



Follow the money not the cash: Comparing methods for identifying consumption and investment responses to a liquidity shock☆



Dean Karlan ^{a,b,c,d,*}, Adam Osman ^e, Jonathan Zinman ^{b,c,d,f}

^a Yale University, United States

^b IPA, United States

^c J-PAL, United States

^d NBER, United States

^e University Illinois at Urbana-Champaign, United States

^f Dartmouth College, United States

ARTICLE INFO

Article history:

Received 22 November 2014

Received in revised form 15 September 2015

Accepted 28 October 2015

Available online 6 November 2015

JEL classification:

D12

D22

D92

G21

O12

O16

Keywords:

Loan use

Consumption

Investment

Liquidity constraint

Liquidity shock

Fungibility

Microcredit

Microenterprise

ABSTRACT

Measuring the impacts of liquidity shocks on spending is difficult methodologically but important for theory, practice, and policy. We compare three approaches for tackling this question: directly asking borrowers how they spend proceeds from a loan (direct elicitation); asking borrowers using a list randomization technique (indirect elicitation) that allows them to answer discretely in cases where loan uses are at odds with lender policies or social norms; and, a counterfactual analysis in which we compare household and enterprise cash outflows for those in a treatment group, randomly assigned to receive credit, to a control group. The counterfactual analysis yields an estimate that about 100% of loan-financed spending is on business inventory. For the direct and indirect elicitations, we find evidence of both strategic misreporting and “following the cash”: borrowers likely report what they physically did with cash proceeds, rather than counterfactual spending.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

What are the impacts of liquidity shocks on the consumption and investment decisions of households and small businesses? Answers to

this question have implications for the theory, practice, and regulation of credit, as well as for modeling intertemporal consumer choice. They shed light on perceived returns to investment, and on the extent to which constraints bind more for some types of household spending than others. Estimating impacts of liquidity shocks matters in many domains, for example in understanding household leveraging and deleveraging decisions in the wake of credit supply shocks,¹ as well as evaluating interventions such as business grants,² unconditional cash transfers,³ and microcredit expansions.⁴

☆ Approval from the Yale University Human Subjects Committee, IRB0510000752 and from the Innovations for Poverty Action Human Subjects Committee, IRB #07October-002. The authors thank financial support from the Bill and Melinda Gates Foundation, the Consultative Group for Assistance to the Poor (CGAP) and AusAID. The authors thank Kareem Haggag, Romina Kazadjian, Megan McGuire, Faith McCollister, Mark Miller, and Sarah Oberst at Innovations for Poverty Action for project management and field support throughout the project, and the senior management and staff at First Macro Bank and FICO Bank for their support and collaboration throughout this project. The authors retained full intellectual freedom to report and interpret the results throughout the study. All errors and opinions are those of the authors.

* Corresponding author.

E-mail addresses: dean.karlan@yale.edu (D. Karlan), aosman@illinois.edu (A. Osman), jzinman@dartmouth.edu (J. Zinman).

¹ See e.g. Hall (2011), Eggertsson and Krugman (2012), and Mian and Sufi (2011).

² See e.g. Fafchamps et al. (2014), Karlan et al. (2013), and De Mel et al. (2008).

³ See e.g. Benhassine et al. (Forthcoming), Blattman et al. (2012), Haushofer and Shapiro (2013), Karlan et al. (2013).

⁴ See e.g. Angelucci et al. (2015), Attanasio et al. (2015), Augsburg et al. (2015), Banerjee et al. (2015a,b), Crépon et al. (2015), Karlan and Zinman (2010), Karlan and Zinman (2011), Tarozzi et al. (2015), and an overview in Banerjee et al. (2015a,b).

Papers that track responses to liquidity shocks often focus on estimating medium- and long-term effects by measuring spending patterns, balance sheets, or summary statistics of financial conditions several months or years post-shock.⁵ This reduced-form evidence has proven quite useful, but it often leaves the mechanism underlying any change unidentified. For each possible outcome observed many months post-liquidity shock, there are many paths from the liquidity change to that outcome. We compare different ways to identify what happens after the liquidity shock in order to identify the mechanisms driving the longer-term results. Identifying mechanisms is important because different paths can have different welfare implications.

To take an example closest to the setting we examine, many microcredit impact evaluations do not find significant effects of microcredit on enterprise scale or profitability one or two years post-intervention, even when the loans are targeted to microentrepreneurs.⁶ There are at least three possible explanations for these findings: (1) impacts only materialize over longer horizons due to compounded benefits, adjustment, etc. This hypothesis often motivates researchers and program advocates to highlight the value of longer-term outcome data; (2) microentrepreneurs do *not* actually invest marginal liquidity in their businesses, perhaps because they are credit constrained on the margin and have household investment or consumption smoothing with a higher expected return on investment (in utility terms) than business investment; (3) microentrepreneurs *do* invest microloan proceeds in their businesses, but these investments do not end up earning a positive net return.

The second and third explanations highlight the potential value of “following the money” from liquidity to spending decisions to reveal mechanisms underlying the paths from shock to outcomes. If the second explanation is accurate, that motivates further attempts to identify causes, consequences, and cures for credit constraints. If the third explanation is accurate that motivates further attempts to understand why entrepreneurs make investments that, ex-post at least, do not yield a positive net return on average (Anagol et al., 2013; Karlan et al., 2013; Moskowitz and Vissing-Jorgensen, 2002).⁷

But how exactly one might go about measuring spending in the immediate aftermath of a liquidity shock is not immediately obvious, methodologically speaking. There are several challenges.

Administrative data is rarely available for the appropriate sample, timeframe, or spending frequency, and even more rarely sufficiently comprehensive in its coverage of different types of consumption and investment. This makes survey design very important. Yet money is fungible, and household and (micro)enterprise balance sheets are often complex, making it difficult for respondents to identify the effects of the liquidity shock on their spending. Similarly, surveys that simply ask about past purchases produce noisy data, and measurement error increases with the length of the recall period (Nicola et al., 2012). Moreover, surveys can produce biased, rather than merely noisy, data if respondents have justification bias,⁸ neglect to report some accounts to save time or effort, worry about surveyors sharing information with tax authorities or a lender that “requires” loans be used for particular purpose, or feel stigma about using debt for consumption purposes (Karlan and Zinman, 2008). In short, data constraints, strategic

reporting, and respondent (mis)perceptions may all make it difficult to follow the money.

We address these challenges by comparing results from three different methods for following the money obtained by borrowers subjected to a randomized supply shock from one of two microlenders in the Philippines. The three methods are: 1) direct elicitation (i.e., ask individuals how they will, or did, spend the money from a loan); 2) indirect elicitation (i.e., in a setting where someone may deliberately be untruthful, employ a survey technique that indirectly elicits their response); and 3) counterfactual analysis (i.e., compare cash outflows of those who borrowed to a counterfactual group that did not borrow).

We compare results from these three approaches, and consider the counterfactual analysis to yield evidence closest to the unbiased “truth”. We examine three categories of loan uses: buying household goods, making enterprise investments, and paying down other debt. The direct elicitation questions ask about (intended) loan usage with questions along the lines of: “Will you/did you spend at least \$X of your loan on Y?” We ask these questions using four different enumerators and timing combinations: by bank staff at application and shortly after the disbursement of the loan, and by independent surveyors two weeks and two months after loan disbursement. The indirect elicitation questions employ a “list randomization” technique that makes it feasible for respondents to respond truthfully to sensitive questions without actually revealing details about their behavior (Karlan and Zinman, 2012). This is important in our setting because the lenders require that loan proceeds be used for business purposes only. Our counterfactual analysis takes advantage of a randomized trial in which marginal applications were randomly assigned to either treatment or control. Then, at both two weeks and two months post-randomization, independent surveyors asked about all cash outflows from the individual’s household or business that exceeded a certain amount, and compare treatment to control to estimate the impact of the liquidity shock on specific outcomes.

Comparing results across these methods sheds light on several questions. We infer how borrowers believe they should *report* their loan usage by comparing results across the four direct elicitation combinations. We infer how borrowers actually *perceive* the impact of the loan on their spending decisions (versus how they should report it) using indirect elicitation (the list randomization). And we infer how these perceptions, at least as elicited, differ from *actual* spending: the counterfactual analysis which compares responses to short-run spending questions across treatment and control groups.⁹

We find important differences, particularly for paying down other debt. Only 2.3% of borrowers report *to the bank* that they use loan proceeds to pay down other debt. Yet 19.1% (s.e. = 4.9%) of our treatment group implicitly (via list randomization) reports doing so. Thus many perceive themselves to be paying down other debt, but do not report this to bank (naturally, since this would be frowned upon by the bank). However, counterfactual analysis suggests that the misreporting is actually closer to the truth, once fungibility is taken into account: only 1.6% (s.e. = 2.9%) of individuals actually paid down more debt relative to the control group. Similar results hold for spending on household items (although the results are less precisely estimated).

We estimate that the treatment effect on spending is entirely on business investment, specifically inventory. This average treatment effect can account for the entire average loan amount two weeks post-randomization, with even larger but more noisily estimated treatment effects at two months post-randomization. This result highlights how our preferred method of identifying counterfactual spending can complement longer-run follow-up data; e.g., in our setting, it will be

⁵ See e.g. Souleles (1999); Souleles (2002); Shapiro and Slemrod (2003); Johnson et al. (2006); Agarwal et al. (2007); Shapiro and Slemrod (2009); Sahm et al. (2012); Parker et al. (2013).

⁶ See the studies cited in the above footnote, with the exception of Karlan and Zinman (2010), which examines untargeted consumer loans.

