High-Frequency Engel and Supply Curves in General Equilibrium: Experimental Evidence from Large Universal Cash Transfers in Malawi

Dennis Egger, Tilman Graff, Miriam Laker, John Walker

Executive summary

This project will study the largest Randomized Controlled Trial (RCT) implemented to date by the NGO GiveDirectly (GD). GD will provide universal unconditional cash transfers of USD 550 (125% of annual household income) to 110k adults in Malawi's Chiradzulu district, representing 50% of the population and 60% of district GDP. This program represents the next phase in GD's ambition to scale up to treating entire countries over the next decade.

Our study aims to understand the effect of transfers at this scale and over time. On the supply side, does the demand shock cause disruptions or inflation? Or does it increase trader entry, mobility, imports, investment, hiring, or a reduction in excess capacity? On the demand side, do consumers re-allocate their spending to markets with lower inflation?

To answer these questions, cash transfers will be randomized at different geographic levels, generating variation in cash flows into markets. High-frequency market surveys will track expenditure, inflation, and market activity, while monthly surveys will monitor recipient spending patterns. Matching data across these surveys, we will construct short-run supply curves at the firm and market levels offering insight into supply impacts.

We will then test whether providing information to a randomly selected group of traders about future increased demand for their products reduces inflationary pressure. The market surveys will track price effects, and we track traders' responses using a newly developed app at daily or weekly frequency. This will provide insight into how microenterprises respond to demand uncertainty, with implications for macroeconomic dynamics in low income countries.

Research team

The research team has ample experience with large-scale cash transfer RCTs. <u>Dennis Egger</u> has worked on GiveDirectly's largest cash transfer project to date in Kenya since 2016. This was the first experiment to study spillovers and general equilibrium impacts of cash, and the first paper is now published in Econometrica (<u>Egger et al. 2022</u>). The proposed project is an order of magnitude larger, and proposed much higher-frequency data collection, allowing us to shed additional light on how local markets adjust, whether the limited inflationary impacts (and large multipliers) in Kenya generalise to a new setting, and much larger transfers,

Tilman Graff is a PhD student in economics at Harvard and has worked on the Kenya project for several years. Importantly, his work on the structural model behind the paper entitled 'Slack and Economic Development' (joint with Dennis Egger, Edward Miguel, Felix Samy Soliman, Nachiket Shah, and Michael Walker) is one of the key motivations behind this new study. The model shows how the large 2.4 multiplier from cash with limited inflation found in the Kenya study can be rationalised through slack (underutilization of labor and capital) in small firms. The paper also suggests, however, that multipliers of larger transfers in different contexts may be smaller. The proposed study will test this, as well as shed light on how relatively cost-effective ex-ante information to traders may help smooth out supply shortages and inflationary pressures.

Miriam Laker is the Research Director of GiveDirectly, and has been involved in <u>several RCTs and</u> <u>evaluations</u> that GiveDirectly has run over the years. GD has experience with running large-scale lump-sum cash transfer programs across 10 countries in Africa, including in <u>Malawi</u>.

John Walker is a DPhil student in Economics at Oxford University. They have previously worked as a pre-doc at the Centre for the Study of African Economies. In that role, they analysed a <u>large cash</u> transfer RCT (R&R at Restud). Their recent work tries to identify households that will receive little benefit from cash transfers given their characteristics (<u>AEA P&P presentation, 2025</u>). They have previously worked on <u>simulating the effects of cash transfer programmes</u> at scale.

Intervention(s)

The intervention will be a universal (untargeted), unconditional cash transfer program delivered by the NGO GiveDirectly (GD) to adults in Malawi. GiveDirectly's intervention is highly scalable and efficient. GD's cash program has been used by GiveWell and others as a 'benchmark' intervention against which the impact of a dollar spent on other interventions is measured. In its recent update, GiveWell has scored GD's program as 3-4x as effective as previously thought, primarily due to longer-term consumption impacts, spillovers in the local economy, and health benefits (in particular, child mortality). So, tongue-in-cheek, a dollar is worth more than a dollar in the longer term.

