Call Me Maybe: Experimental Evidence on Using Mobile Phones to Survey African Microenterprises*

Rob Garlick†, Kate Orkin‡, and Simon Quinn§

March 15, 2016

LATEST VERSION · ONLINE APPENDIX
PRE-ANALYSIS PLAN · QUESTIONNAIRES

Abstract

We provide the first experimental evidence on the usefulness and viability of high-frequency phone surveys of microenterprises. We randomly assign microenterprises to three groups, who are interviewed face-to-face at monthly intervals (mimicking a standard method of collecting data from microenterprises), face-to-face at weekly intervals, and by mobile phone at weekly intervals. We find high frequency data collection is useful: it captures extensive volatility in a number of measures not visible in less frequent data collection. We also find it viable: on most measures and at most quantiles of the distribution, data patterns are indistinguishable in interviews conducted weekly or monthly and face-to-face or by phone. We administer a uniform face-to-face endline survey and verify that high-frequency data collection (either over a mobile phone or in person) does not substantially alter microenterprise owner behaviour over a 12-week period. These results demonstrate an important role for phone-based data collection in development research, particularly given the substantially cost differences that we document.

*We benefitted from comments from Nicola Branson, Markus Eberhardt, Simon Franklin, David Lam, Murray Leibbrandt, Gareth Roberts, Volker Schoer, Duncan Thomas and seminar audiences at the Development Economics Network Berlin, Duke University, NEUDC 2016, the University of Cape Town, the University of Oxford, and the University of the Witwatersrand. We thank Bongani Khumalo, Thembela Manyathi, Mbuso Moyo, Mohammed Motala, Egines Mudzingwa and the staff and fieldwork at the Community Agency for Social Enquiry (CASE) for their survey fieldwork; Mzi Shabangu, Arul Naidoo and the GIS section at Statistics South Africa for enumeration area maps; and Rose Page, Richard Payne and Gail Wilkins at the CSAE. This project was funded by the Private Enterprise Development for Low-Income Countries (PEDL), a joint research initiative of the Centre for Economic Policy Research (CEPR) and the Department For International Development (DFID), and we thank Chris Woodruff and the PEDL team for their support.

†Department of Economics, Duke University; **robert.garlick@duke.edu.**

‡Department of Economics, Centre for the Study of African Economies and Merton College, University of Oxford; **kate.orkin@merton.ox.ac.uk.**

§Department of Economics, Centre for the Study of African Economies and St Antony's College, University of Oxford; **simon.quinn@economics.ox.ac.uk.**

1 Introduction

We run the first randomized controlled trial to test the usefulness and viability of phone interviews and high-frequency interviews of microenterprises. We draw a representative sample of microenterprises in Soweto, South Africa, and randomly assign respondents to three treatment groups. The first group is interviewed face-to-face at monthly intervals, to mimic a standard method of collecting data from microenterprises. The second group is interviewed face-to-face at weekly intervals. This preserves one key feature of the benchmark method — that respondents are interviewed in person — but allows us to test the consequences of collecting data at a much higher frequency. The third group are interviewed at weekly intervals by mobile phone.

We find mobile phone interviews to be *accurate*: on most measures and at most quantiles of the distribution, data patterns are indistinguishable in interviews conducted weekly, whether by mobile phone or face-to-face). Using an endline administered face-to-face, we find little evidence that high-frequency data collection (either over a mobile phone or in person) alters microenterprise behaviour. We are powered to detect relatively small differences between different interview methods (median minimum detectable effect across outcomes is 0.07 standard deviations) so our findings are not explained by low sample sizes or noisy data. We also find interviews at weekly intervals to be *useful*: they capture extensive volatility in a number of measures that is not visible in less frequent data collection. We conclude that mobile phone data collection offers considerable cost savings with little, if any, reduction in data quality.

1.1 Motivation

Development economists are likely to find high-frequency data collection useful. First, high-frequency interviews improve *statistical power*, particularly for noisy and relatively less autocorrelated outcomes such as business profits, household incomes and expenditures, or for relatively rare and episodic outcomes, like shocks (McKenzie, 2012). Multiple measures at short intervals enable researchers to average out noise, increasing power. Second, high-frequency

follow-up measures in experiments can be very useful for measuring how treatment effects *evolve over time*: see, for example, Frison and Pocock (1992), McKenzie (2012) and Franklin (2015).[1] Third, high-frequency data is useful for measuring the *volatility* faced by microenterprises — and, in turn, for characterising heterogeneity in volatility faced by different microenterprises. Households and microenterprises in developing countries often face highly variable flows, lumpy expenses and extreme seasonal variation.[2] Indeed, many arguments in favour of certain policy interventions — such as microfinance — centre on their ability to enable households to smooth consumption (Rosenzweig and Wolpin, 1993; Gertler, Levine, and Morett, 2009; Banerjee, Duflo, Glennerster, and Kinnan, 2015; Karlan and Zinman, 2011). A reliable and cheap method of high-frequency data collection will enable researchers to characterize volatility; better understand patterns in volatility over time, seasonally and in response to shocks; better understand whether interventions reduce volatility; and examine whether interventions work differently for households experiencing more or less volatility.