⁷ Now consider the opposite state of the world: say an evaluation of 12-month impacts does find that a microcredit expansion produces larger, more profitable businesses. The mechanism need not be investment in business assets per se (inventory, physical capital, etc.) Rather, it could be investments in human capital (training, health, child care, etc.) that enable the entrepreneur or business “helpers” from her family to be more productive.

⁸ E.g., my business did not grow from last year to this year, so I won’t report (to the surveyor, or even perhaps to myself) that I actually did try to grow my business by investing in new assets earlier this year.

⁹ Counterfactual analysis in this case means comparing the difference in spending between treatment and control two weeks and two months post-randomization. We find that the difference in spending between the groups is about the same as the average loan provided to the treatment group by the bank.

interesting to see whether the short-run increases in inventory translate into long-run increases, and into higher profits.

Why do people *think* they have flouted bank rules by spending on non-business purposes—assuming that list randomization elicits *perceived* truthful responses—even though our counterfactual analysis suggests that they have not? We suspect that they respond in a more proximate, mechanical sense by following the cash: I took the loan proceeds and bought X with the cash (even if X would have been bought anyhow, through some other means). This finding highlights the risks of making inferences based on direct or indirect survey questions about loan uses. These sorts of questions appear in many surveys and have been used to make inferences like “most borrowers use payday loans to cover ordinary living expenses over the course of months, not unexpected emergencies over the course of weeks.”¹⁰

All told, our key contributions are methodological. Our study highlights the limits of direct and indirect survey questions about loan uses, and the value of short-run, high-frequency data collection on substantial outflows following a liquidity shock.

2. Market overview

We collected data with the cooperation of two different banks in the Philippines, one in Metro Manila (covering mostly peri-urban areas) and another in northern Luzon. Both banks are for-profit institutions that offer individual liability microloans at about 60% APR. Loan sizes range from 5000 pesos to 50,000 pesos, with a mean (median) of 13,996 (10,000) in our sample. Loan maturities range from three to six months, with weekly repayments of principal and interest. Both banks require that applicants have an existing business, and be between 18 and 65 years old.

The Metro Manila bank has operated in the region since the 1960s. It had microloans outstanding to about 2700 borrowers as of July 2013. This portfolio represents a small fraction of its overall lending, which also includes larger business and consumer loans, and home mortgages. Until the end of 2012, the bank's microlending activities received subsidized technical assistance from a USAID-funded program.¹¹ The second bank has operated in mostly rural areas of northern Luzon since the 1980s. It had microloans outstanding to 26,000 borrowers in 2011 and offers other financial products as well.

The microloan market in the Philippines is somewhat competitive, as described in Karlan and Zinman (2011). There are informal options as well, including moneylenders. For our purposes the key fact is that that rejected borrowers do not simply obtain credit elsewhere: our banks' random assignments to credit actually do produce a substantial change in the total/net borrowing of applicants (see Section 3.6 below).

Our sample is comprised of 1661 marginal loan applicants who were randomized into loan approval or denial (see Section 3.2 for details on the randomization). Table 1 Column 1 provides baseline descriptive statistics gleaned from loan applications. 85.1% of the sample are women, 71.2% are married, and 32.9% are college educated.¹² The average applicant is 41.7 years old and has owned her business for 6.6 years. Nearly half of the businesses are “sari-sari” (corner/convenience) stores. Average weekly cash flow in the businesses is 5491 pesos (a bit more than \$125).

3. Methods and results

3.1. Overview

To better understand how borrowers deploy loan proceeds, and report thereon, we follow individuals from when they first apply for a loan until two months later. By that endpoint, we suspect that most of any proceeds will have been spent; this seems like a reasonable assumption given the high interest rates and short maturities. Along the way we use a variety of different methods to try to get at the same underlying question: how did the loan change the client's spending relative to a counterfactual in which the loan was not available?

The data come from four different interactions, with the same individual, over the course of about two months. Fig. 1 summarizes the timeline and the data collected in four distinct steps: (1) an application for a bank loan by the individual (direct elicitation); (2) a short survey of approved applicants at their first repayment, administered by a loan officer (direct elicitation); (3) a questionnaire by an independent surveyor two to three weeks after the loan application (direct and indirect elicitation methods, and cash outflows for counterfactual analysis) (4) a questionnaire by an independent surveyor about two months after the loan application (direct elicitation and cash outflows for counterfactual analysis).

In principle, with a larger sample size, testing different elicitation methods would be better done across-subjects than within-subjects, to avoid the possibility that subjects prefer to give consistent responses. Note that such an effect pushes against our core finding that responses differ in interesting ways across elicitation methods.¹³ Thus our results may underestimate the true extent to which elicitation methods influence inferences about the uses of marginal liquidity.

3.2. Sample creation and randomization

Our sample is comprised of 1661 marginally creditworthy microloan applicants to the two banks described in Section 2. Individuals applied from one of 16 bank branches at the Northern Luzon lender, or 8 branches at the Metro Manila lender, between July 2010 and March 2012. Each loan application is digitized by bank staff and credit-scored by underwriting software. For the purposes of this study, relatively small numbers of applicants with the highest (lowest) scores were automatically approved (rejected). The remaining applicants (about 85% of the pool) were randomly assigned to approval (with 90% probability) or rejection (with 10% probability).¹⁴

This random allocation of loans to marginal clients serves as the identifying instrument for our analysis of the expenditure data described in Sections 3.5 and 3.6 below. Table 1 Column 3 confirms that the treatment and control groups are observably identical, in a statistical sense: regressing treatment assignment on treatment strata and the complete set of baseline characteristics in Table 1, we do not reject the hypothesis that the characteristics are jointly uncorrelated with treatment assignment (p -value = 0.477).

3.3. Data collection step 1: at application, by loan officer (direct elicitation)

The first pieces of data on loan uses come from loan applications. Applications are extensive and take the form of loan officers interviewing applicants, reviewing their documents, and entering data into a small netbook computer. This process typically takes at least an hour to

¹⁰ <http://www.pewtrusts.org/en/research-and-analysis/reports/2012/07/19/who-borrows-where-they-borrow-and-why> For a recent example from a developing country context, see Groh and McKenzie (2014) which examines a macroinsurance scheme tied to a credit product, and, to be fair, acknowledges the difficulty of making inferences from survey questions on loan usage.

¹¹ The program was administered by Chemonics, Microenterprise Access to Banking Services (MABS).

¹² Females were not directly targeted by the bank. Enterprises of this size in the Philippines have greater female ownership; larger loans are serviced by a different part of the bank.

¹³ On the other hand, if respondents notice they are being asked the same question in different manners they may instead change their answer if they believe that they are asked again because they provided the “wrong” answer the first time. This would bias us toward finding an effect.

¹⁴ Although a 50/50 split to treatment and control would have provided more statistical power, the banks were interested in aggressively expanding their pool of borrowers.

Table 1
Orthogonality of treatment to applicant characteristics and attrition.

