Predictive Targeting: Increasing the efficiency of social service delivery in Brazil

Ishani Desai
Rohit Kumar
Shammi Quddus
Anne Shrestha

Harvard Kennedy School
Advanced Quantitative Methods I (API-209)
Final Exercise
Professor Dan Levy
Fall 2014

I. EXECUTIVE SUMMARY

The Ministry of Social Development and Fight Against Hunger (MDS) manages several social security programs in order to serve the poor and extreme poor of Brazil. The Bolsa Familia is the largest cash transfer in the world that serves 12 million households - 26% of the population of Brazil. Without official income data, a program of this scale depends on self-reported incomes to assess household eligibility. This complicates the process of targeting because households have an incentive to underreport income with the hope of qualifying for benefits. In this memo, we propose a **predictive targeting model** that uses a set of easily observable poverty indicators and removes the dependency on self-reported income for targeting.

The poor households in Brazil have many differentiating characteristics compared to the non-poor. These differences can be used to create a precise predictive targeting model. For example, households are more likely to be poor if headed by a female or if the household head cannot read and write. Poorer households tend to be younger and have a lower proportion of adult household members being currently employed. Unsurprisingly, the quality of their dwelling is of lower quality, they have lower access to public utilities, and own fewer material assets.

The proposed predictive targeting model uses **45 of these differentiating characteristics** to predict the probability of a household being poor. Information on these 45 indicators will be collected through a one-page survey and entered into an algorithm that will assign each household a 'poverty score' between 0 and 1. The threshold score proposed is 0.20 which means a household will be eligible for benefits if the probability of its being poor is equal or greater than 20%. This cutoff results in an undercoverage (exclusion of deserving poor households) and leakage (inclusion of non-poor households) of **39%** each. This is a significant improvement on an undercoverage rate of 59% and leakage rate of 49% of the Bolsa Familia Program reported in 2004. With a **concurrent reduction of both undercoverage and leakage**, the MDS can extend its benefits to a greater number of deserving poor and minimize wastage by eliminating a greater number of non-poor households.

We propose that the predictive targeting model be implemented using the **existing Cadastro Unico infrastructure,** which is a three-tiered system. Currently, survey data is collected at the municipal level and sent to the Caixo Economica Federal. The MDS uses the income data to carry out an unverified means testing (UMT) to determine eligibility. This infrastructure is ideal for the implementation of predictive targeting as it has the advantage of being decentralized (compared to a national survey like the PNAD) and well experienced to transfer survey data in a digitized form to the MDS.

The proposed predictive targeting model requires a short questionnaire to collect information on the proxy indicators of poverty, which would replace the much longer survey instrument of the Cadastro Unico. Hence, by using predictive targeting, not only can the MDS **improve the targeting** of its safety nets, it can do so more **efficiently** and **cost-effectively**. We also propose supplemental measures to bolster targeting; namely, a **rolling enrollment** feature to extend BFP's role not merely as a social safety net but also as a social insurance and **different cut-off rates** in areas with a high incidence of poverty.

II. POLICY MEMO

A. INTRODUCTION

Roughly 18 million Brazilians live in poverty despite a steady decline in the poverty rate from 16% in 2007 to 9% in 2012.[1] The Ministry of Social Development and Fight Against Hunger (MDS) has made numerous efforts in the past decade to streamline its social safety net to reach the poorest in the country. Efficient poverty targeting is important for the MDS for many purposes, but it is particularly relevant for the Bolsa Familia Program (BFP) the largest social safety program in Brazil, covering 12 million families and almost 26%[2] of the population.

Currently, targeting relies on an unverified means test (UMT) using the municipal level Cadastro Unico household surveys. Beneficiaries are chosen based on self-reported income on the Cadastro. Income is not a very reliable indicator because families have an incentive to underreport, parts of income may be in-kind, and there may be seasonal variations. Although the BFP has been ranked as the best among the Latin American countries with regards to targeting efficiency, a 2004 UNDP report states that the BFP undercoverage and inclusion error rates were 59% and 49% respectively.[3]

An alternate way to target families is to check their status on indicators that are highly predictive of poverty. The challenge lies in finding indicators that show high statistical correlation to poverty but at the same time are easy to collect by the municipalities. This memo proposes a predictive targeting model that aims to achieve that balance and improve upon existing targeting performance of the MDS.

