An Evaluation of Geographic Targeting in *Bolsa Alimentação* in Brazil

Daniel O. Gilligan
International Food Policy Research Institute
and University of Maryland

Alinne Veiga
International Food Policy Research Institute

Draft: May 14, 2003

Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Montreal, Canada, July 27-30, 2003

Abstract:

We evaluate the effectiveness of targeting for Brazil's *Bolsa Alimentação*, a nutrition-oriented cash transfer program conditioned on beneficiary participation in health activities. Geographic targeting of program funds relied on adjusted estimates of municipality child stunting prevalence, or a malnutrition map. This evaluation provides new estimates of municipality child stunting prevalence for Brazil. The improved estimates indicate moderate budgetary misallocation from geographic targeting. However, when geographic targeting errors are combined with those arising from an inconsistency between geographic and household targeting objectives, undercoverage of children at greatest risk of stunting is potentially large.

1 Introduction

This paper evaluates the effectiveness of geographic targeting for *Bolsa Alimentação*, the Brazilian government's nutrition-oriented conditional cash transfer program. *Bolsa Alimentação* (BA) is designed to reduce nutritional deficiencies and infant mortality among poor households in Brazil. The program offers demand-side incentives to use nutrition-oriented services through monthly money transfers to very low income families with young children or to pregnant or lactating women. For poor households with young children, the cash transfer—R$15 (roughly US$4.24)[1] per month per eligible child for up to three children—is conditional on mothers committing to a 'Charter of Responsibilities' that requires regular child check-ups, compliance with vaccination schedules, and health and nutrition education. Households with pregnant or lactating women receive R$15 per month per eligible woman for six months, conditional on receiving prenatal care. When fully implemented, the program is expected to benefit more than 2.77 million children with bolsas totaling R$499.1 million and 803,000 women with bolsas of R$72.3 million in a total of 2.5 million Brazilian households. As of December 2002, the Ministry of Health reported that 1.39 million bolsas worth R$118.8 million had been distributed.

Beneficiary households with young children were selected for the program in a two-stage targeting process. The first stage involved geographic targeting, in which the Ministry of Health (MOH) allocated program funding to each of Brazil's 5,561 municipalities according to the estimated number of children in the municipality suffering from malnutrition. Because data on malnutrition prevalence are not available for most municipalities, the government relied on adjusted estimates of municipality child stunting prevalence calculatedin a study by University of São Paolo (USP) researchers Benício and Monteiro (1997) using household survey and census data. Geographic targeting of this kind has led to substantial improvements in impact on the target population in similar programs, such as Mexico's PROGRESA program.[2] In the second stage of targeting, each municipality identified beneficiary households primarily using a means test: per capita household income less than half the minimum wage (a threshold of R$90/month) and presence of children aged 6 months to 6 years 11 months.

Recent evidence offers strong support for geographic targeting of social programs.[3] Baker and Grosh (1994) document substantial improvements in program impact through geographic targeting of needs-based programs. Ravallion and Wodon (1999) argue that geographic targeting may be justified because location-specific differences in levels of well-being cannot be explained by individual characteristics observable to policymakers. Morris et al (2000) show that a nutrition program in Honduras was more effective when located in communities with the highest malnutrition rates. An evaluation of Mexico's PROGRESA program by Skoufias et al (2001) showed that PROGRESA's 'marginality index' approach to geographic targeting improved program impact and provided targeting accuracy comparable to a methodology for predicting local poverty rates that is analogous to the approach used for predicting malnutrition in BA.

---

[1] Throughout this report, US$ figures are based on the December 31, 2002, exchange rate reported by the US Treasury (R$3.535/US$).
[2] See Skoufias et al (2001).
[3] See Coady et al (2002) for a review of targeting methodologies. Out of 100 social safety programs reviewed, 43 programs relied on geographic targeting, usually in conjunction with other household targeting criteria.

This report evaluates geographic targeting for the child component of *Bolsa Alimentação*. Based on our understanding of the program design, geographic targeting of pregnant or lactating women was limited, so it is omitted from this evaluation. Despite this exclusion, our analysis applies to the bulk of program resources since children represent the primary group of beneficiaries and 87.4 percent of the annual budget. The Ministry of Health calculated the number of child bolsas allocated to each municipality by applying the estimates of municipality child stunting rates from the USP study to current population figures, with a few adjustments. Therefore, the evaluation focuses on the accuracy and robustness of the USP model of child malnutrition prevalence. We will refer to this as the "original" model. In what follows, we present estimates of stunting prevalence from new models that use more recent data with a larger set of household- and municipality-level predictors.

The general approach to predicting malnutrition prevalence in the original model and in this evaluation is based on an imputation technique known as "small area estimation". It has recently been applied to geographic analyses of poverty, in what is commonly called "poverty mapping". Using data from a nationally representative sample of households that includes anthropometric measures of nutritional status for children, a model of malnutrition prevalence is estimated, possibly allowing for differences in effects between urban and rural areas and across the great regions of Brazil. The set of right-hand-side variables in this model may include child-, household- and municipality-level characteristics. However, the choice of variables is restricted to those that are available in both the sample survey and the census or to municipality-level variables from other sources. The expected malnutrition prevalence for each municipality is then imputed through a procedure that combines the parameter estimates from the first-stage model with recent census data for the predictors. Municipality average prevalence is the mean of child-level predictions in the census if household census data are available. Otherwise, researchers must make predictions using the municipality sum or mean of the regressors from the census, which leads to less precise estimates.

The greatest limitation of the original estimates of malnutrition prevalence is that they are based on a simple empirical model of malnutrition prevalence that uses only a handful of regressors. The evaluation model developed here shows that such stringent restrictions on the set of predictors is unnecessary and weakens predictive accuracy. Another weakness of targeting performance is that the malnutrition estimates used are based on data collected a decade before the start of the program. These estimates are derived from the 1989 PNSN, the *Pesquisa Nacional de Saúde e Nutrição* (National Survey on Health and Nutrition), and the 1991 census. The PNSN is a *nearly* representative household sample survey that covered the urban areas of the five great regions of Brazil (North, Northeast, Center, Southeast and South) and all the rural areas except for the sparsely populated North region.

