Changing aspirations through poverty measurement: The Poverty Stoplight Program

Working Paper, April 15, 2018

Katharina Hammler[1] / Martin Burt[2]

Abstract

Fundación Paraguaya's (FP) Poverty Stoplight (PS) is a multidimensional poverty measurement tool and mentoring approach aiming to empower people to lift themselves out of poverty by changing their aspirations. We use FP's administrative database containing information from over 8,900 micro finance clients to evaluate whether the PS program is effective in helping program participants overcome poverty. We argue that combining the PS, which is a dashboard metric, with the Alkire-Foster (AF) methodology provides advantages for different types of users. Using the AF methodology, we construct two multidimensional poverty indices based on the 50 indicators of the PS, one of which capturing moderate poverty, the other one extreme poverty. Based on the results of OLS and instrumental variable estimation, we find evidence that participation in the PS program is indeed associated with a decrease in poverty.

1 Background

Fundación Paraguaya (FP) is Paraguay's largest non-governmental developmental organization. FP works in the areas of microfinance, entrepreneurship, financial literacy, and self-sufficient vocational education, with the overarching goal of eliminating multidimensional poverty, both within Paraguay, and in many other countries where FP has

[1] Director of Monitoring and Evaluation at Fundación Paraguaya, khammler@fundacionparaguaya.org.py
[2] Founder and CEO of Fundación Paraguaya, burt@fundacionparaguaya.org.py
We thank Ana Vaz, Christoph Jindra, Bilal Malaeb, the participants of the OPHI Multidimensional Impact Evaluation Workshop 2018, Michael Walton, and Ana Revenga for their valuable comments on earlier drafts of this paper.

been active. To support its work, FP developed the Poverty Stoplight (PS, "the Stoplight") in 2011. The PS is both a multidimensional poverty measurement tool and a mentoring approach which claims to make the overwhelming reality of poverty digestible and actionable by allowing families to measure their own multidimensional poverty and to develop and implement a clear plan to overcome it.

The PS survey is designed to be an empowering process carried out as a collaboration between loan officer and client[3]: Through a tablet-based, visual survey which uses a series of graphics, clients self-assess their level of poverty in 50 indicators grouped into 6 dimensions (Income & Employment, Health & Environment, Housing & Infrastructure, Education & Culture, Organization & Participation and Interiority & Motivation). Each indicator has three pre-defined levels: Red (extreme deprivation[4]), Yellow (moderate deprivation) and Green (not deprived). The loan officer guides the client through the survey, presenting the three levels for each indicator and asking the client to choose which level best represents the situation of her family. The results are presented immediately after the survey is completed as a poverty dashboard that summarizes in the stoplight colors red, yellow, and green where the client is deprived. Based on these results, the client and her family then select their own priority areas for improvement, and FP helps them identify practical solutions to their problems in an integrated and empowering mentoring program. Together with the loan officer and mentor, the family first identifies the most likely source of the problem (for instance, whether it is due

[3] The survey tool can be and is used in many different contexts and by many types of respondents. However, as this evaluation is concerned with the program targeting FP's women microfinance clients, the term "client" refers to this group.

[4] The PS refers to the three levels of each indicator as extreme *poverty*, moderate *poverty* and non-*poverty*. In the language of the AF methodology, these are referred to as extreme or moderate *deprivation*, respectively, as they present the individual indicators and not the overall welfare level. In the interest of enhanced clarity, this paper adopts the language of the AF methodology, speaking of *deprivations* for individual indicators and of *poverty* for the overall welfare assessment of a client.

to internal or external factors, whether it is a matter of a lack of knowledge and skills, a lack of motivation, a lack of resources, and so on). Then, together they work on appropriate ways of addressing the problems that where identified, drawing on resources from all sectors (private companies, government support, NGOs, family, community…). For this mentoring process, the loan officer has at least one monthly face-to-face interaction and one weekly contact with the client.

The central idea behind this approach is that the poor are not necessarily poor because they lack resources, but (also) because of aspiration failures (Appadurai 2004; Ray 2006): In this view, preferences and behaviors are determined by the social environment. According to Ray, individuals form aspirations windows containing the states they think they can attain, which is heavily influenced by the experiences and lives they can observe from people who seem similar enough, or relatable, to themselves. The difference between an individual's current state and their aspirations window is referred to as the aspirations gap. The theory predicts that individuals will only work to overcome their aspirations gap if the efforts required to close it appear small compared to the improvements they expect for their lives. This implies that individuals will only be driven to improving their situation if they believe that an improvement is achievable (extended aspirations window), and if the aspirations gap is neither too small (which would mean small benefits from improvement) nor too big (which would mean a lot of effort is required). This theory has been further developed by Dalton et al. (2016), who show formally that poor individuals can be stuck in behavioral poverty traps due to aspiration failures, and that under certain circumstances helping individuals to increase their aspirations can be sufficient for them to be able to overcome poverty. Similarly, psychologist Albert Bandura, and based on his work Grenny et al. (2013), argue that change only happens if individuals can answer two questions affirmatively: First, *is it worth it*? And second, *can I do it*? The PS mentoring program is designed to support the processes that facilitate such positive change, by a) expanding the aspirations window (showing "green" as an attainable goal, demonstrating that the goal is in fact within reach through positive deviants in the community or in the relatable environment of the individual, promoting empowerment); and b) decreasing the (perceived or actual) cost and/or increasing the perceived benefit of achieving a goal (showing the value of being "green", helping to identify resources and strategies to achieve goals). Through this process, the voices, actions, and aspirations of the poor themselves become the essential motors for transformation.

This approach does not imply that the PS shifts all responsibility for overcoming poverty onto the poor's shoulders. Rather, the approach's conceptual framework identifies several levels on which changes are necessary, such as on the level of the individual, the community, or the municipality or even state. For problems whose cause is out of the immediate influence of clients, loan officer and client search for ways of accessing the necessary resources, such as applying for certain benefits or petitioning the government.

There is some evidence that the PS program helps families overcome poverty. However, most of this evidence is anecdotal, and the available quantitative studies are based on administrative data from clients who were purposefully selected to participate in the program (Budzyna and Magnoni 2013; Burt 2014). This study will be the first one to evaluate the PS program using rigorous econometric techniques and a dataset of clients that were randomly selected for participation.

The paper contributes to the literature in several ways. First, it provides empirical evidence that interventions focusing on aspiration failures can play an important role in global poverty elimination efforts. Second, it demonstrates the synergies between two poverty measurement methodologies, the Poverty Stoplight and the Alkire-Foster (AF) method, for the purpose of program evaluation: the ordinal data produced by the Poverty Stoplight gives program managers and program participants alike a descriptive and easy way to track a family's poverty status, while also allowing for an intuitive conversion into an AF index based on the very same cut-offs. Furthermore, in addition to the concepts of poverty *incidence* (the number of poor) and poverty *intensity* (number of deprivations of the poor), an analysis of poverty *severity* can naturally emerge from the PS data, based on the PS's "extreme deprivation" (red) and "moderate deprivation" (yellow) levels. This creates extreme poverty and moderate poverty metrics, similar to the Global MPI and Destitution measures (Alkire and Robles 2017). This approach enables further insights into which client groups benefit most from this program.