Demographic variables	Baseline			2-Week Survey			2-Month Survey		
	Control	Treatment	Difference	Control	Treatment	Difference	Control	Treatment	Difference
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Male	0.139 {0.394}	0.151 {0.386}	0.011 (0.033)	0.114 {0.380}	0.139 {0.380}	0.025 (0.035)	0.111 {0.380}	0.139 {0.378}	0.028 (0.037)
Age	40.987 {9.70}	41.809 {9.12}	0.821 (0.79)	40.501 {9.44}	41.940 {9.10}	1.439 * (0.84)	41.058 {9.10}	42.357 {9.17}	1.299 (0.89)
Marital status – single	0.137 {0.363}	0.084 {0.290}	–0.052 * (0.029)	0.135 {0.361}	0.086 {0.290}	–0.049 (0.031)	0.141 {0.290}	0.072 {0.272}	–0.069 ** (0.034)
Marital status – married	0.675 {0.459}	0.716 {0.439}	0.040 (0.037)	0.684 {0.453}	0.719 {0.434}	0.036 (0.040)	0.672 {0.434}	0.725 {0.424}	0.053 (0.043)
Marital status – widowed/separated	0.115 {0.294}	0.125 {0.314}	0.010 (0.025)	0.109 {0.277}	0.126 {0.309}	0.017 (0.025)	0.138 {0.309}	0.137 {0.311}	0.009 (0.028)
Education – college	0.303 {0.486}	0.284 {0.468}	–0.020 (0.040)	0.301 {0.486}	0.258 {0.458}	–0.042 (0.043)	0.274 {0.458}	0.237 {0.450}	–0.037 (0.047)
Education – some college	0.298 {0.447}	0.266 {0.435}	–0.032 (0.037)	0.281 {0.438}	0.277 {0.441}	–0.004 (0.039)	0.293 {0.441}	0.284 {0.444}	–0.009 (0.043)
Education – graduated high school	0.325 {0.423}	0.383 {0.470}	0.058 * (0.036)	0.325 {0.422}	0.391 {0.473}	0.066 * (0.039)	0.324 {0.473}	0.398 {0.475}	0.073 * (0.042)
Education – some high school or less	0.074 {0.325}	0.067 {0.293}	–0.008 (0.027)	0.095 {0.347}	0.074 {0.301}	–0.021 (0.031)	0.109 {0.301}	0.081 {0.307}	–0.028 (0.034)
Number of dependents	1.869 {1.554}	1.895 {1.448}	0.026 (0.130)	1.907 {1.599}	1.915 {1.463}	0.008 (0.146)	1.947 {1.463}	1.939 {1.464}	–0.008 (0.164)
Number of employees	0.828 {2.024}	0.800 {1.953}	–0.028 (0.180)	0.723 {1.412}	0.786 {2.035}	0.063 (0.149)	0.705 {2.035}	0.761 {2.202}	0.057 (0.170)
Age of primary business	6.451 {5.232}	6.647 {6.095}	0.196 (0.443)	6.595 {5.309}	6.732 {6.056}	0.137 (0.485)	6.392 {6.056}	6.888 {6.163}	0.496 (0.520)
Primary business weekly cashflow	4950 {5400}	5552 {6184}	602 (453)	4590 {4612}	5334 {5824}	744 * (424)	4540 {5824}	5259 {5803}	719 * (431)
Primary business location – commercial	0.305 {0.468}	0.385 {0.489}	0.080 ** (0.039)	0.322 {0.473}	0.384 {0.488}	0.062 (0.042)	0.319 {0.488}	0.387 {0.489}	0.068 (0.046)
Primary business location – residential	0.695 {0.468}	0.615 {0.489}	–0.080 ** (0.039)	0.678 {0.473}	0.616 {0.488}	–0.062 (0.042)	0.681 {0.488}	0.613 {0.489}	–0.068 (0.046)
Primary business arrangement – owned	0.625 {0.481}	0.609 {0.485}	–0.016 (0.040)	0.606 {0.484}	0.615 {0.482}	0.009 (0.043)	0.612 {0.482}	0.629 {0.477}	0.017 (0.047)
Primary business arrangement – rent	0.327 {0.453}	0.340 {0.463}	0.013 (0.037)	0.341 {0.456}	0.332 {0.458}	–0.009 (0.040)	0.335 {0.458}	0.319 {0.451}	–0.016 (0.044)
Primary business type – retail	0.235 {0.394}	0.235 {0.404}	0.000 (0.033)	0.230 {0.392}	0.221 {0.394}	–0.010 (0.035)	0.217 {0.394}	0.219 {0.389}	0.002 (0.037)
Primary business type – small grocery/convenience store	0.522 {0.499}	0.462 {0.500}	–0.060 (0.041)	0.509 {0.500}	0.474 {0.500}	–0.035 (0.044)	0.498 {0.500}	0.480 {0.500}	–0.018 (0.048)
Primary business type – wholesale	0.013 {0.153}	0.020 {0.162}	0.007 (0.013)	0.014 {0.165}	0.016 {0.156}	0.002 (0.015)	0.019 {0.156}	0.012 {0.145}	–0.007 (0.017)
Primary business type – service	0.093 {0.318}	0.134 {0.348}	0.040 (0.026)	0.093 {0.315}	0.135 {0.345}	0.042 (0.028)	0.107 {0.345}	0.148 {0.345}	0.041 (0.029)
Primary business type – manufacturing	0.015 {0.109}	0.022 {0.143}	0.007 (0.009)	0.017 {0.117}	0.025 {0.151}	0.008 (0.010)	0.007 {0.151}	0.019 {0.139}	0.011 (0.009)
Primary business type – vending	0.118 {0.302}	0.125 {0.322}	0.007 (0.025)	0.137 {0.324}	0.127 {0.324}	–0.010 (0.029)	0.153 {0.324}	0.120 {0.321}	–0.033 (0.033)
P-value on joint F-test: all RHS covariates listed = 0			0.4776			0.5932			0.4888
Attrition				0.142 {0.351}	0.155 {0.373}	0.013 (0.029)	0.268 {0.444}	0.326 {0.349}	0.058 (0.037)
Observations	1661	1661	1661	1388	1388	1388	1095	1095	1095

Notes: This table reports the mean of each of the listed variables for control and treatment groups during each of the three data collection phases. Standard deviation is reported in brackets below the mean. Columns 3, 6 & 9 provide the difference between the means, with robust standard errors in parentheses below, based off of a regression that includes dummies for bank and randomization window. The P-value is from a joint test of regressing all the variables on treatment assignment along with dummies for bank and randomization window. In this regression we omit the following categories in the relevant variables: 'Single' is the omitted marital status category. 'College graduate' is the omitted educational attainment variable. Commercial is the omitted primary business location variable. 'Own' is the omitted primary business property arrangement. 'Retail' is the omitted primary business type variable.

complete and includes questions on income, household composition, assets and liabilities, and business cash flows.

The banks added three questions on loan uses to their applications at our behest. The applicant was first asked: (1) Do you plan to spend 5000 pesos or more of your loan on any one household item? ¹⁵ (2) Do you plan to spend 2500 pesos or more of your loan on servicing any other debt? Later the applicant was asked to provide a full listing of intended usage of the loan. The former two questions are designed to identify non-trivial non-business uses of loan proceeds, keeping in mind that the median loan size is 10,000 pesos, and that

borrowers may split loan proceeds among several different types of expenditures. ¹⁶

This first step allows us to see how the applicants report their intended loan usage to the banks. These data will not be very informative about true intentions if applicants believe that their responses may affect the lender's decision. For example, applicants might reasonably infer that banks prefer to lend exclusively for business purposes, and

¹⁶ Appendix Fig. A1 shows the kernel density of requested loan amounts. The amounts asked about (5000 and 2500 pesos) were meant to be large enough to be remembered clearly and to avoid everyday purchases, but small enough to capture potential uses of loan proceeds.

¹⁵ Exchange rate at time of surveys was US\$1 = 43 Philippine Pesos.

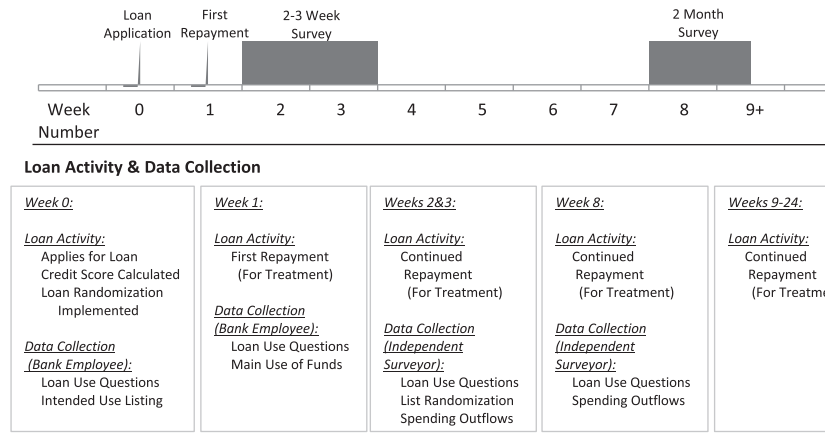


Fig. 1. Study timeline.

answer no to the questions about household and refinancing uses, regardless of their true intentions.

Table 2 Column 1 shows that very few applicants report non-businesses loan uses on their loan applications. Only 1.8% report planning to use their loan on a household transaction of 5000 pesos or more (Panel A), and only 2.3% report planning to use their loan to pay down debt of 2500 or more (Panel B).¹⁷ Column 1 shows results for the treatment group only, for comparability with subsequent analysis. Results do not change if we include the control group.

Is the low reported prevalence of non-business uses on loan applications driven by strategic underreporting? Results below from steps 3 and 4 suggest yes, although only to a point. Before detailing those results we examine whether borrowers change their reporting behavior to the bank after they obtain a loan.

3.4. Data collection step 2: at first loan repayment, by bank credit officer (direct elicitation)

The second pieces of data on loan uses come from a very short survey, administered by loan officers to a subset of borrowers, at the time of first repayment (about one week after loan disbursal). The loan officers asked two questions designed to parallel the key questions from the application: (1) Did you spend, or do you plan to spend, 5000 pesos or more of your loan on any one household item? (2) Did you spend, or do you plan to spend, 2500 pesos or more of your loan on servicing any other debt?

This step allows us to check for differences between what applicants and borrowers tell the bank. We might see such differences if applicants misreported strategically in the first step and the main driver of that behavior was concern about getting approved for the first loan.¹⁸ On the other hand, several factors push against finding differences, including repeat contracting, and any desire among borrowers to appear consistent in their reporting behavior.

We find that the reported prevalence of non-business uses post-loan is essentially unchanged from the loan application. Less than one percent report using their loan on a large household transaction, while

2.9% report using it to pay down other debt.¹⁹ This step was implemented only at one bank and only for a short period of time and so the sample size is small, with only 238 borrowers surveyed. This is because data collection proved onerous for the bank, and the bank discontinued it after we observed the strong similarity in reporting behavior between this step (post-loan) and step one (application).

3.5. Data collection steps 3 and 4: 2-week and 2-month surveys, by independent surveyor (all three methods)

The third and fourth pieces of data on loan uses come from two surveys, administered by an independent surveyor about two weeks and two months after loan application, of both treatment and control group individuals. Surveyors located individuals at their place of business or home and invited them to take a survey on behalf of Innovations for Poverty Action (IPA), a research organization. Surveyors were not aware of any connection to the banks. Surveyors informed people in the sample frame that IPA obtained a list of potential survey respondents from a database of local businesses.²⁰

Both surveys focus on direct elicitation of loan uses and the measurement of all recent substantial outflows, although the second survey is a bit shorter. Both were administered by the same surveyor. The scripts for key questions are reproduced in Appendix 1. Relative to the two-week survey, measuring outflows at two months has the potential advantage of allowing more time for all loan proceeds to be spent. It also has several potential disadvantages: more time for the control group to find alternative sources of financing (weakening power), a longer recall period (increasing measurement error), and/or more time for any short-run returns on investment to affect spending decisions (confounding inferences about the *direct* effect of borrowing on spending).