The models developed for this evaluation include a number of refinements to the general approach used by BA. First, more recent and complete data are used. The evaluation models are based on the 1996 PNDS, the *Pesquisa Nacional sobre Demografia e Saude* (National Survey of Demography and Health) and a 12 percent sample of household records from the long form of the 2000 Census, or *Censo Demográfico*. Like the 1989 PNSN, the 1996 PNDS sample is nationally representative with the exception of the rural North region. The availability of household records from the census, even for a 12 percent sample, makes it possible to add a number of important household demographic variables as predictors to the malnutrition model. It also improves prediction by the obvious increase in coverage of the population. These data are supplemented with a large set of variables of municipality characteristics from the 2000 census and elsewhere found in the BIM data set, the *Base de Informações Municipais*. The updated data should more accurately represent the nutritional status of children at the time that bolsas were distributed. A limitation of these data is that the PNDS sample and the census were collected 4 years apart. This could reduce the precision of malnutrition prevalence estimates if the distribution of the regressors or their effects on malnutrition has changed during that time. The second refinement to the BA methodology is that the evaluation models are subjected to a broader set of specification tests: (i) variable selection is tested by a number of criteria including stepwise elimination of regressors and minimization of RMSE; (ii) the use of models separated by age, urban/rural and regional location is tested against a unified model; and (iii) the role of unobserved municipality-level effects is addressed. In response to these tests, we present three sets of estimates of municipality stunting prevalence that account for different specifications of the first-stage model in order to demonstrate the robustness of the approach. The merits of each set of estimates are discussed and the figures are compared to those used to allocate the program budget. Although the models used in the evaluation are still subject to considerable error in estimating malnutrition prevalence, we present evidence that suggests they are more precisely estimated than the original model used to allocate the program budget. The evaluation models appear to perform better in differentiating malnutrition levels across municipalities.

Results of this analysis show that, although the original malnutrition estimates used by MOH undoubtedly led to targeting gains and boosted program impact relative to a program without geographic targeting, it is likely that significant misallocation of the program budget occurred in some areas. For example, the state of Maranhao, which had among the very worst malnutrition rates, also had the third largest *real*-value shortfall in its BA budget at R$8.7 million, according to the evaluation methodology.

The evaluation is organized as follows. Section 2 describes the approach to estimating municipality malnutrition prevalence in the original model and, to the extent that we understand them, subsequent adjustments performed by the Ministry of Health that yielded the final allocation of bolsas to municipalities. Section 3 describes the empirical models of malnutrition prevalence used to evaluate the geographic targeting, addresses a number of empirical issues, and presents results of the model. Section 4 compares the allocations of bolsas that would arise based on these predictions to the actual allocation. Other measures of targeting performance are also presented. We offer some conclusions and recommendations in Section 5.

2 The Methodology for Geographic Allocation of Bolsas under *Bolsa Alimentação*

In keeping with the objective of *Bolsa Alimentação* to reduce childhood malnutrition in Brazil, the Ministry of Health targeted the child component of program funds to the municipalities in proportion to their prevalence of malnourished children. The municipalities were then responsible for identifying eligible children in the second stage of targeting. Allowing municipalities to manage household targeting was consistent with the new policy of decentralized decision-making in which greater authority is vested in the municipalities. This overall approach to targeting, and particularly linking geographic targeting to malnutrition prevalence, was a sensible approach to targeting given the program objectives. The use of malnutrition prevalence as the indicator for geographic targeting raises two issues in program design: (i) the choice of malnutrition measure, and (ii) the method of estimating malnutrition prevalence at the municipality level. These issues are now addressed in turn.

The Choice of Malnutrition Indicator

There is broad agreement that anthropometric measures of physical status offer the best practical indicator of nutritional status. When summarizing the nutritional status of a population, the choice of anthropometry measure depends on the characteristics of the population and on the goal of the analysis. For a target population of children under the age of five or six, researchers typically use either weight-for-age or height-for-age. The goals of this program argue for using height-for-age for two reasons. First, the WHO Expert Committee on Nutrition recommends using height-for-age for regional targeting of economic and health interventions (WHO, 1995). Height-for-age is a better indicator of long-run nutritional achievement and so is a more reliable one-time indicator for a program that will last several years. Second, the prevalence of *stunting* (a height-for-age Z-score less than -2) is a more serious nutritional problem in Brazil than is being *underweight* (a weight-for-age Z-score less than -2). In the nationally representative 1996 PNDS sample, the prevalence of stunting among children aged 6-59 months is 11.3%, while the share of children in this age group that are underweight is 6.2%. The difference in these measures of nutritional status probably indicates that inadequate retention of micronutrients is a bigger source of malnutrition than inadequate consumption of calories in Brazil.

To evaluate nutritional status using height-for-age, the sample must be compared to a growth standard or reference population. In the Brazilian PNDS data, anthropometry measures are based on the Standard Deviation-derived Growth Reference Curves constructed from the NCHS/WHO Reference Population. The three most common measures of height-for-age are Z-scores (standard deviation scores), percentiles, or percent of median values for the reference population. We use height-for-age Z-scores (HAZ)—or the number of standard deviations from the reference median value— as recommended by WHO (1995), as the dependent variable in all of the models below. A shortcoming of using percentiles of the reference distribution is that an identical change in percentiles represents different changes in absolute height, depending on the location on the distribution. This makes summary statistical measures such as mean or standard deviation difficult to interpret. Also, prediction of low height-for-age through regression is made more difficult by the use of percentiles. Near the extremes of the reference distribution percentile changes will be small, although these can represent large changes in absolute height. In a country such as Brazil with fairly low stunting prevalence, this effect could be important.

The Z-score, on the other hand, is appropriate for monitoring individual status or for summary population statistics. Interpretation is not dependent on the location in the distribution, since a fixed Z-score interval represents a fixed change in height. For the BA program, the original model used HAZ as the measure of nutritional status in its prediction of municipality malnutrition prevalence as recommended by WHO (1995). However, the Ministry of Health then adjusted these figures so that the average prevalence in three aggregated regions was equivalent to the regional prevalence of children with weight-for-age below the tenth percentile (WAP10), as discussed below.