Against these advantages, there are several large potential costs to using weekly surveys. First, some evidence suggests that different survey methods induce *substantially different reporting by respondents* in developing country contexts. For example, Das, Hammer, and Sánchez-Paramo (2012) find that length of recall period has a large impact on reported health and use of health services. Similarly, Beegle, De Weerdt, Friedman, and Gibson (2012) find significant lower reported consumption using recall modules than consumption diaries, particularly among poorer households, while Caeyers, Chalmers, and De Weerdt (2012) find that reported consumption is higher for paper-based interviewing that computer-assisted interviewing. Baird

[1] Similarly, high-frequency measures can be useful to test how treatment effects differ between periods of higher and lower microenterprise volatility: for example, Drexler, Fischer, and Schoar (2014) and Karlan and Valdivia (2011) find business training programs improve sales in bad months, but on average find no significant treatment effects.

[2] For example, McKenzie and Woodruff (2008) report that the percentage change in monthly profits from one calendar quarter to the next in data from Mexico ranges from -97.6 to +4110. The limited evidence suggests the majority of large changes reported are genuine, rather than measurement error. In one panel of firms with no paid employees in Accra, Ghana, Fafchamps, McKenzie, Quinn, and Woodruff (2012) queried owners if they reported changes in profits or sales from one quarter to the next over a threshold, but owners corrected only between 3 and 13 per cent of the large changes in profits and sales in a survey round. See also Collins, Morduch, Rutherford, and Ruthven (2009).

and Özler (2012) and Barrera-Osorio, Bertrand, Linden, and Perez Calle (2011) find significant over-reporting in self-reports on school attendance and enrolment compared to administrative records and observational data. Relatedly, enumerator error may differ by data capture method. Patnaik, Brunskill, and Thies (2009) find lower error rates when enumerators dictate survey answers over the phone to data capturers than using PDAs. Of six RCTs comparing PDAs to paper in developed world contexts, two found PDAs more accurate, one found paper more accurate, and three found no differences (Lane, Heddle, Arnold, and Walker, 2006).

Second, different survey frequencies may even *change actual respondent behaviour*, although results vary by the type of outcome. Reminder effects have only been found for outcomes where respondents pay limited attention to their behaviour in the absence of the survey. Zwane, Zinman, Van Dusen, Pariente, Null, Miguel, Kremer, Karlan, Hornbeck, Gine, Duflo, Devoto, Crepon, and Banerjee (2011) randomize the frequency of face-to-face surveys on health behaviour and diaorrhea incidence during a trial of water chlorination products. Among households surveyed biweekly for 18 rounds, there is less reported child diaorrhea and more uptake of the product, compared to households surveyed in just three of the 18 rounds; the authors attribute this to the effects of being reminded more frequently to invest in water purification. Uptake of the product is measured through sampling chlorine in the household's water as well as self-reports, so differences are not due to social desirability bias. In the US, Stango and Zinman (2013) find that taking a survey which asks about whether clients take an overdraft and pay fees reduces the probability of incurring any overdraft fee by an estimated 3.7 percentage points on a base of 30 per cent, conditional on being in a population which takes surveys for a market research company. Each survey within a two year period reduces the probability by 1.7 percentage points. Effects are stronger for respondents with lower education and financial literacy, which is consistent with the explanation that respondents don't pay attention to overdrafts. Relatedly, market researchers find asking subjects directly for a forecast of their likelihood of engaging in a targeted behaviour is known to improve the likelihood that they engage in that behaviour (Dholakia, 2010).

However, high-frequency surveys on economic outcomes in developing countries find that being surveyed frequently does not change behaviour. Beaman, Magruder, and Robinson (2014) found Kenyan microenterprise owners often forgot to bring cash in small denominations to break larger bills, losing sales between 5 and 8 percent of weekly profits. They administered a weekly survey asking about profits, sales, sales and profits lost to missing change and time spent looking for change, for between 8 and 18 weeks. Simply being surveyed three times about cash management had no significant effect on profits, sales or the amount of change brought to work, although it reduced the number and amount of lost sales due to not having change. However, an information intervention telling firms how much money they were losing reduced lost sales and increased profits. Franklin (2015) finds surveying randomly selected unemployed Addis Ababa youth every week by phone for three months has no effect on job search and employment outcomes compared to surveying them only at the start and end of the period: he argues that respondent awareness is not likely to be an important determinant of employment status.[3]

Conducting high-frequency interviews by phone is cheaper and easier than revisiting respondents multiple times in a short period for face-to-face interviews.[4] Mobile phone data collection may reduce attrition because respondents migrate or travel temporarily, provided migrants keep the same mobile phone number, and is likely to reduce tracking costs. Answering a phone call may be more convenient for respondents than being at a pre-arranged location for an interview.

[3] Research which finds effects on behaviour of being surveyed at all also finds effects differ by the type of outcome. In another part of their study, Zwane, Zinman, Van Dusen, Pariente, Null, Miguel, Kremer, Karlan, Hornbeck, Gine, Duflo, Devoto, Crepon, and Banerjee (2011) find that being surveyed at all had no effect on take-up of loans but did affect take-up of medical insurance. They argue consumption needs were likely salient anyway, but that a survey might make medical risks more salient.

[4] Researchers surveying conflict areas (van der Windt and Humphreys, 2013), remote areas (Bauer, Akakpo, Enlund, and Passeri, 2013) or areas suffering contagious disease outbreaks (Himelein, 2014) have used mobile phone surveys, and these are particularly attractive if existing sampling frames are available.