B. CHARACTERISTICS OF THE POOR

According to the World Bank GNI data, the average per capita income in Brazil is about 2400 Brazilian Reals per month.[4] Based on the *Pesquisa Nacionale por Amostra de Domicilios* (PNAD), the average poor household earns 50 times less than the national average; the mean income of poor households in Brazil is merely 50 Brazilian Reals per month, although this figure is self-reported and is likely to be grossly underreported when compared to consumption data.[5] Almost 30% of the poor housholds have reported zero income as they are unemployed or self-employed.

In terms of demographics, poor households in Brazil tend to be younger, both in terms of the age of the household head and the number of younger people in the household. Of the poor households, 41% are female-headed, 57% are of mixed color, and 29% are white. Over 54% of the poor live in northeast Brazil and over 65% live in urban areas. Overall, poor households are

[1] World Bank. 2014. Country Data: Brazil. http://data.worldbank.org/country/brazil. 24 Nov 2014.

[2] The Guardian "Social Security is necessary and globally affordable, says UN", June 2011. http://www.theguardian.com/global-development/poverty-matters/2011/feb/21/social-protection-innovation-un

[3] International Poverty Center, United Nationals Development Programme. December 2007. Evaluating the Impact of Brazil's Bolsa Família: Cash Transfer Programmes in Comparative Perspective. IPC Evaluation Note Number 1, Brasilia, Brazil.

[4] World Bank. 2014. Country Data: Brazil. http://data.worldbank.org/country/brazil. 24 Nov 2014.

[5] Bénédicte de la Brière & Kathy Lindert. June, 2005. Reforming Brazil's Cadastro Único to Improve the Targeting of the Bolsa Família Program. http://siteresources.worldbank.org/SOCIALPROTECTION/Resources/0527.pdf.

less educated: 88% of the poor household heads reported as having attended school but only 77% of them claimed to be literate. 4% of poor household have at least one child between 5-14 that they do not send to school.

Unemployment is high among the poor in Brazil. Only 38% of the poor household heads reported as having worked in the past week, as opposed to 68% of the non-poor households. Of poor households, 29% are involved in agricultural work whereas of the non-poor household heads only 11% are.

The poor are less likely to live in permanent dwellings and have smaller homes with straw roofs and mortar walls. They have fewer electrical appliances and are less likely to have access to utitlies such as piped water and waste disposal services. Most of the poor in Brazil own at least some basic assets such as fridges and cell phones, however, they own fewer of them than the non-poor.[6]

C. TARGETING METHODOLOGY

1. Predictive targeting model

The targeting model was developed through analyzing the 2012 PNAD survey responses from 362,451 individuals from 110,509 households. The correlation between hundreds of household level characteristics and poverty was analyzed to finalize a list of 45 indicators with the right balance of predictive power and practical relevance. The indicators cover the following aspects of a household:

- *Demographics:* age of household head, average age of the household
- *Education:* years of education and literacy of household head, any children out of school
- *Employment:* whether household head is currently employed, proportion of family currently employed, household head engaged in agriculture, incidence of child labor
- *Access:* access to personal toilet, running water, waste pickup services
- *Region:* lives in urban vs. rural area, lives in the northeast
- *Asset:* ownership of electric stove, car, computer, washing machine, motorbike, radio, fridge by the household, ownership of cellphone by household head and quality of housing materials

The information about the indicators will be collected through a one-page questionnaire, which is a significant improvement on the 13-page Cadastro Unico or the 58-page PNAD. Instead of interviewing the entire household, the survey is designed to be conducted with only the household head which reduces the time and cost of conducting the survey.