Eighty-five percent of our initial sample of 1661 completed the first (two-week) survey. Table 1 Column 6 shows that treatment assignment does not significantly affect two-week survey completion. There are small differences between the treatment and control groups, for instance the proportion who have graduated from high school is slightly higher in the treatment group, but a test of joint significance shows that baseline characteristics do not predict treatment status. Sixty-seven percent of our initial sample completed the second (two-month) survey. Table 1 Column 9 shows that treatment assignment does not significantly affect two-month survey completion. A test of joint significance also confirms

¹⁷ As we show in Section 2 of the paper our randomization was successful and so comparing the reported loan use intentions of the treatment and control group will not be informative at this point. The only place where comparing the responses is useful is in columns 5 and 7, reported spending.

¹⁸ It is worth considering how our inferences will be affected if borrowers change their mind over time about how to use loan proceeds. First, note that each of the independent surveys (two weeks and two months) does not suffer from this potential problem—each survey uses multiple elicitation methods, administered at the same time. Second, it is unclear why mind-changing would be asymmetric; e.g., for every person that changes their mind from a business use to a household use, we might well expect someone else to make the opposite change.

¹⁹ The loan officers also asked the borrowers what they primarily spent their loans on and every borrower replied that they spent it on their business.

²⁰ The goal was to be truthful yet also mask the relationship with the specific partnering bank. The surveyors themselves had no knowledge of the bank connection. It is possible that after multiple similar interactions some respondents could have made the connection between the bank's questions and the surveyor questions.

Table 2
Loan use elicitation methods.

Data Source:	Reported to Bank	Reported in Survey 2-3 Week Following Randomization				Reported in Survey 2 Months Following Randomization	
	(1)	(2)	(3)	(4)	(5)	(6)	
Specific survey question found in:	Appendix 1A	Appendix 1B	Appendix 1C	Appendix 1D	Appendix 1E	Appendix 1F	
Panel A: Household Expenditures:							
Will/Did you use 5,000 pesos or more of your loan on any single transaction for your household?							
Treatment	0.019 (0.004)	0.050 (0.007)	0.130 (0.066)	0.136 (0.011)	0.216 (0.013)	0.227 (0.013)	
Control				0.131 (0.028)		0.180 (0.035)	
Treatment - Control				0.005 (0.033)		0.046 (0.037)	
Sample Size	973	973	973	1095	973	1095	
Panel B: Payoff Other Debt:							
Will/Did you use 2,500 pesos or more of your loan to pay down other debt?							
Treatment	0.025 (0.005)	0.075 (0.008)	0.201 (0.055)	0.144 (0.011)	0.325 (0.015)	0.237 (0.014)	
Control				0.123 (0.030)		0.197 (0.036)	
Treatment - Control				0.021 (0.032)		0.041 (0.039)	
Observations from Treatment	973	973	973	973	973	973	
Observations from Control	0	0	0	122	0	122	

Notes: Column 1 considers our entire sample in treatment whether they were reached for the follow up survey or not. Columns 2 and 5 report the proportion responding “yes” to the direct question asked by the surveyor. If they didn’t report a loan they were coded as saying “no”. Column 3 looks at the difference in responses from a list randomization exercise on those in the treatment group that reported a loan, those that didn’t report a loan were coded as saying “no”. Columns 4 and 6 look at the actual spending reported by both treatment and control groups in each category and reports the difference between the two groups to take into account the fungibility of money. Results are robust to restricting the sample to those that completed all surveys, as well as including surveyor fixed effects. Robust standard errors in parentheses.

that baseline characteristics do not predict treatment status among those who responded to the two month survey.

The two-week survey begins with questions about basic demographics, health and savings. These introductory questions are designed to mitigate the likelihood that respondents infer any connection or association between the survey and their recent loan (application). The surveyor then asked the respondent for details on any outstanding loans, starting with the most recent one. Respondents reporting a loan were then asked about their deployment of loan proceeds using three different methods.

First, the surveyor explicitly asked the two key loan use questions: (1) Did you spend 5000 pesos or more of your loan on any one household item? (2) Did you spend 2500 pesos or more of your loan on servicing any other debt? We expect the proportion of “yeses” here to be higher than those reported to the bank, since incentives for strategic misreporting to an independent surveyor should be lower. Table 2 Column 2 shows that this is indeed the case, to some extent. 5.5% of individuals report using a loan for a large household expense; compared to 1.8% on the loan application (the 3.7 percentage point difference has a p-value <0.001, s.e. = 0.006). 7.7% report using the loan to pay down other debt, compared to 2.3% on the loan application (the 5.4 percentage point difference has a p-value <0.001, s.e. = 0.009). Of course, borrowers may still underreport non-business uses if such uses are stigmatized, or if borrowers suspect a connection between the surveyor and their bank. Such concerns motivate our second elicitation method.

Second, the surveyor administered a list randomization exercise to elicit estimates of group-level proportions of respondents using loan proceeds to pay down debt or buy household goods. List randomization is used across various disciplines to mitigate the underreporting of socially or financially sensitive information (Karlan and Zinman, 2012). The procedure asks a randomly-selected half of the respondents to report the total number of “yes” answers to four innocuous binary questions (Appendix 1), and the other half to report the total number of “yes”

answers to the same four innocuous binary questions plus a fifth sensitive one.²¹ We did this separately for the two different loan use questions: (1) “I spent over 5000 pesos of my loan of a single household transaction” and (2) “I spent more than 2500 pesos of my loan to pay down other debt.” We then estimate the proportion responding “yes” to the sensitive (loan use) question by subtracting the mean count of “yeses” for those who had only had the four innocuous questions from the mean count for those who had all five questions (including a loan use question). As expected, list randomization produces substantially higher estimates of non-business uses (Table 2 Column 3). We infer that 11.5% (s.e. = 5.6%) of respondents report spending at least 5000 pesos of their loan proceeds on a single household transaction, with 19.1% (s.e. = 4.9%) spending at least 2500 of their loan proceeds on paying down other debt.²²

All told, the results in Columns 1–3 suggest that elicitation method can have substantial effects on how borrowers report loan uses. Borrowers report more non-business uses when asked by an independent surveyor rather than a bank, and still more when they can report anonymously. The results suggest that list randomization, administered by an independent surveyor, produces relatively accurate estimates of how borrowers *perceive* their loan uses.

These results thus far do not address the question of how borrower perceptions accord with the reality that is most interesting to many

²¹ The baseline characteristics of the list randomization samples are balanced across treatment and control as reported in Table A1.

²² Those that respond positively to all items in the larger list of 5 questions may be less worried about anonymity as they are identifying themselves as having used the funds in the sensitive manner. Nonetheless, we continue to see under-reporting by this group: 7 out of the 1245 individuals surveyed answered “5” on the debt question. Of these 7, only 1 directly reported they used their loan on refinancing debt when asked. For the household question, 10 out of 1245 individuals answered “5”, of those only 3 directly reported that they used their loan on a large household expenditure. Additionally, this potential for a lack of anonymity would make it less likely for us to find an effect across elicitation methods.

researchers, practitioners, and funders: what did the respondent buy that they would not have in the absence of the marginal loan? Fungibility may make it difficult to construct survey questions that elicit that counterfactual. For example, loan proceeds may be used to purchase inventory in the *proximate* sense of cash from a lender being handed over to a supplier. But if the business owner would have purchased that inventory anyway, the marginal (counterfactual) purchase could be something else entirely; e.g., perhaps the cash flow that would have been used to purchase inventory is now used to purchase health care for an ailing family member.

The difficulty of identifying the counterfactual of interest motivates our third type of survey question: we ask each respondent to list each household and business outflow greater than 1000 pesos from the past two weeks (type and amount).²³ (Note the lack of any reference to loans or loan proceeds: this question asks about spending more broadly.) Together with the random assignment of loan approvals, we use responses to this question to identify the counterfactual: the impacts of the marginal loan on consumption and investment. We treat the results from this method of elicitation as close to the “truth” regarding where the funds went. Although this is an assumption, we find below that the average difference in spending between treatment and control lines up with the average increase in credit availability, lending credence to the idea that this method provides us insight into how marginal liquidity gets spent.

Table 2 Column 4 reports the results, which show a striking *lack* of impact on non-business spending. The treatment (loan approved) and control (loan rejected) groups have identical proportions (0.133) of respondents reporting one or more household expenses ≥ 5000 pesos, for a treatment effect of zero (s.e. = 0.030).²⁴ For debt pay down, the treatment group has a slightly higher proportion (0.142 vs. 0.126), but the 1.6 percentage point difference (s.e. = 0.029) is not statistically significant.²⁵ We find similar results, on a much higher base, in the two-month survey.²⁶ Regarding the base, many more respondents directly report non-business uses, whether directly (Column 5) or on the outflow list (Column 6).²⁷ Regarding the counterfactual of interest, when we compare the treatment group to the control group we find that the control group has an equally high base, statistically speaking. 22.7% of the treatment group report spending at least 5000 pesos on any one household transaction while 18.0% of the control group does so. This difference of 4.7 percentage points (s.e. = 0.037) is not statistically significant (p-value = 0.210). Similarly, 23.7% of the treatment group reports spending more than 2500 pesos on other debt²⁸ while 19.7% of the control group does so. This difference of 4.1 percentage points (s.e. = 0.039) is not statistically significant.

²³ Subjects were not asked to keep a diary of expenses or anything similar. They were simply asked to recall their spending over 1000 pesos over the requested time frame, without any prompts for specific expense types. This may lead to under-reporting of spending, which we would be concerned about if the treatment group was more likely to under-report since they have more spending in general due to the positive liquidity shock of the loan. This may bias us toward underestimating spending of the treatment group, especially on sensitive items. Since the estimated difference in spending between treatment and control is within the range of the additional liquidity provided by the lender, we believe this bias to be small.