But can researchers rely on phone interviews? An extensive literature from developed countries suggests differences between phone and face-to-face interviews are mostly small. Jackle, Roberts, and Lynn (2006) randomly assigned European Social Survey respondents in Hungary and Portugal, sampled from a single sampling frame, to be interviewed either face-to-face or over a fixed line or mobile phone. Although variables which were most likely to be affected were selected for the experiment, there were differences in only 8 of 33 items, and differences were small; they were larger than one standard deviation in only one variable. Other randomised trials also find significant differences by mode in only a few variables, no more than would be expected by chance (Groves, 1979; Sykes and Hoinville, 1985; Körmendi, 2001; de Leeuw, 1992; Nord and Hopwood, 2007). Jackle, Roberts, and Lynn (2006) also find no evidence that mode effects are stronger for respondents with lower cognitive ability.[5]

Even if responses on average don't differ, phenomena such as item non-response, non-differentiation (Krosnick, 1991), acquiescence (Smith and Fischer, 2008) or response order effects (Schwarz, Hippler, and Noelle-Neumann, 1992) may be more prevalent in phone surveys and thus affect certain types of outcomes (such as attitudes questions using scales). However, randomised trials comparing phone and face-to-face surveys find no differences (Jackle, Roberts, and Lynn, 2006; de Leeuw, 1992).

1.2 Contribution and Results

There is very limited research to date on the viability of mobile phone interviewing in developing countries – and, to our knowledge, no research regarding microenterprises. (A household survey conducted biweekly by the World Bank and Gallup in Honduras found no differences between weeks when data was collecting using voice calls and weeks when it was collected

[5] Some political scientists in the US do find more significant differences (Holbrook, Green, and Krosnick, 2003; Shanks, Sanchez, and Morton, 1983). However, most research is done to compare polling results from samples chosen using different methods – face-to-face surveys sampled via conventional listing of randomly selected street blocks and telephone surveys sampled by random digit dialling – and cannot separate the effects of differences in sampling from differences in interview mode. This literature is less relevant if one is concerned with the effects of mode of interview on samples chosen using the same method.

face-to-face from the same respondents (Gallup, 2012).[6])

We collect a random sample of microenterprises in a low-income urban area outside Johannesburg, South Africa, by screening all households within sampled census enumeration areas to find households who were running at least one microenterprise at the beginning of the study. We remove enterprises with more than two full-time employees, which provided professional services, which operated fewer than three days and week and whose owners did not have a mobile phone. We randomly allocate enterprises to be surveyed either face-to-face weekly for 12 weeks, over the phone weekly for 12 weeks, or face-to-face every 4 weeks for 12 weeks. All surveys use an identical questionnaire, which takes approximately 15 minutes to administer and measures seventeen enterprise outcomes. We then conduct a common face-to-face endline with all microenterprises.

First, we find little evidence that interview frequency (weekly vs monthly) or interview medium (face-to-face vs phone) changes behaviour. We test whether endline responses differ between treatment groups and interpret this as a test of whether our interview techniques changed actual microenterprise outcomes. Behavioural changes may arise if regular surveys make information about enterprise performance more salient to owners. Owners interviewed weekly report a slightly and statistically significantly higher stock of fixed assets than owners interviewed monthly. Some insignificant but non-trivial differences in measures are weakly consistent with reminder effects but do not provide strong evidence of behavioural changes. owners interviewed by phone report significantly lower profits and slightly lower spending on stock/inventory, withdrawals for household use, and number of full-time employees. We do not believe these differences by interview medium are consistent with simple behavioural change explanations.

[6] For field notes describing the use of mobile phones in surveys, see Dillon (2012) and Croke, Dabalen, Demombynes, Giugale, and Hoogeveen (2014).

Second, we find mobile phone interviews are accurate compared to a benchmark of face-to-face interviews. We compare the reported outcomes for the panel surveys conducted after the baseline and before the endline. On most measures and at most quantiles of the distribution, data patterns are very similar across weekly phone and weekly face-to-face interviews. Owners interviewed by phone report working fewer hours and taking less money from the enterprise for their household (as well as lower stock/inventory levels), which may reflect differential strength of social desirability bias across interview media.

Third, we provide evidence on the extent and implications of attrition from a high-frequency microenterprise panel. Most microenterprises miss multiple scheduled interviews but the level of attrition does not vary with the medium of interview. Attrition is slightly higher for high- than low-frequency interviews but attrition does not increase through time. We interpret this pattern as evidence of the logistical challenges of conducting high-frequency interviews, not fatigue or irritation by respondents. The reasons for attrition are generally balanced across groups, though phone interviews yield slightly higher rates of refusal and incorrect contact information (presumably due to changing phone numbers). Attrition is predicted by few baseline microenterprise characteristics and the results reported above are robust to adjustment using inverse-probability-of-attrition weights. We conclude that attrition is an important component of panel data collection in this setting, but that high-frequency phone interviews do not produce substantially worse data quality due to attrition than conventional methods.

Our results show that development researchers studying microenterprises should seriously consider phone interviews. These are substantially cheaper than other face-to-face interviews and produce data of comparable quality. These results may also apply to studies of households, particularly given the intertwined nature of consumption and production behaviour in many developing countries (Singh *et al.*, 1986). The viability of phone surveys depends on phone penetration, but this is high and rising rapidly in most developing countries. South Africa, and the population we study, are hardly outliers in this respect.