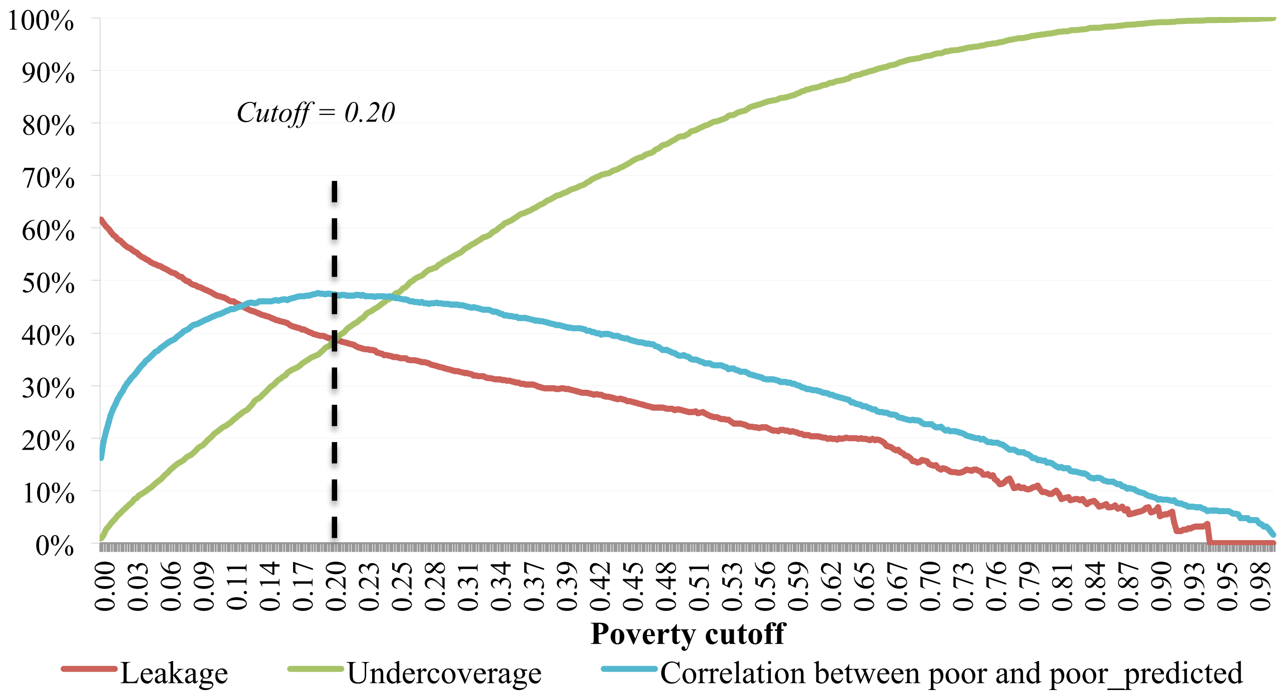
2. Model performance

Based on the survey responses, households receive an aggregate score between 0 and 1 that classifies their probability of being poor. A cut-off score of 0.20 has been chosen since it minimizes the difference between undercoverage and leakage and maximizes prediction power (correlation). This means, if the probability of being poor is 20% or more, that household is eligible to receive social benefits. It also means, at this cut-off, the correlation between the "actual" poor households and the "predicted" poor households is the highest. The threshold can

[6] See Table 2 in the Technical Appendix for a comparison of the poor and non-poor.

be moved if different levels of undercoverage or leakage are desired. Graph 1 shows the relationship between undercoverage, leakage and prediction power (correlation).

Graph 1: Undercoverage, leakage, and correlation with varying poverty cutoffs



Furthermore, Table 1 shows that compared to the Bolsa Familia's Cadastro Unico, the predictive targeting model is able to make notable improvements on the undercoverage and leakage rates.

Table 1: Comparing undercoverage and leakage rates

Targeting Mechanism	Undercoverage	Leakage
Cadastro Unico*	59%	49%
Predictive Targeting Model	39%	39%

* Figures calculated from the 2004 PNAD by Soares, Ribas and Osorio[7]

D. POLICY RECOMMENDATIONS

1. Implement predictive targeting

Compared to the 2004 benchmark of Bolsa Familia's undercoverage of 59% and leakage of 49%, predictive targeting reduces both to 39%, which is a significant improvement for the MDS. This is an opportunity for the Ministry to both increase the net to a greater number of deserving poor and reduce transfers to undeserving households. Hence, it is recommended to use predictive targeting rather than the self-reported income on the Cadastro.

2. Adapt, not replace, the Cadastro Unico survey for predictive targeting

[7] International Poverty Center, United Nationals Development Programme. December 2007. Evaluating the Impact of Brazil's Bolsa Família: Cash Transfer Programmes in Comparative Perspective. IPC Evaluation Note Number 1, Brasilia, Brazil.