Using the share of children that fall below an anthropometric Z-score threshold for nutritional targeting is analogous to targeting the poor with a headcount index of consumption poverty, which measures the share of a regional population with per capita consumption below a poverty line. These measures indicate *prevalence*, but targeting resources to municipalities based on the *depth* or *severity* of malnutrition may also be reasonable. In the poverty literature, a common measure of the depth of poverty is the poverty gap or P_1, the average distance from the poverty line for the poor. The severity of poverty is often measured by P_2, the squared poverty gap. Analogous measures for malnutrition could also be developed. Alternatively, nutritionists have proposed two other measures of the severity of malnutrition: the "standardized prevalence" (Mora, 1989) and the "minimum prevalence" (Monteiro, 1991). The "standardized prevalence" is the proportion of individuals in the observed distribution who fall outside—and to the left of—the reference distribution. The "minimum prevalence" estimate, proposed by a coauthor of this report, is a bit more complicated. It is based on the assumption that in most developing countries there is a malnourished population with an HAZ distribution to the left of the reference distribution and a non-malnourished population with the same HAZ distribution as the reference distribution. The "minimum prevalence" estimate then measures the malnourished proportion of the total population by netting out an estimate of the size of the non-malnourished population. This study will focus exclusively on the measure of malnutrition prevalence in order to be consistent with the existing design of the program.

Estimating Malnutrition Prevalence in the Municipalities: Malnutrition Mapping

As in most countries, there is no representative sample survey for Brazil that provides reliable estimates of child malnutrition rates at a level of regional disaggregation as low as the municipality. Therefore, the government had to identify alternative methods to estimate child malnutrition prevalence across municipalities. One approach would have required conducting a national census of the height of all first-grade students in the country.[4] Morris and Flores (2002) show that a school height census in Honduras provided a reliable and valid tool for small-area targeting of nutrition interventions. Their study found that municipality mean height-for-age Z-scores in a school height census were closely correlated (Spearman's rank correlation = 0.74) with sample survey estimates of municipality mean HAZ in children less than age five. They claim that a school height census should provide a comparably accurate proxy for municipality stunting prevalence because height-for-age Z-scores are normally distributed with little variation in standard deviation across localities. Despite these benefits a school height census was not

[4] See Morris, Saul S., Pedro Olinto and Rafael Flores, "Consultancy to Support the Design and Evaluation of *Bolsa Alimentação*," IFPRI, Final report, September 7, 2001.

conducted in Brazil, presumably because of the short timeframe available for designing and implementing *Bolsa Alimentação* and the cost of administering the school height census.

In the absence of reliable data or a proxy from a school height census, statistical analysis can be used to estimate municipality-level malnutrition prevalence by combining nutrition-oriented sample surveys and census data. The technique involves applying parameter estimates of the relationship of nutritional status to child, household, and municipality characteristics from a national sample survey to census data for the same characteristics in order to obtain estimates of nutritional status at a low level of geography. This approach is closely related to that of small area estimation used by demographers.[5] It has recently become popular as a method for obtaining local-area estimates of consumption or income poverty, or "poverty mapping."[6] The methodology has benefited from recent advances by Elbers et al (2003) in an application to poverty measurement using sample-based estimates of per capita consumption and *household* census records. Their approach involves a careful accounting for the sources of variance in the error term in the imputed local poverty measure. This makes it possible to obtain welfare measures at a very low level of disaggregation, and to be informed about the tradeoff between precision and narrowness of location.[7]

The general technique of poverty mapping can be applied to estimating local malnutrition prevalence, or "malnutrition mapping".[8] A variant of this technique was employed in the original model of municipality malnutrition prevalence used to allocate the program budget. However, household census data was not available at the time that model was developed. This restricted the set of variables available for the prediction and made it impossible to calculate the precision of the malnutrition prevalence estimates. The evaluation models presented here benefit from having the 12 percent sample of household data from the 2000 Census, which enables us to use a large set of predictors for household demographics. However, we received this census data only recently, and so have not yet calculated the standard errors of our malnutrition prevalence estimates, as discussed below.

The Original Model of Malnutrition Prevalence and the Allocation of Bolsas to Municipalities

The original estimates of malnutrition prevalence in Brazil were constructed in the 1997 USP study for use by the Ministry of Health for the nutrition program that preceded *Bolsa Alimentação* called *Incentivo de Combate às Carências Nutricionais* (ICCN). That approach to malnutrition mapping was based on a logistic regression of stunting prevalence in children under age five from the 1989 PNSN sample survey,

(1) $$HAZM2_i^S = X_i^S \beta^S + u_i^S \qquad \text{where } HAZM2_i^S = 1 \text{ if } HAZ_i\text{<-2}$$

[5] See Purcell and Kish (1980).

[6] See, for example, Hentschel et al (1998); Minot (2000); Simler and Nhate (2002); and Minot and Baulch (2002). For regional and UF-level estimates of poverty for Brazil created through small area estimation using the PPV and PNAD data sets, see Elbers et al (2003).

[7] An unresolved issue concerning geographic targeting of nutrition interventions is whether these statistical approaches to estimating local malnutrition prevalence would provide more accurate estimates than a school height census.

[8] Fujii (2002) appears to be the first example of malnutrition mapping, which applies the Elbers et al (2003) approach for Cambodia.

= 0 otherwise,

and where the superscript *S* denotes the sample survey, *i* indexes children, X_i^S is an $n \times k$ matrix of explanatory variables, β^S is a $k \times 1$ vector of parameters, and u_i^S is a $n \times 1$ vector of random disturbances from a logistic distribution with mean zero.

The regressors used to predict stunting prevalence in the sample included up to two indicator variables for household head income level relative to region-specific thresholds; a dummy variable for whether the mother was illiterate; and two dummy variables for household sanitation which were coded as "semi-adequate" and "inadequate". In urban areas, a household with "adequate" sanitation, the reference category, had piped water from the public system and either a sewerage or septic system for waste. "Semi-adequate" referred to piped water but no sewer system or septic pit, and "inadequate" referred to no piped water or sewer or septic. In rural areas, "adequate" was a household with piped water (regardless of source) and "inadequate" referred to the absence of piped water. There was no "semi-adequate" category for rural areas.