However, GiveWell notes that important open questions remain, particularly on i) the persistence of consumption gains to recipient households; ii) the magnitude of consumption gains to nearby non-recipient households; iii) how much we ought to value increasing consumption vs. averting deaths or improving other health outcomes.

This study is designed to address some of these gaps. It will allow us to a) understand how the large spillover effects from Kenya generalise to *universal* (as opposed to targeted in Kenya) transfers, much larger transfers (125% of GDP as opposed to 15% in Kenya), and much poorer settings that are less integrated in trade; b) whether the inflationary impacts remain low, or whether they can be muted by very low-cost information treatments.

More broadly – this project is only the first in a much longer-term research agenda in Malawi. It will be set up to be further scaled with future expansions of the program, and to address other open questions, including c) the long-term impacts of cash at scale, on individuals as well as entire regions, and d) the impacts on health and mortality, and e) the intergenerational impacts on children at birth (child mortality), through education and human capital investments, and finally through increased educational attainment and labor market returns.

There are also efforts underway to test different targeting strategies, and complementary interventions to cash (including. community grants, climate-smart agriculture interventions, additional cash or insurance for widespread shocks, labeled cash - unconditional cash with nudges to promote specific behaviors), complementary infrastructure, or health or community investments to further improve the cost-effectiveness of cash transfers. By supporting the early setup of a research infrastructure, GiveWell would also support these possible future projects.

Research design

This is a large-scale randomised rollout of cash transfers across the entire Chiradzulu District in Malawi. The district comprises 10 Traditional Authorities (TAs), 111 Group village Headman units (GVHs), 934 villages, 90k rural households, 220k adults, 450k population.

Randomization of cash: Randomization will be across 3 different geographic units (villages, GVHs, and TAs) in order to create variation in exposure to treatment not only at the village level, but also at

higher geographic orders, allowing us to estimate treatment effect spillovers across larger geographies.

- Saturation: Variation in the intensity of treatment across TAs and GVHs generates variation in the geographic exposure to treatment of local economies:
 5 high-saturation TAs: ²/₃ of GVH treated – 37 treated GVHs, 19 control GVHs
 5 low-saturation TAs: ¹/₃ of GVH treated – 18 treated GVHs, 37 control GVHs
- Village-level treatment: Within treated GVHs, approximately 50% of villages will be treated. Retaining pure control villages in treated GVHs allows us to estimate within GVH spillovers. Within treated villages, treatment is *universal* and at the individual level, i.e. every adult will be eligible to receive treatment.

This design generates substantial variation in treatment exposure at various different geographies: Treatment will amount to 125% of GDP in treated villages, 60% of GDP in treated GVH, and 30% of GDP in Chiradzulu as a whole. This is substantially larger than any previous cash transfer RCT. For comparison, the Egger et al. (2022) study had 25% of GDP in treated villages, and only 5% of GDP across their study area Siaya County (an area roughly the same size as Chiradzulu district). Beyond 2025, the plan is to enrol additional districts with a very similar design, adding additional clusters and sample size to detect even larger spillovers at the TA level.

Taken together, this study will have a larger sample, larger number of clusters, and a larger amount of cash flowing into the local economy than any previous study, and allow for estimation of the spillover effects of transfers at an even larger scale.

Randomization of demand information: There are approximately 24 'official' markets, and 48 unofficial markets in Chiradzulu district, a total of 72 markets. Markets operate at least weekly, and have an average of about 350 businesses present on a given market day, a total of 25k enterprises and traders. The above cash transfer randomization creates substantial variation in markets' and sellers' exposure to increased demand and spending from cash. This comes from geographic variation (some market catchment areas get up to 125% of GDP, while others get almost 0%), and from variation in spending impacts on different products (in earlier work, for example, food spending increased by 70%, while non-food spending increased by 300%). Traders will thus face very large random differences in demand shocks, from 0 to 300%, allowing us to study the impacts of demand shocks on firms' responses in output, inputs, imports, investment, slack, and prices.