²⁴ The cut-off at 5000 pesos allows us to check for large household expenditures and lines up with the direct questions that are asked of the borrowers. If we instead use a 1000 peso cut-off we get an increase of 0.026 in treatment (SE = 0.045).

²⁵ We are implicitly using the random assignment as an instrument for borrowing over the subsequent two weeks. The top rows of Table 3 confirm that the instrument is a powerful one; e.g., a treatment group member is 16 percentage points more likely to have a formal sector loan than a control group member.

²⁶ The higher base could be due to respondents taking more than 2 weeks to fully spend their loan proceeds, and/or to respondents' increased comfort with the survey or surveyor.

²⁷ We did not include list randomization on the two-month survey.

²⁸ It may seem peculiar that the proportion of respondents who report spending more than 2500 on debt pay down in the explicit question asked by the surveyor (column 6) is higher than the proportion that report this when listing out their spending over the past 2 months (column 7). This may be due to the fact that the outflow list has a 1000 peso threshold, so if someone pays off debt in increments < 1000 pesos but a total amount ≥ 2500 pesos, the outflow list would miss this, whereas that direct question might capture it.

When comparing results across different elicitation methods it is important to note that some confidence intervals are large, in particular for the estimates coming from the list randomization. For instance, while the point estimate for the proportion of individuals who indirectly admit to spending 5000 pesos or more of their loan on a single household transaction is (0.115), the confidence interval is (0.005, 0.225). At the same time, when we instead estimate the proportion of individuals who have spent more than 5000 pesos on a single household item using the data from their expenditure listing we get a point estimate of (0.000), with a confidence interval of (−0.059, 0.059). If we compare the point estimate derived from the expenditure listing to the confidence interval derived from the list randomization we find that the point estimate lies just outside of the confidence interval. If instead we compare the two confidence intervals we do find overlap. On the other hand, when we compare the confidence intervals in the debt repayment case where the list randomization confidence interval is (0.095, 0.287), and the expenditure listing confidence interval is (−0.041, 0.073) we do not find overlap.²⁹ Although some estimates are imprecise, overall we do have sufficient power to infer some differences between what people *report* they did with the loan (Column 2), with what they *perceive* they did with the loan (Column 3) and what they *actually* did with the loan (Column 4), to be informative.³⁰ We do not aim to put too much emphasis on any one estimate, but instead, to draw attention to how different methods provide different results.

Taken together, the results in Table 2 highlight several key findings. Substantively, there is little evidence of substantial non-business uses of microenterprise loans in this particular setting. This is surprising, given low impact on business growth in general from microcredit (Angelucci et al., 2015; Attanasio et al., 2015; Augsburg et al., 2015; Banerjee et al., 2015a,b; Crépon et al., 2015; Karlan and Zinman, 2010), findings from a prior study with one of the lenders here that marginal borrowers *decrease* investment in their microenterprises (Karlan and Zinman, 2011), and mounting concerns that people “over-borrow” to finance consumption (Zinman, 2014).

Methodologically, we find that borrower reporting responds strongly to the elicitation method, and that direct elicitation of loan uses does not produce credible evidence on a key counterfactual—what borrowers purchased that they would not have purchased in the absence of a loan. Rather, we identify the counterfactual using random assignment of credit access coupled with short-term follow-up measurement of substantial outflows.

3.6. So where does the money go?

If the marginal expenditure financed by a loan is *not* on a household item or other debt service (Table 2), it presumably is on some sort of business investment. Can we actually detect an increase in business investment, or do measurement error or reporting biases make it futile to attempt to follow the money with survey data?

Tables 3 and 4 suggest that our methods can in fact identify the marginal spending: business inventory, in this case. We switch from the mean comparisons in Table 2 (Columns 4 and 6) to regressions to improve precision, and estimate OLS intention-to-treat (ITT) models, with Huber–White standard errors, of the form:

$$Y_{it} = \alpha + \beta * treatment_i + \delta * FE_{it} + \varepsilon$$

where i indexes individuals, and t , time. The binary variable, $treatment = 1$ if i was randomly assigned to loan approval, and FE is a vector of randomization strata (a bank indicator, credit score category, application month-year, and the survey month-year). Y is an outcome

²⁹ This simple comparison of confidence intervals does not take into account any covariance between the different questions.

³⁰ These results hold when we restrict the sample to only those reached for both the two week and two month interview. Results are also robust to using surveyor fixed effects. See Tables A2 & A3.

Table 3

Actual loan use: treatment effects on short-term cash outflows (2–3 weeks after loan application).

Dependent variables	(0)	(1)	(2)	(3)
	Control group mean			
Has loan from experimenting lender (admin data)	0.336 (0.474)	0.329*** (0.042)	0.329*** (0.042)	0.329*** (0.042)
Any outstanding formal loan (survey data)	0.559 (0.498)	0.159*** (0.045)	0.159*** (0.045)	0.159*** (0.045)
Number of outstanding formal loans (survey data)	0.657 (0.672)	0.181*** (0.061)	0.166*** (0.061)	0.166*** (0.061)
Total outstanding formal loans (survey data)	6821 (20,706)	1535 (1919)	1725 (1119)	2644*** (788)
Total spending	19,208 (33,094)	5696* (3010)	5374** (2588)	4996** (2136)
Business expenditures	13,746 (20,145)	7031*** (2268)	6280*** (2104)	4523** (1985)
Assets for business	223 (1165)	356* (187)	137 (121)	–93 (94)
Merchandise for business	11,823 (19,365)	6045*** (2173)	5328*** (2013)	3738* (1914)
Business renovations	203 (1790)	120 (203)	–3 (30)	2 (2)
Utilities for business	594 (1562)	303 (252)	92 (119)	63 (98)
Salaries for employees	248 (1286)	159 (135)	102 (126)	0 (111)
Other business expenses	656 (2679)	48 (271)	–16 (228)	109 (146)
Household expenditures	4637 (25,753)	–1676 (1934)	–3 (413)	320 (317)
Household items	727 (2688)	–150 (248)	–38 (142)	27 (98)
Utilities for home	633 (1307)	7 (114)	23 (103)	169** (81)
Home renovation	70 (836)	–1815 (1887)	–79 (103)	–77 (71)
Education expenditure	550 (1776)	60 (174)	6 (165)	–112 (153)
Health expenditure	127 (618)	123 (88)	33 (64)	–42 (54)
Other personal expenses	316 (1284)	163 (151)	32 (106)	85 (75)
Debt repayment	769 (2404)	371 (284)	98 (223)	–59 (206)
Winsorized (1%)	N	N	Y	N
Trimmed (1%)	N	N	N	Y
Observations	143	1388	1388	1374

Notes: Each cell in columns 1–3 presents the treatment effect (i.e., the result on the variable 1 = assigned loan) from a separate regression. Each regression includes controls for the bank and credit scoring band (i.e., the probability of assignment to treatment), and application month and survey month fixed effects. Results are robust to not including the fixed effects. Column 2 winsorizes the dependent variables at the top 1%. Column 3 trims the dependent variables at the top 1%. All self-reported borrowing measures are stock measures at the time of the 2 week survey. Loan and Spending amounts are in Philippine Pesos. Spending amounts are all outflows over 1000 pesos over the 14 days before the survey. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

measuring borrowing (to show the magnitude of the first-stage) or spending, measured at either $t = 2$ weeks or $t = 2$ months post-random assignment. Because inferences about these outcomes may be influenced by outliers, we present results from three different functional forms: Column 1 estimates effects on the level of spending (in pesos); Column 2 “winsorizes” the data, recoding the top 1% of Y 's to the 99th percentile; and Column 3 “trims” the data, dropping observations in the top percentile of Y . We do not use $\log(Y)$ because most of our borrowing and spending variables have many zeros. (See Tables A1–A3.)

Table 3 shows treatment effects on different measures of Y over the two weeks after random assignment. Table 4 shows treatment

Table 4

Actual loan use: treatment effects on long-term cash outflows (2 months after loan application).

Dependent variables	(0)	(1)	(2)	(3)
	Control group mean			
Total spending	70,840 (143,029)	23,577 (17,046)	13,849 (13,643)	22,209** (8868)
Business expenditures	60,665 (138,764)	20,826 (16,518)	11,092 (13,295)	18,774** (8363)
Assets for business	320 (2098)	28 (229)	15 (154)	–45 (94)
Merchandise for business	56,496 (134,126)	19,726 (16,075)	9748 (13,094)	17,978** (8018)
Business renovations	943 (7257)	–561 (828)	–241 (168)	–83 (71)
Utilities for business	1246 (3184)	237 (382)	26 (235)	117 (174)
Salaries for employees	765 (3518)	584 (500)	195 (374)	–172 (316)
Other business expenses	896 (2669)	813 (525)	46 (274)	–160 (252)
Household expenditure	6837 (15,260)	699 (1746)	–63 (1204)	457 (901)
Household items	1303 (4884)	287 (503)	345 (349)	273 (275)
Utilities for home	1370 (2222)	–32 (225)	–47 (207)	30 (185)
Home renovation	1015 (7345)	1065 (1254)	–25 (284)	–196 (136)
Education expenditure	834 (2629)	386 (283)	288 (268)	147 (247)
Health expenditure	1565 (10,770)	–767 (874)	–43 (213)	–3 (132)
Other personal expenses	866 (3581)	164 (432)	2 (264)	17 (198)
Debt repayment	3240 (13,943)	1719 (1618)	622 (1087)	387 (775)
Winsorized (1%)	N	N	Y	N
Trimmed (1%)	N	N	N	Y
Observations	122	1095	1095	1084

Notes: Each cell in columns 1–3 presents the treatment effect (i.e., the result on the variable 1 = assigned loan) from a separate regression. Each regression includes controls for the bank and credit scoring band (i.e., the probability of assignment to treatment), and application month and survey month fixed effects. Results are robust to not including these fixed effects. Column 2 winsorizes the dependent variables at the top 1%. Column 3 trims the dependent variables at the top 1%. Spending amounts are in Philippine Pesos. Spending amounts are all outflows over 1000 pesos over the 2 months before the survey. The two-month survey did not ask about borrowing, administrative data about borrowing is the same data used in Table 3 and so not reported here but results are substantively equivalent. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

effects on the spending measures over the two months after random assignment.