The sample was divided into seven regional strata for estimation: northeast urban, northeast rural, center-south urban, center-south rural, all urban Brazil (to provide estimates for urban municipalities in the North), all rural Brazil (to provide estimates for rural municipalities in the North), and the urban area of Pará state. A Likelihood Ratio Test determined the final set of variables used in each regional model.

The next stage of the analysis involved calculating the predicted stunting prevalence corresponding to every combination of explanatory variables in X_i^S for each regional model, a procedure made feasible by the use of only binary variables as regressors. For each municipality in the 1991 census, the number of children under 5 in each combination of explanatory variables was obtained, as was the number of children in rural and urban areas. Rural and urban prevalence for each municipality was given by the sum of predicted prevalence times the share of children in each category of explanatory variables. The estimates of stunting prevalence for the 4,491 municipalities in 1991 was then calculated as the (child) population-weighted average of urban and rural estimated prevalence for each municipality. The 1,070 municipalities created between 1991 and 2001 received (i) the state weighted average prevalence, (ii) the state simple average prevalence, or (iii) the average prevalence of the municipalities from which they derived, depending on when they were created. It is interesting to note that using a limited set of binary regressors made it possible to fully classify all children into one of the combinations of explanatory variables, ultimately obtaining the number of children in each group. In a sense, this approach fully recovers the household census records for these variables, although only the number of children in the municipality in each category was known. However, this benefit to that approach is outweighed by the loss in predictive power that comes from using only 3-5 binary predictors in each regional model.

Apparently, two adjustments were made by the MOH to these stunting estimates to arrive at the figures used for targeting by BA. First, the stunting estimates were adjusted for the regional average reduction in child stunting prevalence from the period of the 1989 PNSN to that of the 1996 PNDS. For this adjustment, the five great regions of Brazil were grouped into three:

North, Northeast and Center-South. It appears that an adjustment factor was calculated for each of these three regions and applied to the corresponding municipalities' estimates of 1991 malnutrition prevalence. These adjustment factors are believed to be from the report, "Evolução da Mortalidade Infantil e do Retardo de Crescimento nos Anos 90: Causas e Impactos sobre Desigualdades Regionais" – "Changes In Poverty-Related Health Indicators In Brazil: Causes And Impact On Regional Inequities". They are presented in Table 1.

Table 1: Regional Adjustment Factors for Decline in Stunting Prevalence, 1989-1996

Region	Share of children with *HAZ*<-2 in 1989 PNSN, %	Share of children with *HAZ*<-2 in 1996 PNDS, %	Reduction in Stunting, 1989–1996, %	Adjustment Factor
Northeast	27.3	17.9	-34.4	0.656
North	23.0	16.6	-27.8	0.722
Center-South	8.6	5.6	-34.9	0.651

Source: "Evolução da Mortalidade Infantil e do Retardo de Crescimento nos Anos 90: Causas e Impactos sobre Desigualdades Regionais" - "Velhos e Novos Males da Saúde no Brasil," São Paulo p. 400.

We believe that the second adjustment was to multiply the "updated" stunting prevalence estimates by the ratio of the average regional prevalence of children with weight-for-age below the tenth percentile (WAP10) to the average regional prevalence of children with height-for-age Z-scores less than -2 (HAZM2). This adjustment factor was 1.2 for municipalities in the Northeast. Adjustment factors for the other two regions are unknown. We were told that the second adjustment was requested by BA staff who wanted the malnutrition estimates to reflect low weight-for-age because they were considering using observed weight-for-age in the second-stage household targeting as a criterion for eligibility. The tenth percentile threshold for weight-for-age was used to raise the size of the target population to desired levels. Using the standard weight-for-age cutoff for underweight prevalence of less than -2 Z-scores, which corresponds to the 2.3 percentile in a well-nourished population, would have resulted in a target population of 6.2 percent of children according to the 1996 PNDS sample. Using the tenth percentile of weight-for-age as the eligibility threshold increases the target population to 19.7 percent of children. However, the use of this crude adjustment factor applied to estimates of stunting prevalence probably led to inaccurate targeting of weight-for-age prevalence below the tenth percentile. In the end, observed child weight-for-age was not used as a criterion for eligibility for the program because of concerns about the perverse incentives for good nutrition that this would create.

The allocation of bolsas to the municipalities was based directly on these adjusted estimates of malnutrition prevalence. The number of bolsas awarded to each municipality was equal to the adjusted estimate of malnutrition prevalence times a projection of the year-2000 population of children of age at least 6 months and less than 7 years based on the 1991 census.

A number of shortcomings affect the reliability of the results of this methodology. The first is the absence of household census data, which would permit the use of a more robust methodology that would also enable statistical tests of equality of malnutrition prevalence across two municipalities. Another substantial limitation of this approach is the use of only three-to-five regressors, all of them binary, in the first-stage model of stunting prevalence. Such a limited set of binary regressors impedes the model's ability to differentiate the incidence of malnutrition between municipalities. It is possible, for example, that two municipalities have roughly similar distributions of children according to threshold levels of income, female illiteracy, and sanitation, but differ substantially in the distribution of these variables away from the thresholds or with respect to other important characteristics such as infrastructure, health facilities, and political participation. The effect of the two *ex post* adjustments described above is to increase the effect of the region of location of a municipality on its estimated malnutrition prevalence. As just noted, these regional effects were already responsible for a great deal of the variation in municipality malnutrition prevalence estimates because regional models were estimated with few regressors. The projections of the number of children in each municipality are not a significant source of error. The Pearson correlation of these estimates with the number of children in the 2000 census was 0.888. In the next section, we expand and refine the methodology used in the original model to try to obtain more accurate estimates of municipality malnutrition prevalence.