2 Methodology

2.1 Combining the Poverty Stoplight with the Alkire/Foster (AF) Methodology

Fundación Paraguaya's work is focused on eliminating multidimensional poverty. The poverty measurement tool that FP developed for this purpose, the Poverty Stoplight (PS), is a multidimensional dashboard metric: It gathers information on 50 indicators and displays the results in an intuitive format, using stoplight colors to quickly signal deprivations. This format makes the results easily understandable and useful for users in the field, such as for the poor themselves or for field workers of NGOs. However, the PS has no built-in way of aggregating the information. Because the Alkire-Foster (AF) methodology provides that possibility (Alkire et al. 2015), it is a natural addition to the Stoplight. The AF class of poverty metrics follow an axiomatic tradition of poverty measurement, meaning that AF metrics are designed to fulfill a predefined number of desirable characteristics. This axiomatic tradition is combined with a practical and intuitive counting approach, which makes the AF measures so useful for program evaluation purposes in general and for a combination with the PS dashboard approach in particular. As has been pointed out by Ferreira and Lugo (2013), the choice between dashboard approaches and scalar indices of multidimensional poverty is sometimes presented as an either-or decision, but presents a false dichotomy. Both approaches have distinct advantages, and combining the Poverty Stoplight with the AF method is one way to reap the benefits of both.

For this study, we only use a subset of AF measures, namely the headcount ratio, H, the Average Poverty Intensity, A, and the adjusted headcount ratio, M_0 (or rather their constitutive elements, the poverty identification and the censored deprivation counts, see below). An extensive description of these measures can be found, for instance, in Alkire et al. (2015); a short summary follows. AF metrics are based on a dual cut-off approach: First, for each indicator j (out of $j = 1, \ldots, d$), a deprivation cut-off z_j is defined, and it is determined whether an individual i (out of $i = 1, \ldots, n$) is deprived in indicator j by comparing x_{ij}, the achievement of individual i in indicator j, with the deprivation cut-off z_j. These deprivations are collected in the deprivation matrix g^0 such that $g_{ij}^0 = 1$ whenever $x_{ij} < z_j$ and $g_{ij}^0 = 0$ otherwise. Second, the number of weighted deprivations that an individual suffers is added up to the deprivation score c_i, defined as $c_i = \sum_{i=1}^{d} w_j g_{ij}^0 = \sum_{j=1}^{d} \bar{g}_{ij}^0$, where w_j is the weight assigned to indicator j. Third, the identification function $\rho_k(x_{i.}; z)$ is used to identify individuals as poor if they suffer from at least k (weighted) deprivations: $\rho_k(x_{i.}; z) = 1$ if $c_i \geq k$ and $\rho_k(x_{i.}; z) = 0$ otherwise. The (unadjusted) headcount ratio H is then defined as $H = {}^{q}/_{n}$, where n is the total number of individuals, and q is the number of individuals identified as poor by $\rho_k(x_{i.}; z)$.

In order to obtain the adjusted headcount ratio M_0, one first has to go back to the deprivation matrix g^0 and censor all deprivations of individuals not identified as poor. This is done so as to satisfy the desired property that a poverty measure should change if and only if the achievement of a poor person changes; censoring the deprivations of the non-poor assures that improvements in their situation do not influence the poverty metric. Formally this censored deprivation matrix $g^0(k)$ is obtained by multiplying each element of the deprivation matrix g^0 with the identification function $\rho_k(x_{i.}; z)$: for all i and for all j, $g_{ij}^0(k) = g_{ij}^0 \times \rho_k(x_{i.}; z)$. This matrix now contains only the deprivations of those individuals who have been identified as being poor. A censored deprivation score $c_i(k)$ for each individual i can now be obtained as $c_i(k) = \sum_{i=1}^{d} w_j g_{ij}^0\,(k)$; it is the weighted sum of all censored deprivations that an individual suffers. Thus, $c_i(k) = c_i$ when $c_i \geq k$ and $c_i(k) = 0$ if $c_i < k$. These censored deprivation scores are collected in the censored deprivation vector $c(k)$.

From the censored deprivation matrix, one can now obtain the adjusted headcount ratio as the mean of the censored deprivation score vector: $M_0 = \mu\big(c(k)\big) = \frac{1}{n} \times \sum_{i=1}^{n} c_i(k)$. This is mathematically equivalent to multiplying the (unadjusted) headcount ratio H with the average intensity of poverty that is suffered by those identified as being poor, which is defined as $A = \frac{1}{q}\sum_{i=1}^{q} c_i(k)$. Note that intensity is the number of deprivations suffered, not the poverty severity (whether a deprivation is moderate or extreme).

Our AF measure follows the basic structure of the PS: six equally-weighted dimensions, each with a varying number of equally-weighted indicators, adding up to a total of $d = 50$ indicators. This weighting structure implies that the hierarchy of grouping matters for the final weight of an indicator. For instance, as Table 1 below shows, in order to assure equal weight of all six dimensions, indicators assigned to the dimension “Organization and Participation” end up with a final weight of 1/24 each, while indicators in the dimension "Housing and Infrastructure" have each a final weight of 1/72. In the robustness section, we repeat the estimations with an alternative weighting scheme that gives the "traditional" first four dimensions more weight than the "soft" last two dimensions.

In line with the concept of the PS which distinguishes between "extreme deprivation" (red) and "moderate deprivation" (yellow) as well as "no deprivation" (green), there are two measures that capture varying degrees of poverty severity: an "Extreme Poverty" measure that uses the level "red" as the deprivation cut-off ($z_j^1 = red$), and a "Moderate Poverty" measure that uses the level "yellow" as the deprivation cut-off ($z_j^2 = yellow$). Fundación Paraguaya's declared goal is to *eliminate* poverty, and clients only graduate from the Poverty Stoplight program once they are green in all 50 indicators. This corresponds to the union criterion to identifying the poor (Alkire et al. 2015), and the poverty cut-off k is therefore defined as 1/72, which is the weight of the indicator with the smallest weight: one single deprivation is sufficient in order to be defined as poor. This implies that all censored metrics are equal to their uncensored counterparts (e.g., the censored deprivation count is the same as the uncensored deprivation count). The structure of our AF measure is described in some more detail in Table 1 below.

Table 1 Structure of the Poverty Stoplight-AF Measure

Dimension/Indicator			**Weight within dimension**	**Weight w_j**
Dimension: Income & Employment				**Sum: 1/6**
(1) Income above the poverty line	(2) Stable Income	(3) Credit Facility	1/6 each	1/36 each
(4) Savings	(5) More than one source of income	(6) ID card		
Dimension: Health & Environment				**Sum: 1/6**
(7) Access to drinking water	(8) Nearby health post	(9) Nutrition (malnutrition and/or obesity)	1/9 each	1/54 each

(10) Personal Hygiene and Sexual Health	(11) Eye and Dental Health	(12) Vaccinations		
(13) Garbage Disposal	(14) Unpolluted Environment	(15) Insurance/ Community Help		
Dimension: Housing & Infrastructure				**Sum: 1/6**
(16) Safe home	(17) Sanitary latrine and cloaca	(18) Electricity	1/12 each	1/72 each
(19) Refrigerator and other household appliances	(20 Separate bedrooms	(21) Elevated cook stove and ventilated kitchen		
(22) Comfort of the home	(23) Regular means of transportation	(24) Roads accessible in all weather		
(25) Fixed line or cellular telephone	(26) Security	(27) Sufficient and appropriate clothing		
Dimension: Education & Culture				**Sum: 1/6**
(28) Literacy	(29) Children with schooling up to 12th grade	(30) Knowledge and skills to generate income	1/11 each	1/66 each
(31) Ability to Plan and Budget	(32) Communication and Social Capital	(33) School Supplies and Books		
(34) Access to information (radio and TV)	(35) Entertainment and Leisure	(36) Value cultural traditions and heritage		
(37) Respect for other Cultures	(38) Human rights for vulnerable/ defenseless people			
Dimension: Organization & Participation				**Sum: 1/6**
(39) Forms part of a self-help group	(40) Ability to influence the public sector	(41) Problem and conflict-solving ability	1/4 each	1/24 each
(42) Registered to vote and vote in elections				
Dimension: Interiority & Motivation				**Sum: 1/6**
(43) Awareness of needs: life map	(44) Self-esteem	(45) Moral Conscience	1/8 each	1/48 each
(46) Emotional affective capacity	(47) Aesthetic self-expression, beauty and art	(48) Violence against women		
(49) Entrepreneurial spirit	(50) Autonomy and Ability to make decisions			