2 Design and Data

2.1 Context

The study takes place in the township of Soweto, in South Africa's Gauteng province.[7] Soweto's population in October 2011 was approximately 1.28 million people. Residents are almost all Black Africans (99%) and most speak one of South Africa's local African languages (96%).[8] Of the 0.9 million residents aged 15 or older, 41% engage in some form of economic activity (including occasional informal work) and 78% of these adults work primarily in the formal sector. 19% of households report receiving zero annual income and another 42% report receiving less than $10 per day. These income figures reflect South Africa's middle income status and the fact that Soweto includes relatively few recent migrants.[9]

2.2 Sample definition and sampling strategy

We define an eligible microenterprise as any enterprise that: (i) has zero, one or two full-time employees (in addition to the owner); and (ii) does not provide a professional service (e.g. medicine); and (iii) operates at least three days each week; and (iv) whose owner has a mobile phone. The first two conditions operationalize our definition of 'microenterprise', consistent with the general approach in the development economics literature. The third condition excludes microenterprises that are seasonal, occasional (e.g. selling food at soccer games), or run over weekends in addition to wage employment. We impose this condition partly to ensure week-to-week variation in the outcomes of interest. The fourth condition is necessary to allow

[7] 'Township' in South Africa typically refers to low-income urban areas that include both formal and informal housing and were designated as Black African living areas under apartheid's residential segregation laws. Soweto is one of the older townships around Johannesburg. Relative to other townships, it includes more middle-income areas and houses fewer recent migrants.

[8] Statistics South Africa asks people to describe themselves in the Population Census in terms of five racial population groups, Black African, White, Coloured, Indian or Asian and Other or unspecified. We follow this terminology.

[9] Authors' own calculations, from the 2011 Census public release data.

phone surveys and does not bind for any otherwise eligible microenterprises.[10]

We used a three-stage clustered sampling scheme, designed so that our sample is representative of the population of households who live in `low-income' areas of Soweto and own eligible microenterprises. This is discussed in detail in the Online Appendix.

2.3 Data collection and assignment to data collection methods

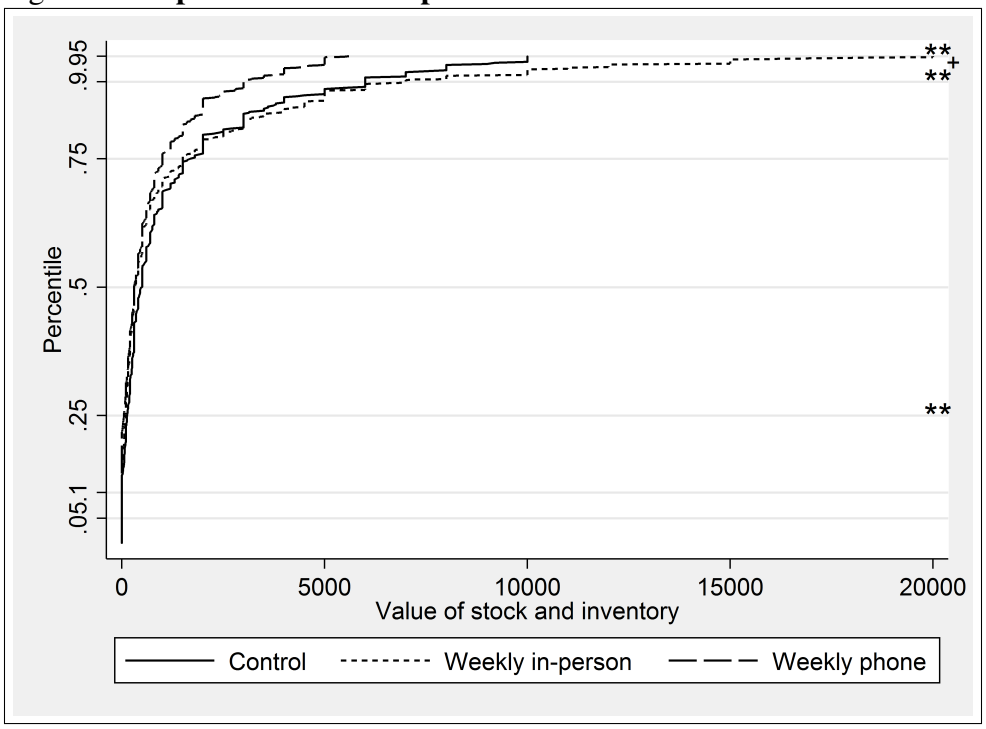
We employed a four-stage data collection process.[11] In the first stage, we conducted a screening survey with all households in the sampled small area layers (SALs). This short survey identified whether anyone in the household owned a microenterprise and, if so, asked further questions about the microenterprise and owner. This survey established whether the microenterprise met the eligibility criteria laid out in section 2.2 and collected phone numbers for the owner and other household members. The screening process took place between September 2013 and February 2014 and realized a sample of 1081 eligible microenterprises. Where a household included multiple eligible microenterprises, we randomly selected one for the final sample, which reduced the sample to 1046.