The first panel of Table 3 shows that we have a strong first stage, similar to that found in Karlan and Zinman (2011) with the Metro Manila lender participating in this study. The treatment effect on the likelihood of having a loan from one of our partner banks is 0.33 (p-value <0.001, s.e. = 0.042). This is measured using administrative data from the bank. The effect is <1 due to approved applicants in the treatment group deciding to not actually go ahead with the loan, and to control group applicants who managed to avail a loan anyway. The remaining outcomes are measured using the follow-up surveys. Treatment effects on measures of total formal sector borrowing are still statistically significant but about one-half the size on borrowing from our partner lenders, due in part to some control group individuals obtaining credit from comparable lenders, and in part to substantial underreporting of debt that is line with what we have found in other studies (Karlan and Zinman, 2008; Karlan and Zinman, 2011; Zinman, 2009; Zinman, 2010).³¹

³¹ 34% of those we know, from administrative data, to have a loan with one of our lenders do not report any outstanding formal sector loans at the two-week follow-up survey.

The next panel of Table 3 estimates the treatment effect on total spending, as measured using our question asking respondents to list all outflows ≥ 1000 pesos during the past two weeks. Depending on our treatment of outliers, the estimate ranges from 4996 to 5696 pesos (with p-values of 0.059, 0.038, and 0.028 and standard errors of 3010, 2588, 2136 respectively). Scaling up these estimates by the difference in borrowing rates from the administrative data (since that data is not subject to underreporting of debt), we get estimated treatment-on-the-treated effects of about 15,000–16,000 pesos. The average loan size is 14,601 pesos, suggesting that our two-week outflow questions do successfully follow the money. They also suggest that borrowers spend all loan proceeds within the first two weeks, which seems plausible given the high interest rate and short maturity.

The rest of Table 3 disaggregates spending into several categories of interest. We confirm the lack of significant effects on household spending and debt pay down found in the earlier means comparisons (Table 2). Most notably, we find increases in business expenditures, in magnitudes commensurate with the treatment effect on overall spending. Disaggregating business expenses into fixed assets, inventory, renovations, utilities, salaries, and other, we find evidence suggesting that the entirety of the (business) spending increase is due to inventory. The ITT estimates on inventory range from 3738 to 6045 depending on how we treat outliers, with p-values of 0.005, 0.008, and 0.049 (s.e.s of 2173, 2013, and 1914 respectively). The focus on inventory may be due to the 3–6 month loan amortization, which may be too short for other types of investments to produce the returns needed to service the debt. We do not mean to suggest that this is a general result, i.e. that microfinance leads to borrowers spending their credit on inventory. We only mean to highlight the insights that can be gained from this type of data collection and analysis.

Table 4 repeats the spending analysis using data from the two-month follow-up survey. The results are qualitatively consistent with the two-week results.³² Point estimates are again more than large enough to offer a complete accounting of the loan proceeds. The pattern of results on spending (sub-)categories again suggests that about 100% of marginal spending is on business inventory. There are two noteworthy differences between the two-month and two-week results. One is that the two-month results are less precise. This is most likely due to the relative difficulty of recalling spending over a two month period. The second is that the two-month point estimates on total business expenditure, and inventory, specifically, are much larger. This could be an artifact of wide confidence intervals or respondent reporting. Or it could be capturing a true multiplier whereby treated individuals reinvest increased profits from the initial inventory increase, or obtain additional financing from other sources, to further increase inventory.

In any case, the suggestion that quantitative effect sizes may differ substantially over as short a period as six weeks—two weeks vs. two months—highlights the utility of short-run and high-frequency follow-ups for capturing and interpreting spending dynamics in the aftermath of a liquidity shock.

4. Conclusion

Discussions of outcome measurement following liquidity shocks often focus on how *longer-run* data may be needed to measure key impacts (e.g., of investments that require longer gestation periods, or learning). We take a different tack, and test three different methods for measuring the short-run responses.

The first method uses direct elicitation questions about intended loan usage on the banks' loan applications, shortly after loan disbursal, and nearly identical direct questions asked of borrowers, by

independent surveyors, with no link to the bank, two weeks and two months after loan disbursal. The second method uses indirect elicitation questions through two "list randomization" questions, asked by independent surveyors two weeks after disbursal, that make it feasible for respondents to respond truthfully to sensitive questions without actually revealing details about their behavior. The third method uses a counterfactual analysis via the lenders' randomizations and the two-week and two-month independent follow-up surveys, by comparing a listing of recent expenditures (with no reference to recent borrowing) across the treatment and control groups.

The results suggest three key findings in our setting. First, respondents report strategically. They report very little non-business uses of loan proceeds to the bank, significantly more to independent surveyors when asked direct questions, and yet significantly more to independent surveyors when presented with lists of statements that allow them to report what they believe to be the truth without directly revealing what they spent. Second, even when borrowers are more likely to respond with what they *perceive* to be the truth, their answers to questions about "did you spend X or more of your loan on..." are different than the counterfactual of greatest interest to economists and policymakers, where additional expenditure was actually spent. Third, we estimate that the treatment effect is actually entirely on business investment, specifically inventory. This treatment effect can account for the entire loan amount 2-weeks post-randomization, with even larger but more noisily estimated effects at 2-months post-randomization.

We believe the main implication of our results is methodological: researchers should consider collecting spending data on both treatment and control subjects very shortly after an exogenous liquidity shock. In particular, our study highlights the value of *shorter-run, high-frequency* data collection on substantial outflows following a liquidity shock. To take just two examples, if we are interested in the possibility of over-borrowing, the methods used in this paper can be used to address the question of "over-borrowing on what"? In the settings studied here, the answer appears to be "not on consumption". If we are interested in why expanding access to microcredit does not reliably lead to business growth and increased profits, the methods here can be used to address the question "is this because borrowers invest in something else, or because they invest and fail?" In the settings studied in this paper it appears that any downstream lack of business growth is not for lack of trying.

Future work can build on this by combining estimates of liquidity shock spending responses with longer-term follow-up data. In our context, this will enable us to measure whether the short-run investments in inventory produce long-run increases in profits and/or improvements in household outcomes. Another important question is whether alternative direct (and indirect) elicitation methods might help borrowers and researchers zero in on the key counterfactual, and encourage future work in this area. Perhaps asking "what did you spend your loan on that you would not have bought if you had not gotten a loan?" would produce the same inferences, at less expense, than a randomized experiment followed by elicitation of all major household and business outflows.

Appendix 1. Survey questions

1A – Bank Interaction

Panel A: Will you use 5000 pesos or more of your loan on any single transaction for your household?

Panel B: Will you use 2500 pesos or more of your loan to pay down other debt?

1B – 1st Loan Payment & 2 Week Survey

Panel A: Did you use 5000 pesos or more of your loan on any single transaction for your household?

Panel B: Did you use 2500 pesos or more of your loan to pay down other debt?

³² When we compare the estimates from the two-week survey and the two-month survey we are comparing slightly different samples. If we restrict the two-week survey to just those that are also in the two-month survey we find similar, but less precise results, as our sample size decreases.

1C – List Randomization

Panel A: Short Version:

As with our example, I will now read four statements. I would like you to tell me how many are true for you, but do not tell me which ones are true.

1. I have a washing machine in my home.
2. I am originally from this city.
3. I have completed one year or more of formal education post-high school.
4. My household owns a computer.

Long Version:

As with our example, I will now read five statements. I would like you to tell me how many are true for you, but do not tell me which ones are true.

1. I have a washing machine in my home.
2. I am originally from this city.
3. I have completed one year or more of formal education post-high school.
4. My household owns a computer.
5. I used 5000 pesos or more of my loan on any single transaction for my household.

Table A1

Orthogonality of list randomization group to applicant characteristics and attrition.