2.2 Data

The analysis is based on administrative data from FP. Starting in August of 2015, new program participants for the Stoplight program have been selected randomly each month from all active FP women microfinance clients that are in village banking groups that have not defaulted on their loans. Hence, all PS participants are also microfinance clients (but not all microfinance clients are also PS participants). This selection process implies that participants in the PS program are not necessarily representative of Paraguay´s population (they are all FP microfinance clients), nor of FP´s client base (they are in committees that have not defaulted). While not ideal from an evaluation perspective, the decision to randomize based on the no-default criterion was taken because of program management requirements, as village banking groups that are defaulting are typically disbanding and clients drop out of the program, thus becoming unavailable for mentoring activities and for data collection. Practically speaking, this no-default rule poses only a minor problem: Over the study period, out of all clients randomly selected for participation, only 71 had to be replaced because of the no-default criteria (representing less than 1% of randomly selected clients). Additionally, even from an identification perspective this does not pose a threat to the internal validity of our results as the counterfactual (participants who started the program later) is defined by the same group of non-defaulting clients. However, the effect that the program may have on other types of poor individuals, such as non-microfinance clients or microfinance clients in default, cannot be estimated from this data.

Our database consists of the PS results of over 8,900 of FP's women microfinance clients who did their Stoplight baseline survey between August 2015 and June 15, 2017[5]. New participants enter the program every month, and we refer to the program entry survey of each individual as "baseline" and to subsequent surveys as "follow up". Note that the database does not contain data on "true" non-participants, only on earlier and later entrants. Program participants do a follow-up survey after a year, or when their *asesora* (loan officer) thinks that the family has met the program improvement goals, whichever comes first. This implies that

[5] The actual number is close to 9,500, yet around 600 of these women clients were purposefully selected for participation by their loan officer instead of being randomly selected. These clients are excluded from this analysis.

follow-up data is more likely to be available for clients who reduced their deprivations, possibly making the program participation identifier endogenous in a model that tries to estimate the effect of program participation on the poverty level. In the methodology section we describe how we use months-to-the-end-of-year as an instrumental variable to address this problem[6].

Follow-up data is currently available for around 2,400 women. In about 60% of these cases, more than 100 days elapsed between the rounds; in about 25% of cases, more than 200 days elapsed (see Figure 1). In only around 14% of the cases, eleven months or more elapsed between the survey rounds. Also note that despite program policies, in around 10% of cases 500 days or more passed between survey rounds. Table 2 divides the study period into four semesters and gives an overview of when clients in each baseline survey semester did their last available follow-up survey. The table shows that the time of program participation differs widely among clients. Furthermore, the table shows significant attrition rates (participants who have not done a follow-up survey by June 15, 2017).

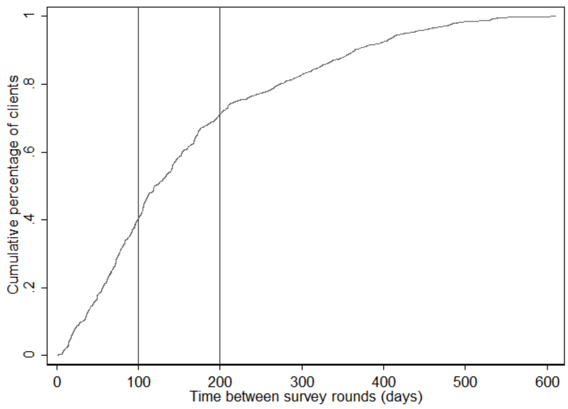

Figure 1 Cumulative distribution of the time difference between baseline and follow-up survey.

[6] We thank Ana Revenga for suggesting this strategy.

Table 2 Number of clients whose last follow-up survey falls in each semester of, by semester of baseline survey

		Last Follow-up			**Total**		
		Jan-Jun 2016	**Jul-Dec 2016**	**Jan-Jun 2017**	**with follow-up [% of total]**	**without follow-up**	**Grand Total**
Baseline	Aug-Dec15	82	260	93	435 [17%]	2,125	2,560
	Jan-Jun16	46	594	130	770 [33%]	1,541	2,311
	Jul-Dec16	0	777	209	986 [61%]	620	1,606
	Jan-Jun17	0	0	207	207 [8%]	2,243	2,450
	Total	128	1,631	639	2,398 [27%]	6,529	8,927

The dataset contains ordinal data on the 50 poverty indicators (coded in the three levels green, yellow, and red indicating non-deprivation, deprivation and extreme deprivation, respectively). This data was used to compute weighted deprivation scores and poverty identifiers as described in section 2.1 above. Additionally, the datasets contains some background information such as zone of residence, date of survey, and loan officer. The date of the survey enters as a time variable in the model, measuring the number of days that have passed between the first survey of the sample (August 7, 2015) and the observation date. The dataset also contains data on program exposure: for each PS indicator, the number of contacts that a program officer had with the client with the goal of overcoming that specific deprivation was recorded. While this data cannot provide any information on the quality of the mentoring activities, we will use it as a proxy to measure program exposure in the robustness section of this paper. Additionally, there is information on family income per capita, which will be used as a control variable. Table 3 summarizes the data, disaggregated by survey round.

Table 3 Description of study sample

	(1) Baseline survey		(2) Last follow-up		(3)
Variable	**N**	**Mean [SE]**	**N**	**Mean [SE]**	**Difference (1) - (2)**
Deprivation count vector, moderate poverty	8924	0.172 [0.001]	2382	0.099 [0.002]	0.073***
Deprivation count vector, extreme poverty	8924	0.062 [0.001]	2382	0.035 [0.001]	0.027***
Poverty identification, moderate poverty	8927	0.985 [0.001]	2398	0.729 [0.009]	0.256***
Poverty identification, extreme poverty	8927	0.808 [0.004]	2398	0.586 [0.010]	0.221***
Poverty intensity, moderate poverty	8788	0.174 [0.001]	1731	0.136 [0.002]	0.038***
Poverty intensity, extreme poverty	7207	0.077 [0.001]	1390	0.060 [0.001]	0.016***
Time (days since first survey in sample)	8927	308.860 [2.163]	2398	474.967 [1.751]	-166.107***
Total number of mentoring contacts	8927	0.000 [0.000]	1853	67.765 [1.469]	-67.765***
Time in program (days since first survey of client)	8927	0.000 [0.000]	2398	162.960 [2.630]	-162.960***
Family income per capita (10,000 Guarani)	8927	80.435 [0.824]	2398	90.574 [1.400]	-10.139***
Months to the end of the year	8927	5.970 [0.038]	2398	4.056 [0.077]	1.914***
Rural area	8773	0.316 [0.005]	2389	0.298 [0.009]	0.017

The values displayed in column (3) are the differences in the means across the groups, not taking into account paired observations. Statistical significance calculated using t-tests and Chi2-tests, respectively.
*Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$*

One important data needs to be addressed before moving on to the actual evaluation. Table 2 shows considerable program attrition: follow-up data is available from within the study period for only about 27% of program participants. We therefore want to test whether participants for whom follow-up data is available differ significantly in observable baseline characteristics from participants from whom no follow-up data is available. Table 4 shows that participants with follow-up data are, if anything, slightly poorer at the baseline than clients without follow-up data: They are 1.7%-points more likely to be moderately poor at the baseline, and their baseline per capita family income is 218,000 Guarani (about 40 USD) lower than that of participants without follow-up data. There are no significant differences in other variables.