To mute potential inflationary impacts, we will cross-randomize ex-ante information about upcoming demand shock (estimated from household level spending responses), again at two levels:

- Saturation:
 36 high-saturation markets: ²/₃ of enterprises treated
 36 low-saturation markets: ¹/₃ of enterprises treated
- 2. Enterprise-level treatment: Treated enterprises, we inform traders upfront (randomly either 1 month, or 1 week) about upcoming expected increases in demand at this market. We get this information from estimated spending impacts at the household level, so these will be highly accurate on average.

Estimation Strategy: We will estimate impacts of cash transfers on a) household outcomes, b) enterprise outcomes, and c) local economy outcomes, including GDP and prices / inflation. Additionally, we will estimate the impacts of providing ex-ante demand information to businesses on their anticipatory behavior, including on inventories, investment, and prices.

GiveWell / Open Philanthropy 2024 Research Proposal

- a) Household outcomes: To estimate direct impacts on households over time, we will compare treated and control households. To estimate the spillover impacts on other households within the same GVH (TA), we will compare control households between treated (high-saturation) and control (low-saturation) GVHs (TAs). This leverages the clustered randomization design. Note that the variation in treatment intensity, (and the size of the randomization clusters and therefore the spatial decay of the spillovers we can estimate) is larger than in any previous cash transfer RCT.
- b) Enterprise outcomes: To estimate impacts of the cash-induced demand shock, we compare enterprises in markets (and sectors) randomly more or less exposed to increased spending. We measure exposure using baseline shopping patterns among treated and control households, and the estimated increase in spending from high-frequency expenditure surveys among households.

To estimate the direct effects of informing traders ex-ante about future demand increases, we compare treated and control enterprises. To estimate the market-level spillover effects of such information on other traders, through competition, information sharing or prices, we compare control enterprises in high- vs. low-saturation markets.

c) Local economy outcomes (market activity, GDP, prices): To estimate impacts on market activity such as overall GDP or prices, we use a similar exposure-based strategy to what we use for enterprises. We compare markets more or less exposed to cash due to our clustered randomization strategy. We may additionally collect other outcomes, such as data from local health providers and clinics, and estimate impacts on those based on baseline exposure.

The empirical strategy and data collection are designed to address two cross-cutting objectives of the study on a variety of outcomes.

- 1. We will be able to measure *long-term impacts (including inter-generational impacts on children).* This is where evidence is currently scarcest, and where new evidence would have the biggest impacts on GiveWell's assessment in terms of cost-effectiveness of cash, or its role as a benchmark.
- 2. We set up the trial to be able to estimate *heterogeneous impacts of cash transfers* including by baseline economic and socio-demographic characteristics by timing of transfer receipt, and by features of the local economy and region. Transfers are universal, creating substantial variation in the type of households that are reached. Together with a large sample size, this will allow us to characterize, for example, *which individuals, which areas, or at what point during the agricultural season benefits from cash are highest*, and which types of targeting approaches maximize the cost-effectiveness of cash transfers in the future, taking into account both the heterogeneous impacts of cash, and the costs (and errors) or targeting. Gender is a core focus of this analysis, both for how household spend transfers and for how effects translate to children and other household members, including through marriage, fertility, and child investments. To estimate heterogeneous impacts, we will rely both on pre-specified and ex-ante interesting dimensions of heterogeneity as well as on novel machine-learning based algorithms for classifying households into distinct groups by predicted impact.

Data collection

1. *Household data:* We will conduct, first, a baseline household census capturing all approximately 90k households within Chiradzulu district, including their location and basic demographics. We will then draw two samples:

- a. *Full sample (12k households):* These households will be administered a long-form baseline and endline survey approximately 1 year after receiving cash transfers (and possibly more rounds thereafter). These surveys will last 2.5 hours and collect data on demographics, health, education, assets, consumption, income and economic activities, social networks, political engagement, mental health and wellbeing, etc. This data will be used to estimate detailed medium- and long-term impacts of cash transfers. We plan surveys after 1, 3, 7, and 10 years.
- b. High-frequency sample (4k households): This subset of households will get short 30min monthly surveys by phone and in-person collecting data on key short-term outcomes, primarily on consumption and spending (by product and geography), labor supply, wages, inflation perceptions, etc. This data will be used to estimate high-frequency spending impacts, consumer switching across markets, supply disruptions, and to estimate market-by-product level demand shocks. We will carefully manage attrition through a combination of phone and in-person encouragements and incentives, which we have trialled extensively in a previous study in Malawi during 2023/24.
- 2. Enterprise data: We will first conduct a baseline enterprise census (jointly with the household census) capturing all enterprises operating in villages and market centers. A large majority of these enterprises will be operating in the 72 market centers or their catchment areas.
 - a. High-frequency and information experiment sample (4,500k enterprises): We will do a longer 45min baseline and endline in-person survey with these enterprises. At baseline, we will distribute smart phones to these enterprises (unless they wish to keep their existing phone), and enlist them into our information experiment. We will then collect extra high-frequency data through a newly developed mobile app, which asks them to report daily location movements, customers, sales, profits, hours, input purchases, etc. in a game-ified and incentivized (remunerated) way. To further incentivize reporting, we will provide monthly reports on sales and profits, and other business insights to these participants. This will ensure low attrition, and high data quality over time.
 - b. Monthly market census and surveys (72 markets, 2160 enterprises): To capture market activity, we will visit each of the 72 market centers monthly, and conduct a full count of sellers, their type, GPS, and ownership information. With a subset of 30 sellers per market, we will then collect a larger 30min survey to capture information on customers, pricing, revenues, profits, labor and capital, investment, input purchases, slack or utilization, etc.
- **3.** *Price data:* We will visit each of the 72 markets monthly and collect 2-3 price quotes for approx. 100 goods and services, including wages and services, collected as inconspicuously as possible, and on the market day/busiest time of the market day

Primary Outcomes

At the *household level*, we are interested primarily in poverty reduction / welfare broadly construed as the ultimate goal. Primary outcomes will include the high-frequency profile of spending, including for immediate consumption and investment (into assets, businesses, human capital, and children). This will enable us to speak directly to GiveWell's main measures of effectiveness and welfare, as well as how they may change over time, and in the long-term. Expenditure impacts will also serve as the main shock and input into estimating market-level supply and price responses to cash. In addition, we will collect data on health (augmented, ideally, with clinic and health-post data) and mental health in order to approximate DALYs and QALYs. A key focus will be intergenerational impacts on children of recipients, and we will measure investments into children (nutrition, health, education, and child labor) as well as detailed child outcomes over time.

For *enterprises,* our key measures are entry (i.e. whether cash induces more people to start businesses, or businesses to enter more or different markets), as well as changes in output within-firm. We will decompose this into changes in inputs (hiring of workers and hours worked and investment into new machines or capital), intermediate inputs (such as the import of supplies and materials), and increases in productivity, e.g. through investment, education, or the reduction in idle capacity. A key focus here will be geography – at what level do markets (and firms within them) adjust to demand increases, and which factors determine whether this leads to large multipliers (and low inflation), vs. inflationary pressures or supply shortages. And, lastly, whether information provision can smooth some of these pressures.

At the *local economy level*, we are interested in overall real economic activity, or real GDP, which is a function of both increased demand and supply net of any price changes. Our key outcomes therefore are the impacts of cash on GDP over time (the multiplier), and on inflation.

Power Calculations

Below, we present power calculations for three key outcomes, one at the household level (consumption), one at the enterprise level (profits), and one at the market-level (prices). We assume a significance level of 0.05 and power of 0.8 as is standard. The baseline mean, standard deviations, and intra-cluster correlations for each variable (either at the village or GVH level) are drawn directly from Egger et al. (2022). Sample sizes are as described under 'data collection' above. We compute power for several different estimands, including direct effects, indirect effects, and heterogeneity.