Currently, the self-reported income on the municipal level Cadastro Unico is the primary survey instrument for targeting. A 2005 World Bank[8] report cites that the Cadastro identifies almost a million additional households as being poor compared to the PNAD (where there is no incentive to underreport incomes as there are no benefits associated with the PNAD). On the other hand, being a municipal level survey, the Cadastro can be conducted quicker than the PNAD, which is coordinated nationally. In addition, municipalities have been trained on the software to report their data digitally to the Caixa Economica Federal. Thus, this excellent infrastructure should be leveraged but by switching the old survey with the new instrument to allow predictive targeting.

3. Collect survey data digitally
While the data entry for the Cadastro is done on computers at the municipal level, data collection can be made digital through the use of cost-effective tablets and open-source data collection platforms. This will allow for quicker surveying, real-time data entry, more accurate data collection, geo-tagging, and remove the extra layer of manual data entry.

4. Allow for rolling enrollment of beneficiaries
We recognize that predictive targeting is effective to identify the chronic poor but it is not suited to target the transiently poor. Currently, beneficiaries cannot be enrolled unless a new Cadastro is conducted; thus, households that may have fallen into poverty after the survey have to wait until the next Cadastro to receive benefits. As the new survey instrument is shorter, the Ministry should allow rolling enrollment. While this will increase the budget required for the BFP, it will behave as informal insurance during times of economic shock, enabling the MDS to serve the transient poor when necessary.

5. Lower poverty score cutoff for the north and northeast
The MDS should aim for a lower poverty cutoff score in order to reduce the undercoverage rate below 39% (which will result in a leakage rate higher than 39%) in the north and northeast because the poverty rates (>8%) in those areas are higher than the other three-regions (~2%). The mean income in those regions is also much lower; a little above 731 Reals compared to 1259 Reals in the south, southeast and mid-west. For example, undercoverage of 10% will result in a leakage rate of 45%, but 'leaked' social benefits will go to households who may have just missed the poverty score cutoff and might still be quite poor.

E. CONCLUSION

Predictive targeting offers the Ministry of Social Development and Fight against Hunger to improve the coverage of its social benefits to a greater number of poor households and reduce wastage of resources by eliminating a greater number of non-poor households. The model can be implemented with minimal changes to existing administrative operations. The existing infrastructure for the Cadastro Unico will be used with the only change required being the implementation of a new, shorter survey to replace the existing Cadastro Unico survey. This makes the process of data collection faster while at the same time improving accuracy for the better delivery of social security to those who need it the most.

---

[8] Bénédicte de la Brière & Kathy Lindert. June, 2005. Reforming Brazil's Cadastro Único to Improve the Targeting of the Bolsa Família Program. http://siteresources.worldbank.org/SOCIALPROTECTION/Resources/0527.pdf.

III. TECHNICAL APPENDIX

This appendix describes how the model was constructed and how the data was treated in order to fit its needs. The model is deliberately kept simple in order to ensure that indicators can be easily measured in the field and updated though new surveys to be conducted by the Ministry.

A. Model choice

Our model was constructed using the probit multivariate regression method that minimizes the distance between observed and predicted values and outputs the probability (Y) of the household being poor, i.e. a value between 0 and 1. Essentially, we had to determine the set of variables that have the most predictive power in terms of identifying whether a household is in need of assistance. Therefore, we modeled the dependent variable Y i.e. the probability of being poor as being:

$$\hat{Y} = \widehat{\beta_0} + \widehat{\hat{\beta}_1} * X_1 + \widehat{\hat{\beta}_2} * X_2 + \widehat{\hat{\beta}_3} * X_2^2 + \cdots + \widehat{\hat{\beta}_k} * X_k$$

While the probit function allocates each household a score between 0 and 1, we are left with the decision of choosing where the cutoff for a 'poor' household should be. We check for the changes in undercoverage (percentage of poor excluded) and leakage (percentage of non-poor included) as the cutoff score (or the poverty threshold) is varied, and choose a cutoff that minimizes the tradeoff.

B. Data aggregation and missing data

The PNAD collects data both at the household and individual level. During the household survey, each individual member of the household is also interviewed; hence, certain variables, such as education and employment, are at the individual level while others such as household characteristics and asset ownership are at the household level.

In order to predict whether a *household* is poor or not, we strategically aggregated individual level indicators to the household level to ensure variables that have the greatest predictive power for poverty are included. For example, there are two ways to aggregate gender: (1) percentage of household members who are female or (2) the gender of the head of the household. We choose to do the latter because female-headed households are often at a greater risk of poverty.[9] After the variables of interest in the individual level were selected, each variable was aggregated at the household level using the above methodology.