Table 4 Baseline difference between participants with and without follow-up data

	All		Without follow-up		With follow-up		Diff.
	mean	**se**	**mean**	**se**	**mean**	**se**	
rural	0.316	0.006	0.324	0.0073	0.300	0.0099	0.024
Poverty Identification							
Moderate	0.985	0.001	0.980	0.0017	0.9971	0.001	-0.017***
Extreme	0.808	0.004	0.808	0.0049	0.8078	0.008	0.000
Deprivation Count							
Moderate	0.172	0.001	0.172	0.0015	0.171	0.0024	0.001
Extreme	0.062	0.001	0.062	0.0008	0.063	0.0012	-0.001
Income p.c. (10,000G)	80.435	0.824	86.292	0.992	64.488	1.405	21.804***
Date	308.860	2.163	307.704	2.721	312.007	3.1531	-4.303

* p<0.05, ** p<0.01, *** p<0.001

2.3 Evaluation methodology

Estimation strategy

We use a pipeline design in order to evaluate whether the Poverty Stoplight program has been successful in decreasing multidimensional poverty. After using simple hypothesis tests to compare the poverty levels the poverty level of program participants at their follow-up surveys with the poverty level of comparable FP clients who are newly entering the program, we use multiple regression to control for potential other factors, using the participation identifier as the main explanatory variable. We do this first by aggregating data over the entire study period, and then disaggregated by semester. As beneficiaries are randomly selected for program participation on an ongoing basis (see description in the data section), comparing early and later program entrants while controlling for a general improvement trend gives an approximation of where program participants may be without the program. Using later entrant's respective baseline surveys as a counterfactual allows to estimate a general improvement trend and thus effectively creates a control group.

The model

Our model is defined as follows, and is estimated with cluster-robust standard errors:

$$poverty = \beta_0 + \beta_1 mentoring + \beta_2 time + \beta_3 income + \beta_4 rural + \beta_{5s} oficina_s + e$$

where $poverty$ is the measure of poverty (see explanation below); $mentoring$ identifies the intervention/program exposure (see explanation below); $time$ is the number of days since the first observation in the entire sample (August 7th, 2015); $income$ is the family per capita income level, in 10,000 current Paraguayan Guaraní (currently approximately USD1.8); and $rural$ is the residency area (which is a dummy taking the value 1 if the client lives in a rural area). Furthermore, we include fixed effects for the loan offices, represented by $\beta_{5s} oficina$; the coefficients on these loan office dummies are not of interest for this study and thus not reported, but they are included to control for heterogeneities between loan offices[7].

Outcome variable

The outcome variable, $poverty$, is the deprivation score $c_i(\frac{1}{72})$. This deprivation score is the weighted sum of deprivations that the client suffers. We use two versions of the outcome variable: one for extreme poverty, which is based on the deprivation cut-off $z_j^1 = red$ in each indicator; and for moderate poverty, which is based on the deprivation cut-off $z_j^2 = yellow$ in each indicator.

[7] An ANOVA analysis shows significant differences in the outcome variables between loan offices.

Main explanatory variable

The effect of the program, $mentoring$, is a dichotomous identifier indicating whether or not an observation comes from a follow-up survey, that is, whether client i has already participated in the Poverty Stoplight program at time t.

Note, however, that our program participation identifier cannot account for potential differences in program implementation, such as the intensity of mentoring activities. The robustness section of this paper contains the results of identifying program exposure by the total number of mentoring contacts and by the number of days in the mentoring program.

3 Results

3.1 Descriptive Analysis

The first step in the analysis is a simple comparison of the AF metrics (headcount ratio H, poverty intensity A, and adjusted headcount ratio M_0) between baseline and follow-up surveys. The former provides a counterfactual for the identification of the program effect, as program participants are randomly selected from the client population; in any given semester, the observed baseline poverty levels indicate where participants might be without the program.

Table 5 shows the observed poverty levels of the study population, together with confidence intervals derived through bootstrapping as suggested by Alkire et al. (2015). The same data is graphically represented in Figure 2. The data shows good targeting: Almost all clients in the sample are at least moderately poor at their baseline survey; the large majority are also extremely deprived in at least one indicator (i.e., has at least one indicator in “red”). However both for moderate and for extreme poverty the share of clients who are poor at the baseline decreases over time: from 99% to 97.1% for the former, and from 85.6% to 71% for the latter. As can be concluded based on the confidence intervals, this decrease in the headcount ratio at the baseline is not statistically significant in the case of moderate poverty, but is so for extreme poverty. A decrease in poverty levels over time at the baseline can also be established for poverty intensity (A) and the adjusted headcount ratio (M_0); these decreases are statistically significant both for moderate and for extreme poverty.

In each semester and for both moderate and for extreme poverty, the percentage of poor survey-takers is larger among those doing their baseline compared to those doing a follow-up survey. These differences are statistically significant in all cases, providing a first piece of evidence for the effectiveness of the Poverty Stoplight program. For poverty intensity (the average number of deprivations suffered by those who are poor), the picture is less clear. While those who are identified as poor in the second study semester (between January and June 2016) suffer significantly more deprivations if they are just entering the program than if they are already exiting, we find no such statistical evidence for the later semesters. For the combined effect of poverty incidence and intensity, as captured by the M_0 measure, the effect of a decreased poverty headcount prevails, for both moderate and for extreme poverty, for the second and third study semester, resulting in a significantly lower M_0 among program participants compared to new entrants; in the last semester, there is no statistical difference in the M_0 between baseline and follow-up surveys.

Put together, this descriptive analysis suggests that the program contributes to a decrease in poverty, above all with regards to decreasing the likelihood of being poor; for those who remain in poverty (in our case, those with at least one deprivation), a change in poverty intensity can statistically only be shown for one of the semesters.

Table 5 Poverty incidence and intensity by survey semester and survey round (95% confidence intervals, computed through bootstrapping, in parenthesis)

		Moderate Poverty		Extreme Poverty	
		Baseline Survey	**Follow-up Survey**	**Baseline Survey**	**Follow-up Survey**
Aug-Dec 2015	**H**	99.0%		85.6%	
		[98.5% , 99.4%]		[82.5% - 88.8%]	
	A	22.0%		9.0%	
		[20.5% , 23.5%]		[8.3% - 9.7%]	
	M0	0.219		0.078	
		[0.204 , 0.233]		[0.069 - 0.087]	
Jan-June 2016	**H**	98.9%	87.0%	88.2%	71.5%
		[98.2% , 99.6%]	[78.9% - 95.1%]	[85.5% - 90.8%]	[61.4% - 81.7%]
	A	19.7%	14.8%	8.3%	6.5%
		[18.4% , 21.0%]	[12.6% - 17.0%]	[7.4% - 9.3%]	[5.5% - 7.4%]
	M0	0.193	0.127	0.073	0.046
		[0.182 , 0.205]	[0.101 - 0.152]	[0.063 - 0.083]	[0.036 - 0.055]
July-Dec 2016	**H**	99.2%	71.3%	77.1%	58.4%
		[98.7% , 99.7%]	[67.% - 75.7%]	[73.2% - 81.1%]	[53.4% - 63.3%]
	A	16.3%	14.2%	7.3%	6.1%
		[15.0% , 17.6%]	[13.2% - 15.2%]	[6.7% - 7.9%]	[5.4% - 6.8%]
	M0	0.162	0.101	0.056	0.036
		[0.149 , 0.174]	[0.091 - 0.112]	[0.05 - 0.062]	[0.031 - 0.04]
Jan-June 2017	**H**	97.1%	73.3%	71.0%	55.8%
		[95.2% , 98.9%]	[67.% - 79.6%]	[65.6% - 76.5%]	[50.0% - 61.5%]
	A	11.1%	11.8%	5.5%	5.7%
		[10.1% , 12.1%]	[10.9% - 12.7%]	[5.0% - 6.0%]	[5.0% - 6.4%]
	M0	0.108	0.086	0.039	0.032
		[0.098 , 0.118]	[0.074 - 0.099]	[0.034 - 0.045]	[0.026 - 0.038]