In the second stage, we conducted a baseline survey with all eligible microenterprise owners identified in the screening stage. These interviews were conducted at the enterprise premises to verify that the enterprises existed, whereas the screening survey was conducted at the owners' homes. The baseline survey asked 30 questions about the microenterprise and owner and recorded the GPS coordinates for the enterprise location. We completed the baseline questionnaire between December 2013 and February 2014 with only 895 of the 1046 microenterprise

[10] This pattern is unsurprising given that 87% of South Africans aged 18 or older own a mobile phone (Mitullah and Kama, 2013). This rate is only slightly higher than other African countries where comparable surveys were conducted by Afrobarometer between 2011 and 2013: 76% in Ghana, 81% in Kenya, 74% in Nigeria, 65% in Tanzania, and 62% in Uganda. Political polling firms in developed countries consider 80% phone ownership to be the threshold above which reliable collection of representative surveys over the phone can occur, with some reweighting (Croke, Dabalen, Demombynes, Giugale, and Hoogeveen, 2014).

[11] All questionnaires are available for download at www.robgarlick.com/research .

Figure 6: **Empirical CDFs in repeated interviews: Stocks and inventories**

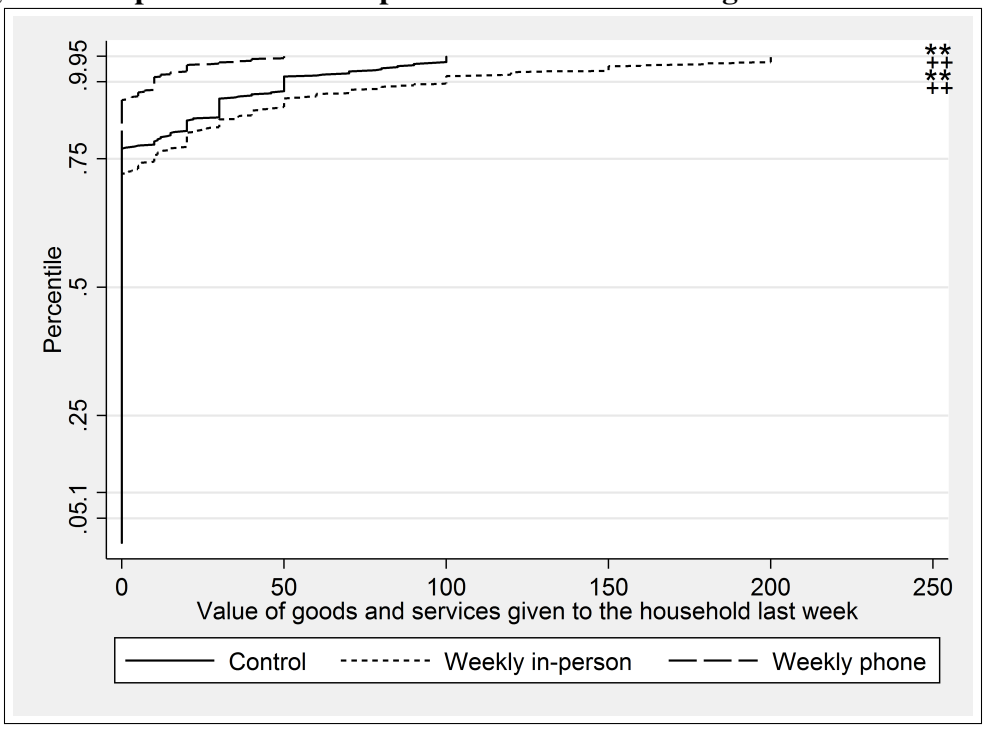


We test for differences at each of the levels shown on the y-axis. We cluster by microenterprise (Silva and Parente, 2013) and use the False Discovery Rate (Benjamini, Krieger, and Yekutieli, 2006) to control for multiple testing across quantiles.

'$+$' indicates significance on the null hypothesis that weekly surveys in-person and by phone are equivalent;
'$*$' indicates significance on a null hypothesis that all three treatments are equivalent.

'$^{+}$' and '*' denote $p < 0.1$; '$^{++}$' and '**' denote $p < 0.05$; '$^{+++}$' and '***' denotes $p < 0.01$.

Figure 7: **Empirical CDFs in repeated interviews: Value given to the household**



We test for differences at each of the levels shown on the y-axis. We cluster by microenterprise (Silva and Parente, 2013) and use the False Discovery Rate (Benjamini, Krieger, and Yekutieli, 2006) to control for multiple testing across quantiles.

'$+$' indicates significance on the null hypothesis that weekly surveys in-person and by phone are equivalent;
'$*$' indicates significance on a null hypothesis that all three treatments are equivalent.

'$^{+}$' and '*' denote $p < 0.1$; '$^{++}$' and '**' denote $p < 0.05$; '$^{+++}$' and '***' denotes $p < 0.01$.