Demographic variables	Debt repayment list			HH expenditure list		
	Short	Long	Difference	Short	Long	Difference
	(1)	(2)	(3)	(4)	(5)	(6)
Male	0.111 {0.386}	0.137 {0.369}	0.026 (0.022)	0.122 {0.379}	0.128 {0.377}	0.006 (0.022)
Age	42.224 {9.04}	41.540 {9.28}	−0.683 (0.53)	41.920 {9.16}	41.774 {9.16}	−0.145 (0.53)
Marital status – single	0.084 {0.306}	0.101 {0.281}	0.017 (0.017)	0.111 {0.273}	0.078 {0.316}	−0.033 * (0.017)
Marital status – married	0.734 {0.440}	0.714 {0.425}	−0.019 (0.025)	0.706 {0.426}	0.737 {0.440}	0.031 (0.025)
Marital status – widowed/separated	0.130 {0.293}	0.110 {0.315}	−0.020 (0.018)	0.122 {0.301}	0.116 {0.307}	−0.007 (0.017)
Education – college	0.254 {0.452}	0.244 {0.459}	−0.009 (0.026)	0.238 {0.459}	0.257 {0.451}	0.020 (0.026)
Education – some college	0.248 {0.453}	0.296 {0.427}	0.048 * (0.026)	0.259 {0.449}	0.288 {0.433}	0.029 (0.026)
Education – graduated high school	0.425 {0.470}	0.392 {0.478}	−0.033 (0.027)	0.428 {0.467}	0.388 {0.480}	−0.040 (0.027)
Education – some high school or less	0.072 {0.298}	0.068 {0.308}	−0.004 (0.018)	0.076 {0.295}	0.065 {0.311}	−0.011 (0.017)
Number of dependents	1.838 {1.465}	1.935 {1.446}	0.097 (0.085)	1.785 {1.499}	1.987 {1.404}	0.201 ** (0.084)
Number of employees	0.743 {2.368}	0.737 {1.647}	−0.006 (0.123)	0.716 {2.440}	0.761 {1.555}	0.045 (0.121)
Age of primary business	6.831 {6.150}	6.690 {6.090}	−0.141 (0.358)	7.001 {5.940}	6.533 {6.304}	−0.468 (0.355)
Primary business weekly cashflow	5220 {5608}	5192 {5879}	−29 (339)	5300 {5100}	5118 {6351}	−182 (340)
Primary business location – commercial	0.362 {0.491}	0.392 {0.484}	0.030 (0.028)	0.372 {0.489}	0.385 {0.487}	0.013 (0.028)
Primary business location – residential	0.638 {0.491}	0.608 {0.484}	−0.030 (0.028)	0.628 {0.489}	0.615 {0.487}	−0.013 (0.028)
Primary business arrangement – owned	0.621 {0.484}	0.607 {0.479}	−0.014 (0.028)	0.611 {0.482}	0.614 {0.482}	0.003 (0.028)
Primary business arrangement – rent	0.314 {0.465}	0.347 {0.449}	0.033 (0.027)	0.318 {0.464}	0.345 {0.451}	0.027 (0.027)
Primary business type – retail	0.214 {0.408}	0.241 {0.383}	0.027 (0.023)	0.221 {0.404}	0.236 {0.389}	0.015 (0.023)
Primary business type – small grocery/convenience store	0.487 {0.500}	0.462 {0.500}	−0.026 (0.029)	0.480 {0.500}	0.467 {0.500}	−0.013 (0.029)
Primary business type – wholesale	0.011 {0.152}	0.013 {0.151}	0.002 (0.009)	0.005 {0.169}	0.018 {0.131}	0.013 (0.009)
Primary business type – service	0.130 {0.351}	0.138 {0.342}	0.008 (0.020)	0.132 {0.350}	0.137 {0.343}	0.004 (0.020)
Primary business type – manufacturing	0.031 {0.147}	0.025 {0.162}	−0.005 (0.009)	0.030 {0.149}	0.026 {0.160}	−0.003 (0.009)
Primary business type – vending	0.128 {0.318}	0.120 {0.328}	−0.008 (0.019)	0.131 {0.314}	0.117 {0.332}	−0.014 (0.019)
P-value on joint F-test: all RHS covariates listed = 0			0.4987			0.8491
Observations	556	632	1188	572	616	1188

Notes: This table reports the mean of each of the listed variables for those who receive the short (4 statements) and long (5 statements) lists for the list randomization. This considers sample for Table 2 – those in the treatment group who admit to having an active loan. Standard deviation is reported in brackets below the mean. Columns 3 & 6 provide the difference between the means, with robust standard errors in parentheses below, based off of a regression that includes dummies for bank and randomization window. The P-value is from a joint test of regressing all the variables on list assignment along with dummies for bank and randomization window. In this regression we omit the following categories in the relevant variables: ‘Single’ is the omitted marital status category. ‘College graduate’ is the omitted educational attainment variable. Commercial is the omitted primary business location variable. ‘Own’ is the omitted primary business property arrangement. ‘Retail’ is the omitted primary business type variable.

Table A2

Loan use elicitation methods (restricting to those reached for all surveys).

Data Source:	Reported to Bank	Reported in Survey 2-3 Week Following Randomization			Reported in Survey 2 Months Following Randomization	
	(1)	(2)	(3)	(4)	(5)	(6)
	Proportion reporting "yes" on loan application	Proportion reporting "yes" in direct self-report to independent surveyor	Implicit proportion reporting "yes" from list randomization.	Proportion reporting "yes" for household (Panel A) or debt (Panel B) outflows	Proportion reporting "yes" in direct self-report to independent surveyor	Proportion reporting "yes" for household (Panel A) or debt (Panel B) outflows
Specific survey question found in:	Appendix 1A	Appendix 1B	Appendix 1C	Appendix 1D	Appendix 1E	Appendix 1F
Panel A: Household Expenditures:						
Will/Did you use 5,000 pesos or more of your loan on any single transaction for your household?						
Treatment	0.018 (0.003)	0.055 (0.006)	0.096 (0.057)	0.132 (0.009) 0.134 (0.028)	0.216 (0.013)	0.228 (0.013) 0.178 (0.037)
Control						
Treatment - Control				-0.002 (0.030)		0.050 (0.037)
Sample Size	1493	1245	1245	1388	973	1095
Panel B: Payoff Other Debt:						
Will/Did you use 2,500 pesos or more of your loan to pay down other debt?						
Treatment	0.023 (0.004)	0.077 (0.008)	0.193 (0.049)	0.142 (0.010) 0.128 (0.028)	0.325 (0.015)	0.239 (0.014) 0.185 (0.036)
Control						
Treatment - Control				0.014 (0.030)		0.054 (0.039)
Observations from Treatment	1493	1245	1245	1245	973	973
Observations from Control	0	0	0	143	0	122

Notes: This table mimics Table 2, but restricts the sample to only those who were contacted for all surveys. Columns 2 and 5 report the proportion responding "yes" to the direct question asked by the surveyor. If they didn't report a loan they were coded as saying "no". Column 3 looks at the difference in responses from a list randomization exercise on those in the treatment group that reported a loan, those that didn't report a loan were coded as saying "no". Columns 4 and 6 look at the actual spending reported by both treatment and control groups in each category and reports the difference between the two groups to take into account the fungibility of money. Robust standard errors in parentheses.

Table A3

Loan use elicitation methods (with surveyor fixed effects).

Data Source:	Reported to Bank	Reported in Survey 2-3 Week Following Randomization			Reported in Survey 2 Months Following Randomization	
	(1)	(2)	(3)	(4)	(5)	(6)
	Proportion reporting "yes" on loan application	Proportion reporting "yes" in direct self-report to independent surveyor	Implicit proportion reporting "yes" from list randomization.	Proportion reporting "yes" for household (Panel A) or debt (Panel B) outflows	Proportion reporting "yes" in direct self-report to independent surveyor	Proportion reporting "yes" for household (Panel A) or debt (Panel B) outflows
Specific survey question found in:	Appendix 1A	Appendix 1B	Appendix 1C	Appendix 1D	Appendix 1E	Appendix 1F
Panel A: Household Expenditures:						
Will/Did you use 5,000 pesos or more of your loan on any single transaction for your household?						
Treatment	0.018 (0.003)	0.055 (0.006)	0.115 (0.056)	0.133 (0.009) 0.133 (0.028)	0.216 (0.013)	0.227 (0.013) 0.180 (0.035)
Control						
Treatment - Control				0.000 (0.030)		0.046 (0.037)
Sample Size	1493	1245	1245	1388	973	1095
Panel B: Payoff Other Debt:						
Will/Did you use 2,500 pesos or more of your loan to pay down other debt?						
Treatment	0.023 (0.004)	0.077 (0.008)	0.191 (0.049)	0.142 (0.010) 0.126 (0.028)	0.325 (0.015)	0.237 (0.014) 0.197 (0.036)
Control						
Treatment - Control				0.016 (0.029)		0.041 (0.039)
Observations from Treatment	1493	1245	1245	1245	973	973
Observations from Control	0	0	0	143	0	122

Notes: This table mimics Table 2 but include surveyor fixed effects. Column 1 considers our entire sample in treatment whether they were reached for the follow up survey or not. Columns 2 and 5 report the proportion responding "yes" to the direct question asked by the surveyor. If they didn't report a loan they were coded as saying "no". Column 3 looks at the difference in responses from a list randomization exercise on those in the treatment group that reported a loan, those that didn't report a loan were coded as saying "no". Columns 4 and 6 look at the actual spending reported by both treatment and control groups in each category and reports the difference between the two groups to take into account the fungibility of money. Results are robust to restricting the sample to those that completed all surveys. Robust standard errors in parentheses.

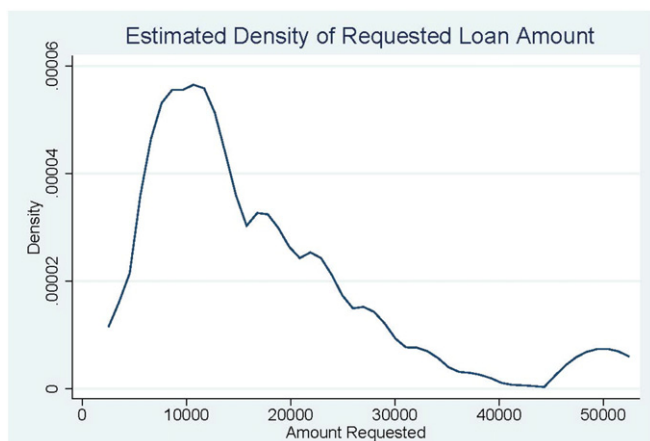


Fig. A1. Kernel density of requested loan amount.