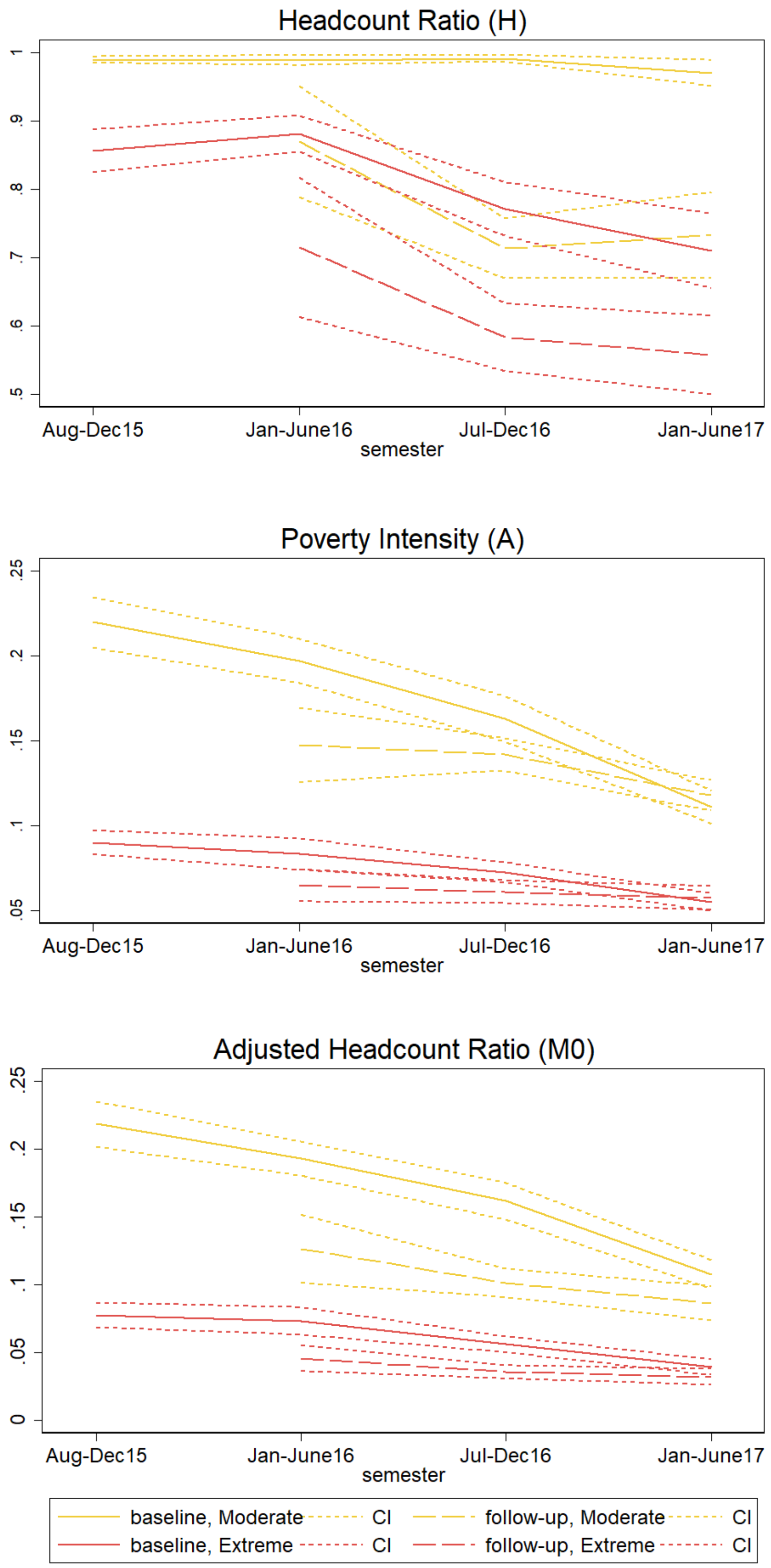

Figure 2 Alkire Foster metrics for baseline and follow-up surveys at each semester (with bootstrap confidence intervals)

3.2 Regression Analysis

Follow- up surveys are only available for a relatively small share of those doing the baseline survey. Additionally, there are some concerns that the overall client population might have changed over the course of the study. Hence, in order to arrive at a more robust estimate of the program effect, we want to control for some client characteristics. This will be the done in the next section through the use of OLS estimations.

Table 6 and Table 7 present the results of OLS regression with cluster-robust standard errors and office-dummies to account for potential differences between loan offices. The models reveal statistically significant effects of program participation in the expected direction: Program participants see their moderate deprivation count reduced by between 2 and 5 percentage points, and their extreme deprivation count reduced by between 0.6 and 2 percentage points, depending on the model.

With regards to control variables, the models show a negative time trend of between around 4 and 17 percentage points for moderate poverty and between 2 and 7 percentage points for extreme poverty, on an annualized basis (the reported coefficients are changes per day). These effects are statistically significant in all models except for semester 2. Furthermore, as expected, per capita income is associated with a reduced deprivation count, which is statistically significant in all models; for instance, according to the aggregated model, the moderate deprivation count is reduced by 0.06 percentage points per 10,000 Guarani (18 USD) more in family per capita income. Finally, the area of residence has a significant effect only for extreme poverty, and for one of the models of moderate poverty.

Table 6 Results of the OLS estimations, dependent variable: deprivation count vector for moderate poverty

	Aggregate	Semester 2	Semester3	Semester 4
Mentoring	-.0345***	-.0506***	-.0255***	-.02**
	(-8.68)	(-5.77)	(-4.24)	(-3.35)
Date	-.00018***	0.000063	-.00047***	-.00012*
	(-10.06)	(0.86)	(-6.04)	(-2.14)
Income p.c.	-.00056***	-.00063***	-.00057***	-.00035***
	(-12.69)	(-13.17)	(-9.91)	(-7.96)
rural	.0128	.00941	.0248	.025***
	(1.99)	(1.20)	(1.97)	(4.07)
Intercept	.261***	.235***	.366***	.193***
	(33.09)	(12.87)	(12.23)	(5.77)
N(total)	1,1143	2,351	3,224	3,083
N(follow-up)	2,373	122	1,618	633
R2	.336	.296	.34	.281

Office-level fixed effects included but not reported. Standard errors clustered at office level.
t-statistics in parenthesis. * p<0.05, ** p<0.01, *** p<0.001

Table 7 Results of the OLS estimations, dependent variable: deprivation count vector for extreme poverty

	Aggregate	Semester 2	Semester3	Semester 4
Mentoring	-.0126***	-.0207***	-.00721*	-.00641
	(-5.82)	(-3.92)	(-2.30)	(-1.69)
Date	-0.000065***	-0.000041	-.00019***	-0.000065*
	(-6.16)	(-1.18)	(-5.51)	(-2.13)
Income p.c.	-.00024***	-.00029***	-.00022***	-.00014***
	(-10.54)	(-8.94)	(-9.02)	(-5.98)
rural	.0129***	.0111*	.0139*	.0173***
	(4.12)	(2.55)	(2.52)	(4.52)
Intercept	.0868***	.0987***	.121***	.0732***
	(19.80)	(10.59)	(8.87)	(3.80)
N(total)	1,1143	2,351	3,224	3,083
N(follow-up)	2,373	122	1,618	633
R2	.265	.286	.256	.25