The variables we chose to work with did not have missing responses with one exception. Due to a skip at the beginning of the household section, dwellings that were not permanent were recorded as missing for all the household characteristics and assets. To handle this, we made intelligent decisions as to what assets and items these households would not have and used averages elsewhere. For example, a household that is not permanent will probably not have a washing machine. Also, aggregating individual level data required careful analysis of skips or branching patterns in the questionnaire. Each individual in the sample were not asked all questions. For instance, adults were not asked questions related to child labor or attendance at

[9] Snyder, Anastasia R., Diane K. Mclaughlin, and Jill Findeis. "Household Composition and Poverty among Female-Headed Households with Children: Differences by Race and Residence." *Rural Sociology* 71.4 (2006): 597-624.

school. Therefore, we tracked back each skip to the parent question and consolidated the responses accordingly.

Table 2: Difference in household characteristics and assets between poor and non-poor

Variable	Difference	Variable	Difference	Variable	Difference
Urban area	-0.086 [0.004]**	At least one child works	0.203 [0.044]**	Uses wood/coal burner	0.147 [0.008]**
HH size	0.011 [0.001]**	At least one child doesn't go to school	0.116 [0.013]**	Uses electric/gas burner	-0.148 [0.007]**
Permanent dwelling	-0.091 [0.032]**	HH head is female	0.01 [0.002]**	Owns motorbike	-0.009 [0.002]**
Number of rooms per person	-0.01 [0.001]**	Age of HH head	-0.001 [0.000]**	Owns car	-0.067 [0.002]**
North Region	0.031 [0.003]**	Years of education of HH head	-0.005 [0.000]**	Percentage of adults in HH	-0.169 [0.007]**
Northeast Region	0.073 [0.002]**	HH head is literate	-0.06 [0.004]**	Wall made of straw/taipa	0.285 [0.021]**
Central-West Region	-0.033 [0.002]**	HH head attended school	-0.028 [0.004]**	Wall made of mortar	-0.063 [0.004]**
Southeast Region	-0.04 [0.002]**	HH head had a job in the past week	-0.063 [0.002]**	Owns computer	-0.065 [0.002]**
South Region	-0.031 [0.002]**	Own toilet in HH	-0.196 [0.009]**	Owns washing machine	-0.07 [0.002]**
Race: mixed	0.031 [0.002]**	Connected to water supply network	-0.072 [0.003]**	Owns fridge	-0.169 [0.008]**
Race: yellow	-0.038 [0.006]**	Roof made of straw	0.312 [0.030]**	Owns color TV	-0.104 [0.008]**
Race: black	0.015 [0.003]**	Roof made of concrete	-0.033 [0.002]**	Owns radio	-0.044 [0.003]**
Race: white	-0.036 [0.002]**	Roof made of tile	0.021 [0.002]**	Owns water filter	-0.026 [0.002]**
Race: indigenous	0.052 [0.019]**	Burns trash	0.108 [0.005]**	Owns land phone	-0.058 [0.002]**
Migration in the past year	0.012 [0.004]**	Uses electric light	-0.205 [0.020]**	Owns cellphone (HH head)	-0.073 [0.004]**
HH head does agricultural work	0.089 [0.004]**	Uses no burner	0.154 [0.017]**	Access to running water	-0.164 [0.007]**
73673 Observations. * significant at 5%; ** significant at 1%					

C. Choosing indicators

While choosing indicators, we began by looking for variables where the poor and non-poor differ in their characteristics. Surprisingly, we found statistically significant differences between the poor and non-poor on most indicators. Hence, we decided to undertake a more subjective approach whereby we prioritized the practicality of implementation. We started with a literature review to get a better sense of the characteristics most associated with poor households in Brazil and used that as a starting point for our analysis. We also chose to work with questions that were direct, simple to observe in the field, and had simple skip patterns on the PNAD. For example, asset based indicators took priority since they are easier to note by the field survey staff.

Variable selection was the most subjective part of model building which was revisited several times when we did not get the desired accuracy levels. After creating a shortlist of all variables, we regressed poverty on these variables using a forward stepwise approach to narrow down our list of predictors to those that were identified as statistically significant (45 total, including quadratic and interaction terms discussed below). Before this, we regressed poor with each individual variable to get a sense of the differences between the poor and non-poor. Table 2 (above) lists the initial variables we considered and shows the differences between poor and non-poor households.