Variable	Effect	Cluster	no. of clusters	Ν	BL Mean (SD)	ICC	MDE in %
Consumption	Direct - Full sample	village	1000	12000	2500 (2000)	0.02	4.5%
	Direct - High-Frequency	village	1000	4000	2500 (2000)	0.02	7.3%
	Direct - Heterogeneity	village	1000	6000	2500 (2000)	0.02	6.1%
	GVH-level Spillover	GVH	111	12000	2500 (2000)	0.007	5.4%
Firm profits	Full sample	Market	72	4500	800 (1250)	0.02	19%
	Monthly market survey	Market	72	2160	800 (1250)	0.02	24%
	Spillover to control	Market	72	2250	800 (1250)	0.02	23%
Log prices	All months	Market	72	864	4.1	0.08	1%
	One month	Market	72	72	4.1	n/a	3%

With the proposed sample sizes, we can detect direct consumption impacts of approximately 5% to 7.3%, even when looking at heterogeneity across groups. For spillovers, we can detect a 5.4% increase in consumption for the control group.

For enterprise profits, both at high-frequency and in the long term as well as for spillovers, we can detect impacts of around 20%. This relatively high MDE is largely driven by how noisy estimated profits tend to be. However, in previous work, we have also found that profit measures tend to be highly correlated over time, so that controlling for baseline profits and revenues is a very promising strategy for increasing power. This is why a thorough baseline survey is necessary. We believe that for expected increases in demand between 0 - 300%, an MDE of 20% for profits or revenue is still economically useful and precise enough.

For prices, we will be able to detect an increase of about 3% in any given month, and an average increase of 1% over the course of 12 months.

Impact of research

As mentioned above, GiveWell uses GiveDirectly's cash transfer program as a benchmark for rating other charities. In its recent update, it estimates that GD's program is 3-4x as cost-effective as previously thought, primarily due to longer-term impacts, spillovers, and health impacts on children. This study will address many of GiveWell's key uncertainties in this space, helping Givewell further improve its model of GD's and cash transfers' impact in a range of dimensions (see arguments above). This is also a chance to support a much longer-term research agenda that will feed into GD's proposed nationwide scale-up, inform the future roll-out of cash transfer programs, and inform policy around cash transfers as a development tool. There will be additional opportunities for adding outcomes (e.g. data on child mortality and health from nearby clinics), or randomize additional aspects that may improve cash's effectiveness overall, including for instance, complementary demand or supply-side intervention such as the information provision to businesses studied herein, but also Al-based agricultural extension, community funds, health or education investments; or new approaches to targeting cash (e.g. during the lean season, for pregnant women, at low-utilization areas, etc.).

This research will also directly inform the effectiveness of GiveDirectly's planned scale-up within Malawi (and elsewhere) over the next decade and beyond. The learnings from this study are embedded in a long-term research partnership between the Oxford Centre for Macro-Experimental development and GD. GD is committed to build research and evaluation into each phase, and incorporate research findings into next phases to maximize impact and further improve the cost-effectiveness of cash. Concretely, should ex-ante trader information lower inflationary pressure or supply disruption, this may become common practice within GDs program. Should particular targeting strategies or complementary cash+ interventions prove to be much more effective, these may be rolled out at a much larger scale in the future.

Human subjects research

This research will involve human subjects. We will obtain IRB approval at Oxford University, and locally in Malawi for any work that we do, and more generally follow best-practices.

Budget

See the proposed <u>budget</u> attached. 100% of funds requested from GiveWell would be spent on research. The intervention is funded by donors of GiveDirectly.

Alternate funding sources

We have already raised an additional USD 1.4 million from the <u>UKRI Future Leaders' Fellowship</u> awarded to Dennis Egger for setting up the Oxford Centre for Macro-Experimental Development. The fellowship will additionally pay for PI time, and the hiring of a research manager, post-doctoral researcher, and pre-doc ensuring substantial additional personnel, management, oversight, and research assistance. The fellowship lasts for 4 years and encompasses the broader project, and the setup of the RCT and data collection infrastructure for the next decade. As such, it cannot be used fully for fieldwork only, and needs to be complemented by additional funding. We ask for funding from GiveWell specifically for all price, market, and enterprise related activities of this next phase of the program to complement the core infrastructure. This ensures that GiveWell's money will be spent highly cost-effectively, exclusively on fieldwork.