Office-level fixed effects included but not reported. Standard errors clustered at office level.
t-statistics in parenthesis. * p<0.05, ** p<0.01, *** p<0.001

3.3 2SLS Models

We are concerned about the possible endogeneity of the program participation identifier, because follow-up surveys may be more likely to be done if the loan officer thinks that the program goals were reached (i.e., if she thinks that the deprivation count has decreased). We address this problem using instrumental variable estimation, taking advantage of the fact that

Fundación Paraguaya incentivizes its loan officers to help their clients overcome poverty by entering those who meet a specified yearly target into a lottery to win prizes. As a result, the number of follow-up surveys increases as the end of the year approaches and loan officers focus harder on meeting their targets[8]. Therefore, the number of months left until the end of the year can be used to instrument program participation, as a given survey is more likely to be a follow-up survey as the end of the year approaches.

A valid instrument has to fulfill to requirements: First, it has to be correlated with the endogenous variable; and second, it must not be correlated with the error term of the structural model, e. While the later requirement cannot be formally tested, the former one can be, by showing for the following reduced-form model that $\pi_1 \neq 0$:

$$mentoring = \pi_0 + \pi_1 months + \pi_2 time + \pi_3 income + \pi_4 rural + \pi_{5s} oficina + v$$

Indeed, this is the case for all models: the null hypothesis that $\pi_1 = 0$ can be rejected at the highest significance level (p<0.001); the results of these first stage estimations are reported in Table 8. We are also confident that our instrument fulfills the first requirement, i.e., that $Cov(months, e) = 0$. While it is possible that some of the indicators are sensitive to changes in the seasons, it is not clear why the month of the year should be systematically correlated with the unobservable factors that affect the deprivation count in the structural model (which was presented on page 15). We therefore estimate the model using months-to-the-end-of-the-year as an instrument in a 2SLS estimation.

[8] In order to minimize incentives to misreport, Fundación Paraguaya uses a system of random checks to verify whether any reported improvements have indeed taken place. If misreporting on behalf of the loan officer is detected, her entire portfolio of follow-up surveys is discarded. These checks have not revealed any substantial misreporting in the study period, which is why we are not worried about misreports due to conflicts of interest.

Table 8 Results of the First Stage estimations

	Aggregate	Semester 2	Semester3	Semester 4
Date	.00096***	.00775***	.0224***	-.00771***
	(67.59)	(7.62)	(24.66)	(-7.65)
Income p.c.	.0002***	.0003**	.00068***	-.00024**
	(4.14)	(3.09)	(6.45)	(-2.73)
rural	-.00441	.00206	.0387*	-.0517**
	(-0.59)	(0.18)	(1.98)	(-2.80)
Months	-.041***	.218***	.596***	-.269***
	(-38.57)	(7.12)	(20.14)	(-9.27)
Intercept	.163***	-3.65***	-10.3***	7.17***
	(9.96)	(-7.37)	(-22.37)	(8.43)
N	11,143	2,351	3,224	3,083

Office-level fixed effects included but not reported.
t-statistics in parenthesis. * p<0.05, ** p<0.01, *** p<0.001

Both for moderate poverty (Table 9) and for extreme poverty (Table 10), the results of the 2SLS estimations show a highly significant effect of participating in the program in the aggregated model: the programs reduces the moderate deprivation count by 3.7%-points, and the extreme deprivation count by 2.8%-points, respectively. However, no statistically significant effect of program participation can be found when analyzing the effect for each study semester separately.

As far as control variables are concerned, as expected we find a negative time trend (i.e., M_0 decreases over time); income is negatively correlated with the number of deprivations that a client suffers (e.g., in the aggregate model for moderate poverty, a 100,000 Guaraní increase in per capita family income reduces the deprivation count by 0.6 points); and living in a rural area is associated with higher levels of poverty.

The last part of the table reports the results of the Durbin-Wu-Hausman test of endogeneity of the program identifier. As it turns out, the null hypothesis of exogeneity can be rejected only for one case, for the aggregated model of extreme poverty. At a reasonable level of confidence (5%), there is no evidence for endogeneity in any of the other models. We therefore conclude that the OLS models are appropriate.

Table 9 Results of the 2SLS estimation, dependent variable: deprivation count vector for moderate poverty

	Aggregate	Semester 2	Semester3	Semester 4
Mentoring	-.0374***	-.0663	-.018	.0277
	(-5.79)	(-1.73)	(-1.76)	(1.36)
Date	-.00017***	0.000075	-.0005***	-.00014***
	(-26.41)	(1.75)	(-10.25)	(-3.69)
Income p.c.	-.00056***	-.00063***	-.00057***	-.00034***
	(-35.13)	(-13.46)	(-20.41)	(-17.39)
rural	.0127***	.00952	.0245***	.0276***
	(4.18)	(1.79)	(4.18)	(6.09)
Intercept	.261***	.232***	.373***	.197***
	(48.83)	(16.29)	(21.99)	(8.86)
N(total)	11,143	2,351	3,224	3,083
N(follow-up)	2,373	122	1,618	633
Durbin-Wu-Hausman Test of Endogeneity				
F-Statistc	0.088	0.109	0.106	3.443
p-value	0.770	0.744	0.748	0.076

Office-level fixed effects included but not reported.
t-statistics in parenthesis. * p<0.05, ** p<0.01, *** p<0.001

Table 10 Results of the 2SLS estimation, dependent variable: deprivation count vector for extreme poverty

	Aggregate	Semester 2	Semester3	Semester 4
Mentoring	-.0282***	-.00936	.00188	.0111
	(-8.21)	(-0.46)	(0.36)	(1.02)
Date	-0.000054***	-0.000050*	-.00022***	-0.000077***
	(-15.14)	(-2.06)	(-8.81)	(-4.12)
Income p.c.	-.00023***	-.00029***	-.00023***	-.00014***
	(-27.99)	(-11.19)	(-16.25)	(-12.34)
rural	.0124***	.011***	.0135***	.0183***
	(7.80)	(3.46)	(4.42)	(7.10)
Intercept	.0869***	.101***	.129***	.0748***
	(30.31)	(12.27)	(14.85)	(7.05)
N(total)	11,143	2,351	3,224	3,083
N(follow-up)	2,373	122	1,618	633
Durbin-Wu-Hausman Test of Endogeneity				
F-Statistc	8.894	0.262	0.869	1.238
p-value	0.007	0.614	0.361	0.277

Office-level fixed effects included but not reported.
t-statistics in parenthesis. * p<0.05, ** p<0.01, *** p<0.001

4 Discussion

4.1 Main findings

We found evidence that participating in the Poverty Stoplight program is associated with better odds of overcoming poverty: the results from the descriptive analysis and the OLS estimation suggest a strong general improvement trend in our study population, and that participation in the PS program adds to this positive trend. In concrete terms, our results suggest that over half a year—which is the typical time span between a baseline and a follow-up survey—the average deprivation count of FP’s general micro finance client population decreases by between 2 and 8 percentage point for moderate poverty, and by between 1 and 3 percentage points for extreme poverty. A our index consists of the 50 PS indicators, this translates to between around 1 and 4 less (weighted) PS indicators in “yellow” or “red”, and between 0.5 and 1.5 less in “red”, respectively. The PS program adds to this general time trend a reduction of between 2 and 5 percentage points for moderate poverty, and of between 0.6 and 2 percentage points for extreme poverty, respectively (i.e., between 1 and 2.5 indicators less in “yellow” or “red”, and between 0.3 and 1 indicators less in “red”).