3 The Evaluation Model for Geographic Allocation of Bolsas under *Bolsa Alimentação*

In order to evaluate the accuracy of the adjusted estimates of malnutrition prevalence used by *Bolsa Alimentação*, we estimate revised models of stunting prevalence that include a number of changes in methodology, more recent data, and a substantially expanded set of predictors that should provide more precise malnutrition estimates. The first-stage stunting models are estimated using the 1996 Brazilian Demographic and Health Survey (PNDS) supplemented by IBGE's BIM data, a large data set of municipality characteristics on topics such as school access and quality, health care providers, and political participation. Parameters from this model are then applied to data from the 2000 *Censo Demográfico* and the BIM data to obtain estimates of municipality stunting prevalence. This section presents the details of three models of municipality stunting prevalence. For each model, we discuss the choice of explanatory variables, results of various specification tests, and a summary of results. The three models differ by choice of regressors and by allowance for regional differences in effects of predictors on the probability of being stunted. These three models are presented in order to provide an indication of the sensitivity of the methodology to model specification, and to allow the government to consider the assumptions that underlie any of the estimates it might ultimately use. We use these results to demonstrate the limitations in the accuracy of prediction in these types of models and consider the implications for targeting.

Methodology

The general approach to constructing estimates of municipality stunting prevalence is the same for all three sets of estimates presented below. The procedure occurs in two stages. Our first-stage model predicts the height-for-age Z-score for the sample of 3,610 children aged 6-59 months in the 1996 PNDS,

(2) $$HAZ_{ij}^{S} = X_{ij}^{S}\beta^{S} + u_{ij}^{S}, \text{ where } u_{ij}^{S} = \eta_{j}^{S} + \varepsilon_{ij}^{S}.$$ [9]

[9] Although the target population for *Bolsa Alimentação* is children aged 6 months to 6 years 11 months, the PNDS only provides anthropometry data for children aged 6-59 months. A benefit of this restriction is that it enables comparability of samples between these new estimates and the original study.

Here, i indexes individuals, and j indexes municipalities. X_{ij}^{S} and β^{S} are as defined in (1). u_{i}^{S} is a vector of random disturbances with two components: a municipality-specific effect, η_{j}^{S}, and a classical zero-mean individual disturbance, ε_{ij}^{S}, with variance σ^{S}. A continuous model of height achievement is estimated at this stage rather than a discrete threshold model of stunting prevalence, as in the original model, because the continuous dependent variable takes better advantage of the information available and because the estimates are less sensitive to distributional assumptions about the error term.

In the second stage, we calculate the expected height-for-age Z-score for children in the census sample using the parameters estimated from the PNDS and the census data for the corresponding regressors, and accounting for the possibility of unobserved municipality effects,

$$(3) \qquad E\left[HAZ_{ij}^{C}\right] = E\left[X_{ij}^{C}\hat{\beta}^{S} \mid \eta_{j}^{S}\right],$$

where the superscript C denotes census data, X_{ij}^{C} represents census sample observations for predictors, and η_{j}^{S} controls for omitted municipality effects from the first-stage regression. Accounting for unobserved municipality effects at this stage borrows from the approach of Elbers et al (2003). If tests for the presence of omitted municipality effects are significant, these can be controlled for in the estimation.

In the final stage of the calculation, the expected probability of being stunted for the ith child in the jth municipality is given by,

$$(4) \qquad P_{ij} = \Phi\left(\frac{\mu - X_{ij}^{C}\hat{\beta}^{S}}{\hat{\sigma}^{S}}\right),$$

where Φ is the cumulative normal density function, $\mu = -2$ is the Z-score cutoff for stunting prevalence, and $\hat{\sigma}^{S}$ is the estimated standard error of the first-stage model. Based on (4), the estimated municipality stunting prevalence is the average estimated prevalence of all children in the municipality,

$$(5) \qquad P_{ij} = \frac{1}{M_{j}}\sum_{i=1}^{M_{j}} P_{ij},$$

where M_{j} is the number of children in the jth municipality. In practice, it is necessary to test whether the parameter estimates in (2) differ by age of the children, urban and rural area, or location. If so, separate models for (2) are estimated on each cohort, and the estimate of municipality stunting prevalence in (5) is a population-weighted estimate of cohort-specific stunting prevalence.

Variable Selection and Comparability

The selection of variables for estimation of the HAZ model in (2) was guided by the need to maximize the predictive accuracy of the model, while being sensitive to concerns about overfitting the first-stage estimates. This objective suggested a broader set of regressors than that used in the original study. We sought variables at various levels of aggregation in order to account for local differences in child, household and municipality characteristics. The list of potential child- and household-level explanatory variables was restricted to those available in both the 1996 PNDS and the 2000 Census. With household level census data available, it was still possible to include many control variables for household demographics. The PNDS and census variables are not comparable on these topics. The analysis also benefited from having access to more than 400 variables of municipality characteristics from the BIM data, providing a large set of local data to control for municipality-level effects.

The process of selecting variables first involved identifying a set of child-, household-, and municipality-level variables available in both data sets that were plausible predictors of stunting. From this core set of potential right-hand-side variables, various techniques were used to identify which set of regressors would be included in the final models. These tests of variable selection and the resulting models are described in the next section. Here, we describe the core set of variables included in the selection process.

Where possible, continuous rather than binary variables were used to provide greater variability across observations. In addition, many continuous variables were made non-linear through use of quadratic terms or clustering of levels into categories. For example, household size included both linear and quadratic terms and various specifications of mother's education were tried, including grouping education levels into dummy variables for groupings of completed years of education (0-3, 4, 5-7, 8, 9-11, >12 years).

The list of child- and household-level variables included as potential predictors in the HAZ model includes child age as a quadratic or in monthly groupings; child gender; an interaction of gender and age; child birth order; a dummy for female household headship; a quadratic for mother's age; dummy variables for groupings of numbers of completed years of mother's education (0-3, 4, 5-7, 8, 9-11, >12 years); a dummy for whether the mother's husband has at least 4 years of education; a quadratic for household size; the under-5 dependency ratio; number of rooms and number of bedrooms in the dwelling; number of people per room; a dummy variables for piped water; a dummy variable for well or spring water in the house for any use; and a dummy for electricity. Household asset variables considered were dummy variables for presence of a radio, a refrigerator, a television, a VCR, a washing machine, and a car. From the large set of municipality explanatory variables available, 37 regressors were selected on the following topics: numbers of hospitals and clinics per capita; birth rate; death rates by age/reason; primary schools (enrollment, student-teacher ratio, class size); literacy rate; employment rate; fiscal expenditures by type; sanitation; housing type; population density; and household head income. We also included an interaction term for municipality infant mortality rates and household observations on mother's education in completed years. This allows the nutritional benefits of mother's education to depend on the local health environment.