Next, we experimented with the form of our regression equation. This included the use of quadratic variables and interaction among variables. For example, since the returns on education are often non-linear, we used a quadratic form for the education variable. We also used interaction among certain variables such as the 'availability of running water' with the variable 'urban'. Interaction takes into account differences that might arise due to systemic constraints that prevent the same patterns of behavior from replicating in rural and urban areas (or regions). For instance, even a rich household in a rural area may not have running water simply because of the lack of availability of public utilities infrastructure in the area. Therefore, by interacting these variables, we accounted for differences in the behavior of people (in similar income brackets) across rural and urban settings.

D. Prediction methodology

Using the indicators that are statistically significant, we get a final list of variables that can be used to assess whether a household is in need of assistance. The survey Part F of the Technical Appendix contains 28 questions that must be asked to each household to get the relevant information. Table 3 shows the indicators in our proposed model that will help predict poverty. Note that we use the 'poverty line' variable in the model to create a normalized poverty score. Brazil is divided into 21 different regions, each with a different poverty line.

Table 3: Variables in the proposed targeting model

Variable	Coefficient	Variable	Coefficient
Urban	-0.004	Owns a motorbike	-0.134**
Female HH head	-0.119**	Owns a car	-0.307**
HH head age	0.083**	Owns a cellphone	-0.128**
HH head age-squared	-0.001**	Owns a fridge	-0.235**
HH head attended school	0.097*	Owns a radio	-0.108**
HH head employed past week	-0.59**	Owns a washing machine	-0.35**
Literate HH head	-0.187**	Owns a washing machine x Urban	0.196*
HH head works in agriculture	0.627**	Owns a water filter	-0.046*
Average age of HH	-0.062**	Owns a color TV	-0.134*
Average age of HH-squared	.00042**	Owns a computer	-0.305**
At least one child in HH works	0.433*	Has own toilet	-0.246**
Household size	-0.004	Has tile roof	-0.182**
Household size x Urban	-0.047**	Has concrete roof	-0.202**
Proportion of people who are literate in HH	0.404**	Has straw roof	-0.045
At least one child in HH doesn't go to school	0.099	Has straw/taipa wall	0.226*
Has permanent dwelling	0.861**	Has mortar wall	0.026
HH average years of education	-0.11**	Has mortar wall x Urban	-0.155*

HH average years of education-squared	0.006**	Has running water	-0.224**
Rooms per person in HH	0.065**	Has running water x Urban	0.169
Regional poverty line	0.01**	Uses gas/electric burner	-0.328**
Northeast region	0.293**	Connected to water supply network	0.058
Proportion of adults in HH who worked last week	-1.707**	Connected to water supply network x Urban	0.001
Burns trash x Urban	0.194*	Constant	-1.106**

* significant at 5%; ** significant at 1%

As discussed earlier, the probit function allocates each household a score between 0 and 1. We then chose a cutoff that minimizes the tradeoff between undercoverage and leakage. Using this cutoff, we predict the 'poor' (called 'poor_predicted' in the model) variable for all households in the data set.

E. Evaluating model performance

To evaluate the performance of our model, we looked at three main outcomes: (1) fit with the model sample, (2) undercoverage and leakage, and (3) fit with the cross-validation sample.

1. Fit with the model sample: We use two metrics to judge the fit with the model sample. The first is McFadden's pseudo R^2 (from probit). Since a higher value indicates a better fit with the data, we used this to compare between the different variants of the model. However, a high R^2 does not necessarily indicate that the model has a good fit because the regression line can systematically over and under-predict the data (indicating bias). Therefore, we also look at the correlation between the variable 'poor' (as given to us) and the variable 'poor_predicted' (as predicted by our model), i.e. what percentage of the households officially designated as poor by the survey does the model also predict as being poor? This gives us a measure of the fit with the sample data. Our final model has the pseudo R^2 of 0.386 and correlation of 47.33% between poor and poor_predicted at a poverty cut-off of 0.20.