For the 2SLS models, this positive program effect could only be shown on an aggregate level, not for individual semesters – yet we showed that program participation does not seem to be endogenous, which is why the more efficient OLS is preferred.

Note that the entire study population are micro finance clients, which might explain at least part of the overall decrease; the results suggest that participating in the PS program in addition to the micro finance program increase participants’ welfare even further. This overall result is encouraging. It supports the theory that changing people’s aspirations and providing mentoring can help people overcome poverty.

4.2 Robustness of Results

We carry out three types of robustness analysis. First, we want to check how robust our results are to an alternative measurement of program participation, given that our participation dummy hides the fact that the time in the program and the number of mentoring contacts received varies considerably among participants. We will thus measure program participation by the number of mentoring contacts received and by the time-of-exposure (number of days between baseline survey and follow-up survey).

Second, as described in the background section, the Poverty Stoplight gives a prominent role to non-traditional dimensions of poverty, Organization & Participation and Interiority & Motivation. This makes it unique in the field of multidimensional poverty measurement, but also begs the question whether any observed improvements are concentrated in these "soft" dimensions. We therefore decompose the M_0 at the baseline and follow-up surveys to see where most changes happen, and check how our results vary if the non-traditional dimensions are given less weight. Instead of six equally weighted dimensions (with equally weighted nested indicators), this alternative AF measure gives the four "traditional" dimensions Income & Employment, Health & Environment, Housing & Infrastructure, and Education & Culture a combined weight of 90%, while the remaining two dimensions Organization & Participation and Interiority & Motivation receive only 10% of the total weight. Each dimension within each of these two groups, and each indicator within each dimension, receives equal weight; the poverty cut-off is reduced to 1/160 to maintain the union approach to poverty identification.

Third, we are concerned that the comparisons between follow-up survey and baseline survey in the pipeline design might not be valid if the client population changes over time. As new participants are selected randomly among all eligible clients of a given month, a change in the client population will result in biased comparisons. As was shown in the data section, the poverty metrics at the baseline survey differ significantly between program semesters; it is conceivable that other observed and unobserved characteristics differ as well. To address this issue, we limit our sample to those clients who were already micro finance clients and eligible for program participation at the start of the program, in 2015. While this reduces the sample size drastically, it also assures that the characteristics of the study population remain constant over the study period.

Table 11 shows the results of the first exercise: program participation is first measured through the total number of mentoring contacts that a client received, and then through the time of exposure (the number of days spent in the program). The results by and large replicate the findings of the main analysis: receiving more mentoring contacts, and spending more time in the program, are both associated with a reduced deprivation count vector. The same analysis was carried out using 2SLS, in which case the results replicated the results of the 2SLS estimation in the main section (results not reported).

Table 11 Results of the OLS estimations, alternative program identification

	Aggregate	**Semester 2**	**Semester3**	**Semester 4**
Dependent variable: Deprivation count for moderate poverty				
Mentoring contacts	-.0394***	-.0588***	-.0332***	-.0222**
	(-9.82)	(-6.07)	(-5.52)	(-3.76)
Time in program	-.0394***	-.0588***	-.0332***	-.0222**
	(-9.82)	(-6.07)	(-5.52)	(-3.76)
Dependent variable: Deprivation count for extreme poverty				
Mentoring contacts	-.0154***	-.0221**	-.0104**	-.00726
	(-6.43)	(-3.27)	(-3.39)	(-1.99)
Time in program	-.0154***	-.0221**	-.0104**	-.00726
	(-6.43)	(-3.27)	(-3.39)	(-1.99)
N	11,144	2,351	3,224	3,083

Date, income, and rural, intercepts, and office-level fixed effects included but not reported.
t-statistics in parenthesis. * p<0.05, ** p<0.01, *** p<0.001

Figure 3 shows the contribution of the individual PS dimensions to the overall M0 metric, at the baseline and the follow-up surveys. The graphs show that for both moderate and extreme poverty, the biggest decreases in deprivations can be found in the dimensions Income & Employment, and to a lesser extent Health & Environment and Housing & Infrastructure. In fact, the “softer” dimensions “Organization & Participation” and “Self-Awareness & Motivation” contribute much less to overall poverty already at the baseline, and are clearly not driving the decrease in poverty that the analysis reveals. In that light, it is not surprising that with the use of the alternative weighting scheme, the program effect is, if anything, slightly stronger than with the original weighting scheme (results presented in Table 12). Overall, we can be confident that our results are not only driven by the “soft” indicators included in the PS, which are possibly more prone to misreporting. Again, the analysis was also carried out using 2SLS, with results that are not reported here but that mirrored the results of the 2SLS estimation in the main part of the paper.

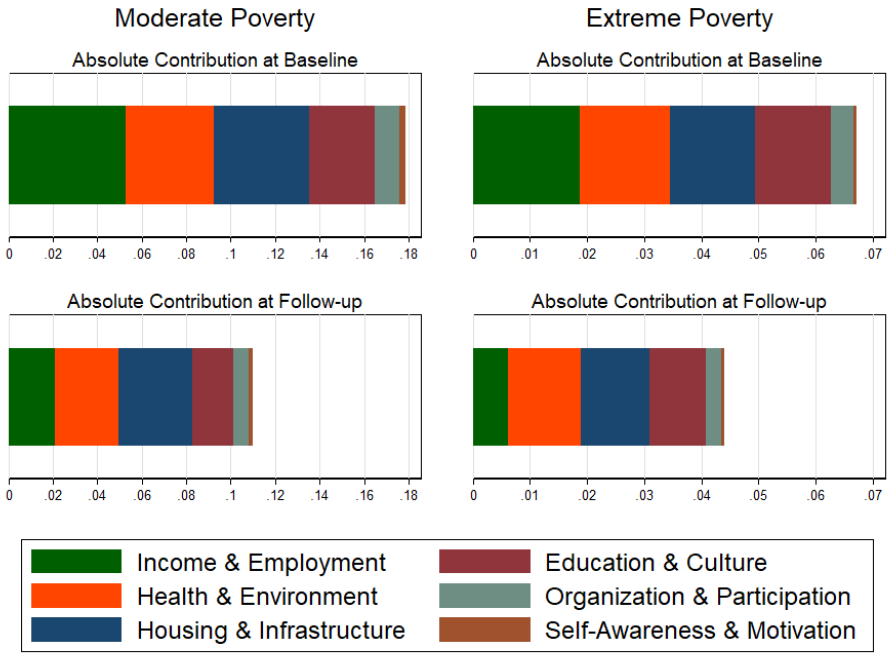

Figure 3 Absolute contribution of PS dimensions to M0, at baseline and follow-up surveys

Table 12 Results of the OLS estimation with alternative weights

	Aggregate	Semester 2	Semester3	Semester 4
Dependent variable: Deprivation count for moderate poverty				
Mentoring	-.039388***	-.058798***	-.033184***	-.022183**
	(-9.82)	(-6.07)	(-5.52)	(-3.76)
Dependent variable: Deprivation count for extreme poverty				
Mentoring	-.015362***	-.022078**	-.010417**	-.007263
	(-6.43)	(-3.27)	(-3.39)	(-1.99)
N	11,144	2,351	3,224	3,083

Date, income, and rural, intercepts, and office-level fixed effects included but not reported
t-statistics in parenthesis. * p<0.05, ** p<0.01, *** p<0.001