Based on the recommendations of the WHO Expert Committee on Nutrition (WHO, 1995) on the use of anthropometric data, observations with child height-for-age Z-scores below –5 or above 3 were treated as outliers and omitted from the sample. This led to dropping 44 observations, all but three of which had Z-scores over 3. Table 2 provides summary statistics for the full set of explanatory variables used as possible regressors in the first-stage model.

Elbers et al (2003) note that in mapping exercises like this one comparability of the regressors between the sample survey and the census data is important to the quality of the prediction. Variables should have similar definitions in the two data sets and should have similar distributions. Some variables under consideration were dropped during this analysis for these reasons. An important example is a set of more precise sanitation variables on type of latrines and sewerage. Information on sanitation is not gathered in the same way in the PNDS and the 2000 census. Other important variables that were maintained in the set of regressors presented a significant challenge in this regard. For example, mothers of children in the census are not identified. In order to construct variables on mother's age and education, it was necessary to infer who the likely mother is based on the information provided. We used the following conditions successively to identify probable mothers in the census:

(i) a female head of family with children under age 7
(ii) the female spouse of head of family if a child under age 7 is the offspring of the head of family
(iii) a female daughter of head of family who has given birth if head of family has grandchildren under 7
(iv) an adult female who has given birth if the head of family has grandchildren under 7

We believe this approach was fairly successful at identifying likely mothers of the children in the census. In the case of the first category, mothers are clearly identified. Women in the second category have a high probability of being the birth mother. Those that are not are likely to be the primary caregiver. We regard conditions (iii) and (iv) as less likely to identify the actual mother. To check this approach, we looked at the data for five UFs and found that 92.7 percent of children had a mother identified by the first two categories alone. The third condition identified a mother for another 2.8 percent of children under 7, and the fourth condition for 0.3 percent of children.

Next, we tested the comparability of the distributions of the variables in Table 2 between the PNDS and census data sets. Two tests were used, a t-test for equality of means and a Kolmogorov-Smirnov test, a non-parametric test of equality of distributions. The results of these tests are provided in Table A.1 in Appendix A. Of the 30 child and household variables tested, equality of the means of the variables was rejected 11 times and the equality of the distributions was rejected by the Kolmogorov-Smirnov test 11 times. Equality of the distributions was rejected under both tests for 8 of the 30 variables. These results suggest reasonable comparability in the data sets, although differences in the distributions of predictors failing the tests, such as household size and asset variables, could be a source of error in the predictions of stunting prevalence. For mother's age and education, there is mixed support for our strategy for identifying likely mothers in the census. Results of the Kolmogorov-Smirnov test fail to reject equality of the distribution of all of the mother's education variables, but rejects equality of the distribution of mother's age in the PNDS and census. It is not possible to know whether possible differences in mothers' ages in the two data sets is due to sampling error or to our definition of mothers in the census. T-tests for equality of means reject equality for 2 of the 5 variables for mother's education, but fail to reject equality of the means of mother's age.

Table 2: Summary Statistics for Variables in Stunting Model*

	Dependent Variable and Sample Stunting Prevalence	**Mean**	**Std Dev**	**Min**	**Max**
HAZ	Z-score of height-for-age	-0.581	1.266	-4.99	3
HAZM2	Dummy for HAZ<-2	0.129	0.335	0	1
	Child and Household Explanatory Variables				
Age7m	Dummy for child age of 7 months	0.015	0.120	0	1
Age8m	Dummy for child age of 8 months	0.018	0.133	0	1
Age9m	Dummy for child age of 9 months	0.024	0.153	0	1
Age10m	Dummy for child age of 10 months	0.023	0.150	0	1
Age11m	Dummy for child age of 11 months	0.023	0.149	0	1
Age1223	Dummy for child age of 12-23 months	0.223	0.416	0	1
Ageyrs	Age in completed years	2.195	1.328	0	4
Ageysq	Age in completed years squared	6.580	5.859	0	16
Girl	Dummy for child gender = girl	0.490	0.500	0	1
Agey_Girl	Age in completed years times girl dummy	1.070	1.438	0	4
Birthord	Child's birth order for live births	2.735	2.220	1	18
Rage	Respondent's (mother's) age	28.000	6.406	15	49
Ragesq	Respondent's (mother's) age squared	825.030	381.281	225	2401
Headfem	Dummy for female household headship	0.123	0.328	0	1
Hhsize	Number of household members	5.565	2.330	2	18
Hhsizesq	Number of household members squared	36.401	35.754	4	324
Redy03	Respondent has 0-3 yrs of education	0.313	0.464	0	1
Redy4	Respondent has 4 yrs of education	0.173	0.378	0	1
Redy57	Respondent has 5-7 yrs of education	0.216	0.411	0	1
Redy8	Respondent has 8 yrs of education	0.086	0.280	0	1
Redy911	Respondent has 9-11 yrs of education	0.176	0.381	0	1
Ped4ov	1 if husband completed at least 4 yrs educ	0.661	0.473	0	1
Headfem	1 if HH head is female	0.123	0.328	0	1
U5deprat	Under 5 dependency ratio	0.328	0.134	0	0.8
Nroom	No. rooms in house	5.015	2.005	1	17
Nbedroom	No. rooms used for sleeping	2.082	0.875	1	8
Peopleroom	No. HH members per room	1.323	0.919	0.231	10
Pipedwater	1 if water piped into house/yard	0.702	0.458	0	1
Wellspringin	1 if well/spring water in house/yard	0.157	0.364	0	1
Electr	1 if has electricity	0.896	0.305	0	1
Radio	1 if has radio	0.801	0.399	0	1
Refrig	1 if has refrigerator	0.622	0.485	0	1
Vcr	1 if has VCR	0.143	0.350	0	1
Wshmach	1 if has washing machine	0.261	0.439	0	1
Tvset	1 if has television	0.525	0.499	0	1
Car	1 if has car	0.178	0.382	0	1

(continued…)

Table 2: (continued)