Figure 8: **Other empirical CDFs in repeated interviews**

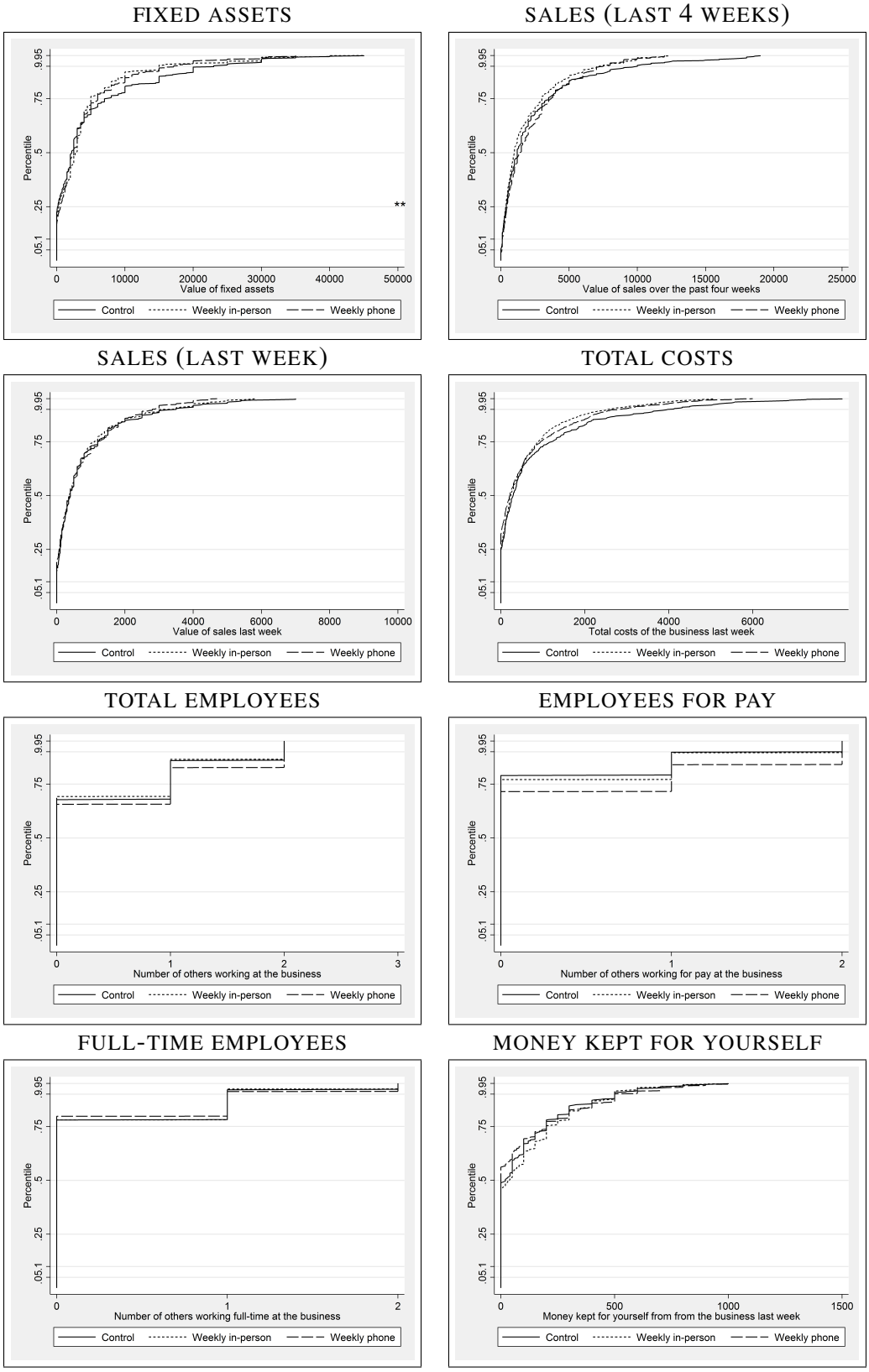


Table 5: **Regression results: Average Treatment Effect for repeated interviews (equation 2)**

	(1) Operating	(2) Stock & inventory	(3) Fixed assets	(4) Profit	(5) Sales last week	(6) Sales last 4 weeks	(7) Total costs	(8) Profit check
Weekly in-person	**0.017**	**0.737**	**1.659**	**-0.110**	**-0.070**	**-0.111**	**-0.078**	**-0.101**
	(0.010)	(0.223)***	(1.598)	(0.061)*	(0.068)	(0.067)*	(0.066)	(0.056)*
Weekly by phone	**0.014**	**0.120**	**0.316**	**-0.098**	**-0.093**	**-0.092**	**-0.049**	**-0.020**
	(0.010)	(0.155)	(0.431)	(0.074)	(0.072)	(0.074)	(0.069)	(0.064)
Observations	4070	3989	3988	3986	3985	3987	3987	3984
Weekly treatments equal (p)	0.566	0.006***	0.287	0.843	0.673	0.748	0.538	0.029
All treatments equal (p)	0.262	0.004***	0.568	0.191	0.429	0.251	0.475	0.027
MDE: Weekly in-person	0.015	0.270	0.079	0.072	0.072	0.077	0.070	0.060
MDE: Weekly by phone	0.015	0.273	0.080	0.072	0.073	0.078	0.071	0.060
Lee bound: Weekly in-person (lower)	0.014	-0.303	-0.229	-0.287	-0.267	-0.284	-0.295	-0.288
Lee bound: Weekly in-person (upper)	0.039	0.909	2.106	-0.060	-0.010	-0.056	-0.045	-0.070
Lee bound: Weekly by phone (lower)	0.011	-0.303	-0.210	-0.278	-0.236	-0.237	-0.273	-0.259
Lee bound: Weekly by phone (upper)	0.038	0.138	0.434	-0.069	-0.065	-0.057	-0.029	0.008