Table 13 Results of the OLS estimation with restricted sample

	Aggregate	Semester 2	Semester3	Semester 4
Dependent variable: Deprivation count for moderate poverty				
Mentoring	-.0272***	-.0326	-.0156*	-.0123
	(-5.18)	(-1.94)	(-2.24)	(-1.36)
Date	-.0002***	7.5e-05	-.00052***	-.0001
	(-10.14)	(0.97)	(-7.78)	(-1.24)
Income p.c.	-.00055***	-.00053***	-.00048***	-.0003***
	(-12.44)	(-8.38)	(-7.43)	(-4.86)
rural	.0185	.018	.0384*	.0232*
	(1.93)	(1.94)	(2.33)	(2.19)
Intercept	.247***	.181***	.367***	.173**
	(30.39)	(8.82)	(12.75)	(3.50)
N(total)	4,234	945	1,040	740
N1(follow-up)	927	69	633	225
R2	.364	.365	.376	.315
Dependent variable: Deprivation count for extreme poverty				
Mentoring	-.00892**	-.0111	-.00148	-.00148
	(-2.89)	(-1.39)	(-0.46)	(-0.24)
Date	-.000072***	-.000022	-.0002***	-.000055
	(-6.06)	(-0.41)	(-4.96)	(-0.96)
Income p.c.	-.00023***	-.0002***	-.00018***	-.00012***
	(-11.46)	(-7.96)	(-7.02)	(-4.06)
rural	.0176***	.0187*	.0214**	.0182***
	(4.74)	(2.62)	(3.29)	(4.44)
Intercept	.0832***	.0702***	.116***	.0601
	(18.71)	(5.48)	(7.12)	(1.82)
N(total)	4,234	945	1,040	740
N(follow-up)	927	69	633	225
R2	.288	.324	.286	.260

Office-level fixed effects included but not reported.
t-statistics in parenthesis. * p<0.05, ** p<0.01, *** p<0.001

Finally, Table 13 shows the results of the OLS model when the analysis is restricted to those who were eligible to participate in the program from the very program start. This reduces the sample size considerably. In these models, the effect of program participation is significant only in the aggregate, and, in the case of moderate poverty, in semester three. The coefficients maintain their expected signs, but are slightly smaller than in the main models. Given the small sample sizes in the individual semesters, particularly with regard to data from follow-up surveys, we estimated the power that the respective samples had to detect a program effect of the (non-statistically significantly) estimated size. As it turns out, the power levels are very low, at between 0.05 and 0.6. Hence, it remains possible that the lack of significance is due to an insufficiently large sample size.

4.3 Limitations

The main shortcoming of this evaluation study is the lack of two rounds of observations for clients who did not participate in the program, which makes it harder to isolate the program effect. Even though all participants, including later entrants who provide the counterfactual for earlier ones, are randomly selected from the same pool of clients, there might have been some shifts in the client population, or systematic attrition, both of which pose threats to the validity of our results. We partly mitigated these problems by controlling for some client characteristics, using an instrumental variable to address the potential endogeneity of program participation, measuring program exposure in a nuanced way, and running the analysis on a restricted but constant sample of clients.

An additional shortcoming of the present study is the lack of further control variables that may reasonably be assumed to influence poverty status, and in some cases program participation. For instance, at the client level the database does not allow us to control for hard-to-measure concepts such as motivation or effort. Additionally, one would like to control for factors such as the characteristics of the loan officers or the social and economic environment in the region. Note, however, that the office level fixed effects control for some of these higher-level effects.

Finally, the validity of our results might also suffer because the average time difference between baseline and follow-up surveys was only around 5.5 months, and only in about 10% of the cases for which baseline and follow-up data was available, a year or more passed between the two rounds. This time difference may well be too short for sustainable changes in multidimensional poverty. Additionally, it is important to keep in mind that the PS program is part of FP's microfinance program which provides additional support for clients to overcome poverty. Out data does not allow us to estimate which effect the PS program might have in the absence of a micro finance program.

5 Conclusions

This paper set out to evaluate the Poverty Stoplight program, estimating its effect on multidimensional poverty with the help of the Alkire-Foster poverty measurement method. Our results indicate that participation in the PS program is indeed associated with a decrease in multidimensional poverty, which suggests that the integrated mentoring approach can be a valuable tool to eliminate poverty. The PS's program theory assumes that poor people can overcome their own poverty if they can affirmatively answer the two questions "is it worth it?" and "can I do it?", which can be reframed in the language of the emerging literature on aspiration failures as problems related to the size of the aspirations window and the perceived costs and benefits of closing it. The results of this study support the notion that aspiration failures can be addressed with a targeted mentoring program, enabling people to overcome poverty. The results indicate that the PS approach is particularly useful for clients suffering from moderate deprivations, yet some benefits can also be shown for clients suffering extreme deprivations.

Much more research is needed, however, to be able to draw firm conclusions. First, the conclusions of this study should be replicated using a true experimental design in which the poverty levels of both the treatment and the control group are measured at program start and at the follow-up, which allows to identify the program effect more accurately. Second, in order to learn more about the mode of action of the Poverty Stoplight, and about whether it truly can be a tool to overcome aspiration failures, more targeted research is necessary that explicitly measures how and if the PS influences aspirations, and if this truly is the mechanism through with the PS decreases poverty. Third, this study focused on a specific population, i.e., on active women microfinance clients in Paraguay. It is unclear how the PS's effectiveness might differ when applied to another population, as microfinance clients might be more receptive to motivational interventions than other poor people. Finally, a longer-term study is necessary to study the sustainability of the PS program's effect on poverty. Such a longer-term study would also allow a closer look at the differences in the role that aspiration failures, as opposed to other challenges, play for people who live in poverty.

6 References

Alkire, Sabina, James Foster, Suman Seth, Maria Emma Santos, and Jose Manuel Roche. 2015. *Multidimensional Poverty Measurement and Analysis*. New York, NY: Oxford University Press.

Alkire, Sabina, and Gisela Robles. 2017. "Multidimensional Poverty Index - Summer 2017: Brief Methodological Note and Results." MPI Methodological Notes 44. Oxford, UK: Oxford Poverty & Human Development Initiative (OPHI). http://www.ophi.org.uk/wp-content/uploads/OPHIMethNote_44_Summer_2017.pdf.

Appadurai, Arjun. 2004. “The Capacity to Aspire: Culture and the Terms of Recognition.” In *Culture and Public Action*, edited by Vijayendra Rao and Michael Walton, 59–84. Stanford, CA: Stanford Univ. Press.

Budzyna, Laura, and Barbara Magnoni. 2013. “Measuring the Social Impact of Fundación Paraguaya.” EA Consultants.

Burt, Martin. 2014. “The Poverty Stoplight: Does Personalized Coaching in Microfinance Help Clients Overcome Poverty?” Unpubslished internal document. Asunción, Paraguay: Fundación Paraguaya.

Dalton, Patricio S., Sayantan Ghosal, and Anandi Mani. 2016. “Poverty and Aspirations Failure.” *The Economic Journal* 126 (590): 165–88. https://doi.org/10.1111/ecoj.12210.

Ferreira, Francisco H. G., and Maria Ana Lugo. 2013. “Multidimensional Poverty Analysis: Looking for a Middle Ground.” *World Bank Research Observer* 28 (2): 220–35.

Grenny, Joseph, Kerry Patterson, David Maxfield, Ron McMillan, and Al Switzler. 2013. *Influencer: The New Science of Leading Change, Second Edition*. 2 edition. New York: McGraw-Hill Education.

Ray, Debraj. 2006. “Aspirations, Poverty, and Economic Change.” In *Understanding Poverty*, edited by Abhijit Banerjee, Roland Bénabou, and Dilip Mookherjee, 409–21. New York, NY: Oxford University Press.