Update to the pre-analysis plan for "Paying for urban services: utility bills, prepaid metering and spending patterns of the poor"

PIs: Kelsey Jack, Kathryn McDermott, Anja Sautmann
Date: October 11, 2020

Update: Below, we specify how we construct the indices for the analysis of heterogeneous treatment effects. The previously published pre-analysis plan did not provide full detail on how the indices would be created from the specified survey questions. We are posting this update prior to conducting any analysis of heterogeneous treatment effects.

Following the original pre-analysis plan, we will conduct the heterogeneity analysis in two steps.

First, we will test for heterogeneity in the response to Round 1 treatment assignment. We examine heterogeneity with respect to pre-treatment purchase frequency (split at the sample median), and with respect to four binary indices associated with our two alternative hypotheses, described in detail below. The indices cover the household's liquidity, the transaction costs associated with purchasing electricity, the degree to which the decision maker exhibits sophistication about intra-household control issues, and the degree to which there is intra-household coordination of consumption.

Second, we will test for heterogeneity in the willingness to pay for different delivery forms of the transfer (cash, one electricity voucher, or two vouchers) measured using Round 2 choices, *conditional* on observing heterogeneous treatment effects consistent with either of our two alternative hypotheses in Round 1 treatment effects. We will not proceed to analysis of heterogeneity in willingness to pay if we do not observe such heterogeneity.

We will estimate heterogeneous treatment effects by interacting Round 1 treatments with binary index variables, using daily data and a post-treatment window informed by a dynamic specification estimated using the whole sample (i.e., not including heterogeneity variables):

$$y_{ijt} = \sum_j \beta_j treatment_{ijt} + \sum_j \gamma_j treatment_{ijt} \times 1(index_i) + \eta_i + \tau_t + \epsilon_{ijt}. \quad (1)$$

Our outcomes of interest, y_{ijt}, include (a) expenditures in ZAR, (b) kWh purchased, and (c) transactions per day. τ are time (day or week) fixed effects that are allowed to vary with geographic strata, and η_i are household fixed effects.

Table 1: Liquidity Constraint Index (Summary)

	N	Mean	Stdev	Minimum	Maximum
Unpredictable income	795	0.366	0.482	0.000	1.000
Cash liquidity	750	0.428	0.495	0.000	1.000
Days liquidity	583	0.448	0.498	0.000	1.000
Non-missing inputs	795	2.677	0.504	1.000	3.000
Liquidity constraint index	795	0.392	0.324	0.000	1.000

Notes: The Liquidity Constraint index was created from the following components: a dummy variable indicating if the household has unpredictable income (unpredictable income indicates constrained liquidity), the amount of cash the household could raise in an emergency (low amounts indicate constrained liquidity), and the number of days it would take the household to raise R1000 in an emergency (high numbers indicate constrained liquidity). Non-binary variables were converted to binary variables based on the median value and coded so that 1 indicates constrained liquidity. The index for the individual household was created as the sum of the non-missing binary variables, divided by the number of non-missing inputs.

Table 2: Liquidity Constraint Index (Correlation)

	Unpredictable income	Cash liquidity	Days liquidity	Non-missing inputs
Unpredictable income	1.000			
Cash liquidity	-0.001	1.000		
Days liquidity	0.051	0.328	1.000	
Non-missing inputs	0.026	0.420	-0.063	1.000

Notes: Pairwise correlations are reported among the components of the index.

Table 3: Transaction Cost Index (Summary)

	N	Mean	Stdev	Minimum	Maximum
Cannot make a night purchase	794	0.401	0.490	0.000	1.000
Unsafe	794	0.422	0.494	0.000	1.000
Inconvenient location	795	0.225	0.418	0.000	1.000
Vendor fee	788	0.330	0.470	0.000	1.000
Non-missing inputs	795	3.989	0.106	3.000	4.000
Transaction costs index	795	0.344	0.247	0.000	1.000

Notes: The Transaction Cost index was created from the following components: dummy variables indicating whether the individual would be unwilling to purchase electricity late at night if the meter ran out (high TC), whether the neighborhood is unsafe (high TC), and whether the electricity vendor charges a fee, requires an accompanying purchase or limits the amount of the electricity purchase (high TC). The index for the individual household was created as the sum of the non-missing binary variables, divided by the number of non-missing variable inputs.

Table 4: Transaction Cost Index (Correlation)

	Cannot make a night purchase	Unsafe	Inconvenient location	Vendor fee	Non-missing inputs
Cannot make a night purchase	1.000				
Unsafe	0.217	1.000			
Inconvenient location	0.109	0.027	1.000		
Vendor fee	-0.041	0.164	-0.316	1.000	
Non-missing inputs	0.031	0.035	0.029	0.035	1.000

Notes: Pairwise correlations are reported among the components of the index.

Table 5: Sophisticated Control Index (Summary)

	N	Mean	Stdev	Minimum	Maximum
Disconnect self	795	0.057	0.231	0.000	1.000
Disconnected others	795	0.053	0.224	0.000	1.000
Non-missing inputs	795	2.000	0.000	2.000	2.000
Sophisticated control index	795	0.055	0.197	0.000	1.000

Notes: The Sophisticated Control index indicates whether a household displays sophistication about self- or other-control problems (and has these problems). The index was created based on two dummy variables indicating whether the decision maker intentionally ran out of electricity to control either their own expenditure or that of other household members (a value of 1 indicates greater sophistication/control problems). The index for the individual household was created as the sum of the non-missing binary variables, divided by the number of non-missing variable inputs.

Table 6: Sophisticated Control Index (Correlation)

	Disconnect self	Disconnected others	Non-missing inputs
Disconnect self	1.000		
Disconnected others	0.502	1.000	
Non-missing inputs	.	.	.

Notes: Pairwise correlations are reported among the components of the index.

Table 7: Intrahoushold Coordination Failure Index (Summary)

	N	Mean	Stdev	Minimum	Maximum
Out of electricity	795	0.348	0.477	0.000	1.000
Share w/BY	795	0.340	0.474	0.000	1.000
Waste elect	787	0.219	0.414	0.000	1.000
Free riders (buy)	731	0.386	0.487	0.000	1.000
Free riders (money)	706	0.421	0.494	0.000	1.000
Non-missing inputs	795	4.797	0.568	3.000	5.000
Intrahousehold coordination failure index	795	0.338	0.249	0.000	1.000

Notes: The Intrahousehold Coordination Failure index was created from the following components: the number of times the household ran out of electricity in the past week, dummy variables indicating whether the household shares the electricity meter with backyard dwellers and whether the individual thinks others in the household waste electricity (a value of 1 in each indicating higher IC issues/poor coordination), the number of meter users who do not purchase electricity and the number of meter users who do not contribute money to electricity purchases (higher numbers in each indicate greater IC issues). Non-binary variables were converted to binary variables based on the median value such that 1 indicates poor coordination. The index for the individual household was created as the sum of the non-missing binary variables, divided by the number of non-missing variable inputs.

Table 8: Intrahousehold Coordination Failure Index (Correlation)

	Out of electricity	Share w/BY	Waste elect	Free riders (buy)	Free riders (money)	Non-missing inputs
Out of electricity	1.000					
Share w/BY	0.139	1.000				
Waste elect	0.068	0.033	1.000			
Free riders (buy)	0.062	-0.111	0.050	1.000		
Free riders (money)	0.130	-0.108	0.103	0.470	1.000	
Non-missing inputs	0.014	0.050	-0.074	0.091	0.010	1.000

Notes: Pairwise correlations are reported among the components of the index.