	Municipality Explanatory Variables°	Mean	Std Dev	Min	Max
Birthpcm	Live births per capita, 1998	0.030	0.013	0.001	0.112
Dthr01_redyrs	Interaction: < 1 death rate & mother's educ.	0.129	0.117	0	1.014
Dthr14_redyrs	Interact: death rate, 1-4 y.o.'s & mother's educ.	0.006	0.006	0	0.081
Hosp00pc	Number of hospitals per capita	4.34E-05	4.18E-05	0	2.79E-04
Hospbed00pc	Number of hospital beds per capita	2.74E-03	2.17E-03	0	3.28E-02
Intesp00pc	Hospital internment per capita, 2000	7.25E-02	4.33E-02	0	3.36E-01
Ambuis99pc	Ambulatory units per capita, 1999	3.67E-04	2.93E-04	0	2.49E-03
Postde99pc	Health posts per capita, 1999	1.14E-04	2.04E-04	0	2.41E-03
Cent99pc	Health centers per capita, 1999	7.87E-05	8.60E-05	0	7.34E-04
Cons99pc	Surgical units or private clinics per capita, 1999	5.24E-06	5.21E-05	0	6.90E-04
Conso99pc	Dental surgical units per capita, 1999	2.14E-06	1.61E-05	0	1.94E-04
Ambus299pc	Hospital ambulatory unit per capita, 1999	2.50E-05	3.53E-05	0	2.91E-04
Postca99pc	Centers w only ambulatory care per capita, 1999	3.54E-07	2.92E-06	0	3.66E-05
Dthrate	Death rate: deaths per population	4.96E-03	1.87E-03	0	9.75E-03
Dthrinfec	Rate of deaths by infection	2.69E-04	1.53E-04	0	1.20E-03
Dthr01	Death rate under 1 year olds	2.33E-02	1.40E-02	0	8.84E-02
Dthr14	Death rate for 1-4 year olds	1.03E-03	8.00E-04	0	7.33E-03
Dthr59	Death rate for 5-9 year olds	3.53E-04	2.99E-04	0	4.35E-03
Dthr1014	Death rate for 10-14 year olds	3.96E-04	3.05E-04	0	2.34E-03
Dthr1519	Death rate for 15-19 year olds	9.79E-04	6.73E-04	0	4.12E-03
Dthr2029	Death rate for 20-29 year olds	1.57E-03	8.29E-04	0	4.98E-03
Primenrollr	Approx primary school enroll rate, 2000	0.990	0.115	0.638	2.060
Primestr	Primary school student-teacher ratio, 2000	24.643	4.485	13.361	43.945
Primestusch	No students per primary school, 2000	238.778	160.290	33.857	793.786
Employ98	Approximate share of adults (>15) employed	0.203	0.161	0.003	1.083
Orcacor97pc	Government revenues per capita	2.462	1.637	0	16.500
FunFUN00	Funding for basic education	3.073	9.918	0	50.479
Shgarbbuck00	Public garbage bins (share HHs), 2000	0.061	0.087	0	0.730
Shgarbpubser00	Public garbage collection (share HHs), 2000	0.637	0.291	0	0.982
Shgarbburied00	Garbage buried (share HHs), 2000	0.013	0.020	0	0.139
Shgarbriver00	Garbage disposal: river (share HHs), 2000	0.005	0.012	0	0.162
Shgarbwast00	Garbage disposal: wasteland (share HHs), 2000	0.111	0.140	0.000	0.792
Shhhouse00	House as dwelling (share HHs), 2000	0.909	0.100	0.352	0.998
Shhhapt00	Apartment as dwelling (share HHs), 2000	0.066	0.095	0	0.616
Shhhroom00	Room as dwelling (share HHs), 2000	0.011	0.013	0	0.078
Popdens00	Population density, 2000	11.245	22.243	0.004	129.086
Lrhhpcinc00	Ln real per capita income, July 2000	5.426	0.677	3.848	6.826

*These summary statistics are unadjusted for sample design including weights and clustering.

°For some of the municipality variables expressed in shares of the population, the population figure in the denominator is projected or is from a different year than the main variable in the numerator. These variables, such as Primenrollr and Employ98, are proxies for the true shares and may take on values greater than 1.

Specification Tests and Empirical Issues

A number of specification tests were performed on the HAZ model in (2). The first set of tests compared a unified model for all Brazilian children to separate models nested by the age of the child, then urban and rural residence, and regional location.

For analysis of nutrition determinants, the WHO Expert Committee on Nutrition (WHO, 1995) recommends a minimum stratification of a sample of children under 5 years old into cohorts above and below age 2.[10] A Wald test was performed to compare separate models for 6-23 month olds (N=1240) and 24-59 month olds (N=2370) to a unified model of all 6-59 month olds (N=3610). The unified model was rejected in favor of separate models for each age cohort.[11] Hereafter, we refer to the cohort of 6-23 month olds as "infants" and to the 24-59 month olds as "toddlers".

Next, the data were separated into urban and rural sub-samples to test for the appropriateness of separate urban and rural models within the two age groups. Within the infant cohort, a Wald test was unable to reject the unified infant model for separate models for urban and rural infants. In the toddler cohort, the results were more mixed. The Wald test rejected the unified toddler model in favor of urban and rural toddler models (Wald statistic= 46.84; $\chi^2(30) = 43.77$). However, a Wald test is inexact and may too easily reject the null in small-to-medium samples. The more conservative bounds test in Kobayashi (1986) failed to reject the unified toddler model.

We then tested for regional differences in coefficients. Observations from the South, Center-West, and Southeast regions were pooled into a Center-South (CS) sub-sample and those from the North and Northeast regions were pooled to create a North-Northeast (NNE) sub-sample. For the infant cohort, both the Wald test and the Kobayashi bounds test rejected the unified infant model in favor of separate NNE and CS infant models. For toddlers, the Wald test failed to reject the toddler model in favor of NNE and CS toddler models, but with a p-value of 0.1008. The Kobayashi test failed to reject the unified toddler model.[12]

These results are quite mixed. There is only weak support for urban/rural models in the toddler cohort, and no support for urban/rural models for infants. However, separate NNE/CS models are supported for infants, but not for toddlers. Since the bulk of this evidence seems to favor separate models for infants and toddlers only, with no further disaggregation, we began with a model that uses only this separation of the data. This model will be referred to as Model 1. Later, we will present another set of estimates that allows for division of the sample into eight models based on the infant/toddler, urban/rural, NNE/CS divisions described. Such an approach

[10] This recommended point of separation reflects, among other things, that the reference populations on which the Z-scores are based are different for 6-23 month olds and 24-59 month olds. Therefore, an empirical finding that the determinants of nutritional status are different above and below this age cutoff may be due in part to the use of different reference groups in the NCHS/WHO data.