2. Undercoverage and leakage: Undercoverage and leakage vary with changes in the cutoff score (i.e. the poverty threshold). If the cutoff is very high, only a few households will be classified as poor and undercoverage will be very high. If instead the cutoff is too low, leakage will be very high and many non-poor households will get classified as poor. We chose the poverty threshold where the undercoverage and leakage curves intersect. This is also the point where the correlation between the variable 'poor' and the variable 'poor_predicted' is the highest. Our final model has an undercoverage rate of 38.7% and leakage rate of 38.7%.

3. Fit with the cross-validation sample: We used the Kaggle platform to test how well our algorithm predicts out of the model-building sample. Our final model correctly predicts poverty for 77.6% of the households on the Kaggle platform.

F. Survey instrument for proposed targeting model

Brazil Poverty Household Survey

Interviewer: For the purposes of this survey, please have the household head tell us about your household's residents, characteristics, and assets.

A0. Household Information

A0.1	A0.2	A0.3	A0.4	A0.5a	A0.5b	A0.5c	A0.6	A0.7
Name	Relation to the household head	Gender (Male - 1 Female - 2)	Age (years)	Attend(ed) school? (Yes-1 No-2)	Years of education	Can he/she read and write?	Did he/she work in the past week?	What in your main occupation?
1. ____________________________	01	1 2		1 2		1 2	1 2	
2. ____________________________		1 2		1 2		1 2	1 2	
3. ____________________________		1 2		1 2		1 2	1 2	
4. ____________________________		1 2		1 2		1 2	1 2	
5. ____________________________		1 2		1 2		1 2	1 2	
6. ____________________________		1 2		1 2		1 2	1 2	
7. ____________________________		1 2		1 2		1 2	1 2	
8. ____________________________		1 2		1 2		1 2	1 2	

Codes for A0.2

01.Household Head	05.Brother/sister inlaw	09. Other relatives
02.HH Head's spouse	06.Son/daughter inlaw	10. Domestic help
03.Own/stepchild	07.Own/step grandchild	11. Other, ______________
04.Own/step sibling	08.Parents/parents inlaw	

Codes for A0.7

01.Own Farm Activities	05.Petty Business/Trade
02.Agricultural Laborer	06.Domestic Worker
03.Casual Labor, Non-Ag	07.Local Government
04.Salaried Employment	08.Other,___________

B0. Household Characteristics

B0.1	Is this household in a rural or urban setting?	Rural	1	
		Urban	2	
B0.2	Which region is this household located in?	_____________		
B0.3	Is this a permanent household/dwelling?	Yes	1	
		No	2	
B0.4	How many rooms are in your household?	___________		
B0.5	What type of roof do you have in your household?	Tile	1	
		Concrete Slab	2	
		Wood equipped	3	
		Zinc	4	
		Seized Wood	5	
		Straw	6	
		Other	888	___________________
		I don't know	666	
		Not Applicable	999	
B0.6	What type of walls do you have in your household?	Masonary	1	
		Milled wood	2	
		Uncoated taipa	3	
		Seized wood	4	
		Straw	5	
		Other	888	___________________
		I don't know	666	
		Not Applicable	999	
B0.7	Do you have your own toilet?	Yes, I have at least one	1	
		No, I share with another household	2	
		No, I don't have a toilet	3	
B0.8	Do you have running water in atleast one room in your household?	Yes	1	
		No	2	Skip to B0.10
		Not Applicable	999	Skip to B0.10
B0.9	What is the source of water in your household?	General distribution network	1	
		Well/Spring	2	
		Other	888	___________________
B0.10	How do you dispose of your trash in your household?	Collected directly	1	
		Burned/buried on property	2	
		Throw in lake/river	3	
		Other	888	___________________
		I don't know	666	
		Not Applicable	999	
B0.11	What type of burner do you use to cook food in your household?	Bottled gas	1	
		Piped gas	2	
		Firewood	3	
		Coal	4	
		Electric power	5	
		Other	888	___________________
		I don't know	666	
		Not Applicable	999	

C0. Household Assets

Does this household own a …. ? (CIRCLE: Yes - 1 No - 2)

C0.1	Motorbike	1	2	**C0.6**	Fridge	1	2
C0.2	Car	1	2	**C0.7**	Freezer	1	2
C0.3	Cellphone	1	2	**C0.8**	Radio	1	2
C0.4	Computer	1	2	**C0.9**	Washing Machine	1	2
C0.5	Color TV	1	2	**C0.10**	Water Filter	1	2