to B as long as $Y > X$. And so on... The decision would likely vary across different settings for the same decisionmaker: in poorer settings with low inequality, you're more likely to care about ineligible HHs, whereas the weight attached to them might go down in richer and more unequal settings. When organizations like GiveWell give you rankings of, say, top charities, they are making these judgment calls for you. In some sense, that's what you want – either because you can't adjudicate all the evidence, or you don't have the time or the willingness to do so. But questions like the one above are key considerations in decision-making: overall program impacts are a combination of direct and indirect effects. The latter accrue to people not targeted by the program. Your judgement will therefore depend on (a) accounting for all the effects and (b) how much you care about people that the program is not targeting.

SOCIAL PROTECTION

Get updates from Development Impact

Enter your email to receive email notifications

Language: English

* I agree with the terms of the Privacy Notice and consent to my personal data being processed, to the extent necessary, to receive these updates.

SUBSCRIBE



Berk Özler

Lead Economist, Development Research Group, World Bank

MORE BLOGS BY BERK



Join the Conversation

Your name

Your Email

The content of this field is kept private and will not be shown publicly

Write a response...

Remaining characters: **1000**

- I have read the [Privacy Notice](#) and consent to my personal data being processed, to the extent necessary, to submit my comment for moderation. I also consent to having my name published.

SAVE