	(1) Employees	(2) Full-time	(3) Paid	(4) Hours yesterday	(5) Money kept	(6) Household takings	(7) Honest	(8) Careful	(9) Written records
Weekly in-person	**0.067**	**0.015**	**0.123**	**0.040**	**0.090**	**0.313**	**-0.152**	**-0.109**	**-0.015**
	(0.077)	(0.077)	(0.079)	(0.071)	(0.066)	(0.123)**	(0.028)***	(0.030)***	(0.017)
Weekly by phone	**0.043**	**-0.009**	**0.163**	**-0.424**	**0.012**	**-0.021**	**-0.378**	**-0.271**	**-0.109**
	(0.075)	(0.081)	(0.078)**	(0.070)***	(0.061)	(0.113)	(0.029)***	(0.031)***	(0.023)***
Observations	3987	3984	3973	3987	3986	3986	4056	4056	3987
Weekly treatments equal (p)	0.735	0.754	0.616	0.000***	0.162	0.009***	0.000***	0.000***	0.000***
All treatments equal (p)	0.681	0.950	0.096*	0.000***	0.280	0.015**	0.000***	0.000***	0.000***
MDE: Weekly in-person	0.104	0.113	0.118	0.082	0.063	0.118	0.038	0.036	0.025
MDE: Weekly by phone	0.105	0.114	0.119	0.082	0.064	0.119	0.039	0.036	0.026
Lee bound: Weekly in-person (lower)	-0.318	-0.361	-0.265	-0.235	-0.249	-0.261	-0.223	-0.203	-0.031
Lee bound: Weekly in-person (upper)	0.152	0.086	0.195	0.359	0.176	0.432	-0.033	-0.009	0.057
Lee bound: Weekly by phone (lower)	-0.201	-0.300	-0.110	-0.616	-0.236	-0.295	-0.446	-0.342	-0.129
Lee bound: Weekly by phone (upper)	0.093	0.037	0.216	-0.291	0.052	0.002	-0.338	-0.234	-0.012

Standard errors in parentheses, allowing for clustering by microenterprise. '' denotes $p < 0.1$; '**' denotes $p < 0.05$; '***' denotes $p < 0.01$.*

Table 6: **Regression results: Effect on sample standard deviation for repeated interviews (equation 3)**

	(1) Stock & inventory	(2) Fixed assets	(3) Profit	(4) Sales last week	(5) Sales last 4 weeks	(6) Total costs	(7) Profit check	(8) Sales check
Weekly in-person	**0.486**	**5.241**	**-0.004**	**-0.042**	**-0.045**	**0.039**	**0.008**	**-159.643**
	(0.129)***	(5.050)	(0.050)	(0.053)	(0.052)	(0.040)	(0.051)	(232.942)
Weekly by phone	**0.151**	**1.638**	**0.061**	**-0.058**	**-0.025**	**0.123**	**0.151**	
	(0.099)	(1.644)	(0.078)	(0.051)	(0.049)	(0.050)**	(0.065)**	
Observations	610	609	608	608	608	608	608	366
Weekly treatments equal (p)	0.012**	0.308	0.293	0.617	0.512	0.038**	0.002	.
All treatments equal (p)	0.001***	0.583	0.576	0.520	0.648	0.039**	0.009	.

	(1) Employees	(2) Full-time	(3) Paid	(4) Hours open yesterday	(5) Money kept	(6) Household takings
Weekly in-person	**-0.031**	**-0.006**	**-0.015**	**-0.000**	**0.053**	**0.494**
	(0.048)	(0.055)	(0.043)	(0.064)	(0.093)	(0.171)***
Weekly by phone	**0.095**	**0.014**	**0.052**	**0.087**	**-0.027**	**0.359**
	(0.051)*	(0.055)	(0.047)	(0.052)*	(0.079)	(0.177)**
Observations	608	608	608	608	608	608
Weekly treatments equal (p)	0.001***	0.663	0.079*	0.076*	0.282	0.477
All treatments equal (p)	0.006***	0.907	0.208	0.087*	0.560	0.010

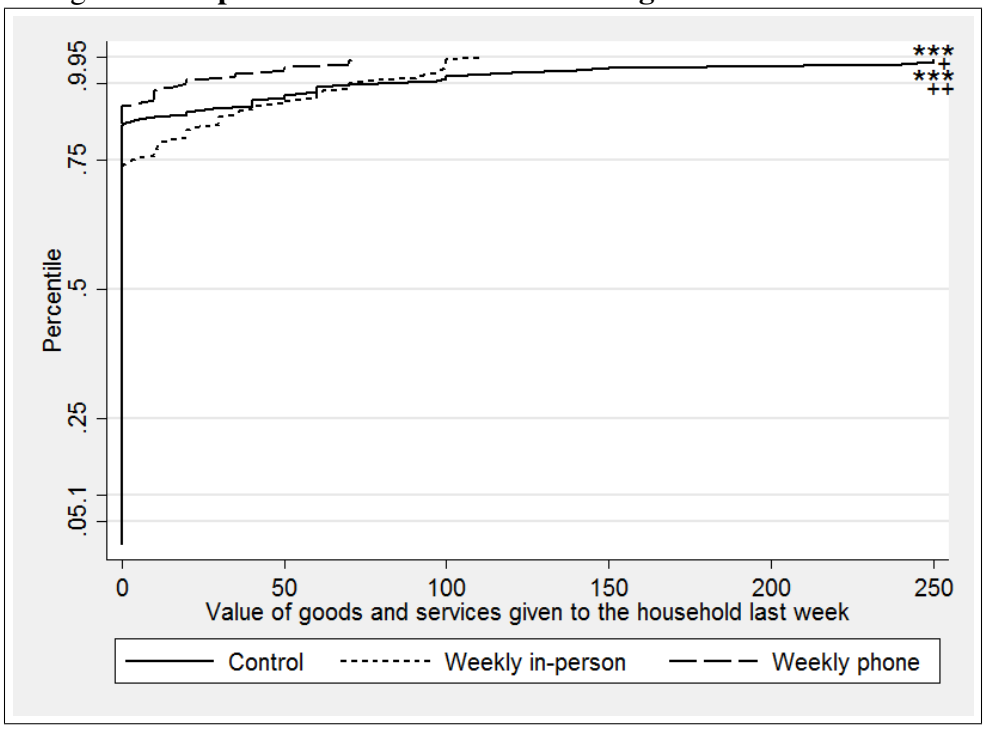
Robust standard errors in parentheses. ' ' denotes $p < 0.1$; '** ' denotes $p < 0.05$; '*** ' denotes $p < 0.01$.*