Panel B: Short Version:

As with our example, I will now read four statements. I would like you to tell me how many are true for you, but do not tell me which ones are true

1. I have visited a hospital or clinic in the last six months.
2. I have more than three siblings.
3. I have purchased some type of insurance in the past five years.
4. My household owns an air conditioner.

Long Version:

As with our example, I will now read five statements. I would like you to tell me how many are true for you, but do not tell me which ones are true

1. I have visited a hospital or clinic in the last six months.
2. I have more than three siblings.
3. I have purchased some type of insurance in the past five years.
4. My household owns an air conditioner.
5. I used 2500 pesos or more of my loan to pay down other debt.

1D – 2 Week Survey

Please list all transactions of 1000 pesos or more that you have made in the last 14 days. List each item with the amount that you spent.

1E – 2 Month Survey

Panel A: In the past two months, did you spend 5000 pesos or more on any single transaction for your household?

Panel B: In the past two months, did you spend 2500 pesos or more to pay down debt?

1F – 2 Month Survey

Please list all transactions of 1000 pesos or more that you have made in the last two months. List each item with the amount that you spent.

References

Agarwal, Sumit, Liu, Chunlin, Souleles, Nicholas S., 2007. The reaction of consumer spending and debt to tax rebates—evidence from consumer credit data. *J. Polit. Econ.* 115 (6), 986–1019. <http://dx.doi.org/10.1086/528721>.

- Anagol, Santosh, Etang, Alvin, Karlan, Dean, 2013. Continued existence of cows disproves central tenets of capitalism? *Natl. Bur. Econ. Res.* 19437 (September) (<http://www.nber.org/papers/w19437>).
- Angelucci, Manuela, Karlan, Dean, Zinman, Jonathan, 2015. Microcredit impacts: evidence from a randomized microcredit program placement experiment by compartamos banco. *Am. Econ. J. Appl. Econ.* 7 (1), 151–182. <http://dx.doi.org/10.1257/app.20130537>.
- Attanasio, Orazio, Augsburg, Britta, De Haas, Ralph, Fitzsimons, Emla, Harmgart, Heike, 2015. The impacts of microfinance: evidence from joint-liability lending in Mongolia. *Am. Econ. J. Appl. Econ.* 7 (1), 90–122. <http://dx.doi.org/10.1257/app.20130489>.
- Augsburg, Britta, De Haas, Ralph, Harmgart, Heike, Meghir, Costas, 2015. The impacts of microcredit: evidence from Bosnia and Herzegovina. *Am. Econ. J. Appl. Econ.* 7 (1), 183–203. <http://dx.doi.org/10.1257/app.20130272>.
- Banerjee, Abhijit, Duflo, Esther, Glennerster, Rachel, Kinnan, Cynthia, 2015a. The miracle of microfinance? Evidence from a randomized evaluation. *Am. Econ. J. Appl. Econ.* 7 (1), 22–53. <http://dx.doi.org/10.1257/app.20130533>.
- Banerjee, Abhijit, Karlan, Dean, Zinman, Jonathan, 2015b. Six randomized evaluations of microcredit: introduction and further steps. *Am. Econ. J. Appl. Econ.* 7 (1), 1–21. <http://dx.doi.org/10.1257/app.20140287>.
- Benhassine, Najy, Devoto, Florencia, Duflo, Esther, Dupas, Pascaline, Pouliquen, Victor, 2015. Turning a shove into a nudge? A 'labeled cash transfer' for education. *Am. Econ. J. Econ. Policy* 7.3, 86–125.
- Blattman, Christopher, Fiala, Nathan, Martinez, Sebastian, 2012. Employment generation in rural Africa: mid-term results from an experimental evaluation of the youth opportunities program in northern Uganda. SSRN Scholarly Paper ID 2030866.
- Crépon, Bruno, Devoto, Florencia, Duflo, Esther, Pariente, William, 2015. Estimating the impact of microcredit on those who take It Up: evidence from a randomized experiment in Morocco. *Am. Econ. J. Appl. Econ.* 7 (1), 123–150. <http://dx.doi.org/10.1257/app.20130535>.
- De Mel, Suresh, McKenzie, David, Woodruff, Christopher, 2008. Returns to capital in microenterprises: evidence from a field experiment. *Q. J. Econ.* 123 (4), 1329–1372.
- Eggertsson, Gauti B., Krugman, Paul, 2012. Debt, deleveraging, and the liquidity trap: a Fisher-Minsky-Koo approach. *Q. J. Econ.* 127 (3), 1469–1513.
- Fafchamps, Marcel, McKenzie, David, Quinn, Simon, Woodruff, Christopher, 2014. Micro-enterprise growth and the flypaper effect: evidence from a randomized experiment in Ghana. *J. Dev. Econ.* 106 (January), 211–226. <http://dx.doi.org/10.1016/j.jdeveco.2013.09.010>.
- Groh, Matthew, McKenzie, David, 2014. Macroinsurance for microenterprises: a randomized experiment in post-revolution Egypt. SSRN Scholarly Paper ID 2502229. Social Science Research Network, Rochester, NY (<http://papers.ssrn.com/abstract=2502229>).
- Hall, Robert E., 2011. The long slump. *Am. Econ. Rev.* 101 (2), 431–469. <http://dx.doi.org/10.1257/aer.101.2.431>.
- Haushofer, Johannes, Shapiro, Jeremy, 2013. Welfare effects of unconditional cash transfers: evidence from a randomized controlled trial in Kenya. M.I.T. Working Paper (http://www.princeton.edu/~joha/publications/Haushofer_Shapiro_UCT_2013.pdf).
- Johnson, David S., Parker, Jonathan A., Souleles, Nicholas S., 2006. Household expenditure and the income tax rebates of 2001. *Am. Econ. Rev.* 96 (5), 1589–1610. <http://dx.doi.org/10.1257/aer.96.5.1589>.
- Karlan, Dean, Zinman, Jonathan, 2008. Lying about borrowing. *J. Eur. Econ. Assoc.* 6 (2–3), 510–521.
- Karlan, Dean, Zinman, Jonathan, 2010. Expanding credit access: using randomized supply decisions to estimate the impacts. *Rev. Financ. Stud.* 23 (1), 433–464.
- Karlan, Dean, Zinman, Jonathan, 2011. Microcredit in theory and practice: using randomized credit scoring for impact evaluation. *Science* 332 (6035), 1278–1284. <http://dx.doi.org/10.1126/science.1200138>.
- Karlan, Dean, Zinman, Jonathan, 2012. List randomization for sensitive behavior: an application for measuring use of loan proceeds. *J. Dev. Econ.* 98 (1), 71–75. <http://dx.doi.org/10.1016/j.jdeveco.2011.08.006>.
- Karlan, Dean, Knight, Ryan, Udry, Christopher, 2013. Consulting and capital experiments with micro and small tailoring enterprises in Ghana. *Working Paper*.
- Mian, Atif, Sufi, Amir, 2011. House prices, home equity-based borrowing, and the US household leverage crisis. *Am. Econ. Rev.* 101 (5), 2132–2156.
- Moskowitz, T.J., Vissing-Jorgensen, A., 2002. The returns to entrepreneurial investment: a private equity premium puzzle. *Am. Econ. Rev.* 92 (4), 745–778.
- Nicola, De, Francesca, Giné, Xavier, 2014. How accurate are recall data? Evidence from coastal India. *J. Dev. Econ.* 106, 52–65.
- Parker, Jonathan A., Souleles, Nicholas S., Johnson, David S., McClelland, Robert, 2013. Consumer spending and the economic stimulus payments of 2008. *Am. Econ. Rev.* 103 (6), 2530–2553. <http://dx.doi.org/10.1257/aer.103.6.2530>.
- Sahm, Claudia R., Shapiro, Matthew D., Slemrod, Joel, 2012. Check in the mail or more in the paycheck: does the effectiveness of fiscal stimulus depend on how it is delivered? *Am. Econ. J. Econ. Policy* 4 (3), 216–250. <http://dx.doi.org/10.1257/pol.4.3.216>.
- Shapiro, Matthew D., Slemrod, Joel, 2003. Consumer response to tax rebates. *Am. Econ. Rev.* 93 (1), 381–396. <http://dx.doi.org/10.1257/000282803321455368>.
- Shapiro, Matthew D., Slemrod, Joel, 2009. Did the 2008 tax rebates stimulate spending? *Am. Econ. Rev.* 99 (2), 374–379. <http://dx.doi.org/10.1257/aer.99.2.374>.
- Souleles, Nicholas S., 1999. The response of household consumption to income tax refunds. *Am. Econ. Rev.* 89 (4), 947–958. <http://dx.doi.org/10.1257/aer.89.4.947>.
- Souleles, Nicholas S., 2002. Consumer response to the Reagan tax cuts. *J. Public Econ.* 85 (1), 99–120.

- Tarozzi, Alessandro, Desai, Jaikishan, Johnson, Kristin, 2015. The impacts of microcredit: evidence from Ethiopia. *Am. Econ. J. Appl. Econ.* 7 (1), 54–89. <http://dx.doi.org/10.1257/app.20130475>.
- Zinman, Jonathan, 2009. Where is the missing credit card debt? Clues and implications. *Rev. Income Wealth* 55 (2), 249–265.
- Zinman, Jonathan, 2010. Restricting consumer credit access: household survey evidence on effects around the Oregon rate cap. *J. Bank. Financ.* 34 (3), 546–556.
- Zinman, Jonathan, 2014. Consumer credit: too much or too little (or just right)? *J. Leg. Stud.* 43 (S2 Special Issue on Benefit-Cost Analysis of Financial Regulation), S209–S237.