Table 7: **Blundell-Bond estimates: Disaggregating by treatment**

	(1) Profit *In-person*	(2) Profit *Phone*	(3) Capital *In-person*	(4) Capital *Phone*
Lag 1	0.477***	0.363***	0.459***	0.380***
	(0.070)	(0.079)	(0.115)	(0.103)
Lag 2	0.305***	0.261***	0.194***	0.195*
	(0.114)	(0.076)	(0.075)	(0.115)
Lag 3			0.158***	0.137***
			(0.048)	(0.046)
Lag 4			0.175***	0.104***
			(0.065)	(0.036)
Time dummies	✓	✓	✓	✓
Observations	520	289	420	418
Firms	132	87	116	89
Arellano-Bond: AR(1) (p-value)	0.000***	0.000***	0.024**	0.003***
Arellano-Bond: AR(2) (p-value)	0.302	0.168	0.325	0.140
Arellano-Bond: AR(3) (p-value)	0.340	0.319	0.783	0.967
Arellano-Bond: AR(4) (p-value)	0.242	0.392	0.323	0.497
Hansen test (p-value)	0.199	0.988	0.288	0.596
H_0: Equal parameter estimates (p-value)	0.548		0.907	

Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
The final row reports Wald tests of the a null hypothesis of equal lag coefficients between in-person and phone data.
To run this test, we impose zero parameter covariance between estimations (since the estimation samples are non-overlapping).

Figure 9: **Empirical CDFs at endline: Value given to the household**



We test for differences at each of the levels shown on the y-axis. We cluster by microenterprise (Silva and Parente, 2013) and use the False Discovery Rate (Benjamini, Krieger, and Yekutieli, 2006) to control for multiple testing across quantiles.

'$+$' indicates significance on the null hypothesis that weekly surveys in-person and by phone are equivalent;
'$*$' indicates significance on a null hypothesis that all three treatments are equivalent.

'$^{+}$' and '*' denote $p < 0.1$; '$^{++}$' and '**' denote $p < 0.05$; '$^{+++}$' and '***' denotes $p < 0.01$.

Table 8: **Regression results: Average Treatment Effect for endline interviews (equation 4)**

	(1) Operating	(2) Stock & inventory	(3) Fixed assets	(4) Profit	(5) Sales last week	(6) Sales last 4 weeks	(7) Total costs	(8) Profit check
Weekly in-person	**-0.004**	**0.380**	**-0.132**	**0.233**	**0.354**	**-0.065**	**0.108**	**0.479**
	(0.026)	(0.243)	(0.078)*	(0.222)	(0.258)	(0.092)	(0.106)	(0.297)
Weekly by phone	**-0.033**	**-0.167**	**-0.106**	**-0.050**	**0.034**	**-0.018**	**0.214**	**0.059**
	(0.028)	(0.162)	(0.091)	(0.141)	(0.155)	(0.124)	(0.147)	(0.169)
Observations	594	546	546	546	546	546	546	546
Weekly treatments equal (p)	0.281	0.082*	0.722	0.024**	0.110	0.703	0.481	0.131
All treatments equal (p)	0.437	0.217	0.241	0.033**	0.276	0.773	0.293	0.260

	(1) Employees	(2) Full-time	(3) Paid	(4) Money kept	(5) Household takings	(6) Hours open yesterday	(7) Honest	(8) Careful	(9) Written records
Weekly in-person	**0.111**	**0.071**	**0.094**	**0.371**	**0.013**	**-0.003**	**0.053**	**0.058**	**-0.016**
	(0.090)	(0.092)	(0.091)	(0.272)	(0.138)	(0.095)	(0.045)	(0.048)	(0.030)
Weekly by phone	**0.012**	**-0.086**	**0.254**	**0.089**	**-0.170**	**0.034**	**-0.029**	**0.005**	**0.052**
	(0.113)	(0.096)	(0.263)	(0.162)	(0.095)*	(0.128)	(0.048)	(0.050)	(0.030)*
Observations	546	546	546	546	546	546	593	593	546
Weekly treatments equal (p)	0.348	0.081*	0.574	0.050*	0.297	0.770	0.077*	0.273	0.018
All treatments equal (p)	0.402	0.219	0.340	0.113	0.203	0.955	0.191	0.397	0.049**

Standard errors in parentheses, allowing for clustering by microenterprise. ' ' denotes $p < 0.1$; '** ' denotes $p < 0.05$; '*** ' denotes $p < 0.01$.*