[11] The Wald statistic was 124.83. The corresponding χ^2 test statistic with 33 degrees of freedom was 47.40. The unified model was easily rejected.

[12] Results of all tests are available from the authors upon request.

can be defended in part by the mixed model specification test results and because the data are representative at this level of disaggregation.[13] This model will be referred to as Model 3.

For Model 1, separate regressions of height-for-age Z-scores were estimated for the infant and toddler sub-samples of the PNDS data using a subset of the regressors listed in Table 2. Both models began with the full set of possible explanatory variables, with the exception that the infant models used dummy variables AGE7M, AGE8M, AGE9M, AGE10M, AGE11M, AGE1223 for child's age and the toddler model used AGEYRS and AGEYSQ. Explanatory variables were selected for inclusion into the models depending on their contribution to predictive accuracy based on several criteria including individual t-statistic and minimization of RMSE.[14] First, a stepwise regression procedure of backward selection was used in which each model was estimated on all potential regressors and then variables with high p-values were removed. Through repeated attempts using different p-value thresholds for variable inclusion, it was determined that a liberal threshold of $p<0.3$ performed best in the stepwise regression in terms of minimizing RMSE. Thus, many variables that would not meet standard levels of significance in a model of nutrition determinants were kept as regressors because they improved the predictive accuracy of the model. Using the list of variables obtained from the stepwise regression as a base, additions and subtractions of individual variables from the list were considered on the basis of their effects on RMSE. One approach involved replacing all child or household level regressors omitted from the models during the backward selection procedure. Returning these variables generally led to little change in RMSE, but in the end these variables were retained on the grounds that they belonged in the model for intuitive reasons. This procedure lead to the use of 39 predictors in the infant model and 33 predictors in the toddler model.

We also tested for the presence of omitted municipality-level effects in Model 1. The presence of significant municipality-level effects in nutritional status that remain unexplained by the model would argue for using simulation methods to add the municipality component of the empirical residuals back into the linear prediction of HAZ status. We tested for unobserved municipality effects in two ways: (i) regressing residuals from the HAZ regression on municipality fixed effects, and (ii) estimating the HAZ model in (2) by random effects. F-tests that the municipality fixed effects were jointly zero cannot be rejected for either the infant or toddler model. There is no evidence of important omitted municipality effects. The large set of municipality variables included in the regression capture the correlation of municipality effects with nutritional status. Since no omitted municipality effects were identified, simulation methods for predicting nutritional status were not employed at this stage.

In the final stage of the prediction model, parameter estimates from each model were matched to data from the 2000 census survey to construct estimated child HAZ levels and corresponding predicted stunting prevalence for each municipality. In the infant and toddler samples of imputed HAZ levels in the census, 2.9 percent and 3.8 percent of children, respectively, had

[13] In the literature on poverty mapping, it is common to estimate separate first-stage models on each stratum in the sample data set, in order to allow for greater variability in predicted poverty rates. Such an approach is not feasible here. The PNDS strata are the urban and rural areas of each UF, many of which have fewer than 20 observations in the data set.

[14] See Theil (1961) on the use of minimized RMSE as a criterion for predictive accuracy.

missing HAZ estimates due to missing data. Also, in the infant cohort from the census, a limited number of predicted HAZ levels (0.07%) fell outside the (-5,3) range used to trim the PNDS sample. In the toddler cohort, even fewer predicted HAZ levels (0.02%) fell outside these bounds. These outliers were dropped in both cohorts. The municipality average prevalence was constructed as a child-population-weighted average of the stunting rates for each age cohort. The imputation exercise created stunting prevalence estimates for the 5507 municipalities in existence in 2000. Between 2000 and 2001, another 54 municipalities were created. These were assigned the median estimated stunting prevalence for their corresponding *Unidade de Federação* (UF).

One concern raised by the approach to variable selection used for Model 1 is that of overfitting the data. Overfitting may occur when right-hand-side variables are selected primarily on the basis of their fit for the present sample. Although all of the variables considered as potential regressors were chosen for intuitive reasons as potential predictors of height achievement, the stepwise selection procedure used to reduce the regressor matrix to only variables with relatively high t-statistics could have fit the PNDS data too closely. Because of sampling error, these variables may not be the best predictors in other samples. This can lead to extreme predictions in other samples, particularly if the distribution of the X variables differs across the two data sets.

Evidence of overfitting is provided if estimating the regression on a smaller sub-sample leads to a large change in R^2. To undertake this test, we drew random samples of half the observations from the infant and toddler cohorts and estimated the regressions again for the same set of X variables. With the decline in sample size, the R^2 changed from 0.202 to 0.218 in the infant model and from 0.264 to 0.274 in the toddler model. These results provide little evidence of overfitting in the infant and toddler models in Model 1.

Despite this informal evidence against overfitting, we developed another set of infant and toddler models of height achievement in the PNDS that relied on a narrower set of predictors. This approach would provide estimates of stunting prevalence that were more conservative with respect to extreme values, and would also serve as a robustness test for Model 1. We refer to this new set of models as Model 2 or the "reduced" model. In these models, no stepwise selection of variables was used. Instead, we reduced the set of child- and household-level variables under consideration by either dropping or merging variables. Variables dropped from Model 2 include female headship, number of rooms and bedrooms, piped water, radio, vcr, washing machine, and quadratic terms for child and mother's ages. The categorical variables for mother's education were also collapsed at the upper tail of the distribution. Starting with this reduced set of child and household regressors, all remaining child and household variables were included in the regressions for Model 2. To identify a more restricted set of municipality-level variables to include, we first estimated the model using only the child and household variables for each of the infant and toddler cohorts. We regressed the residuals from these models on municipality fixed effects in order to capture omitted municipality effects for each cohort. We then regressed these unexplained fixed effects on the set of potential municipality-level variables and selected the 3-5 municipality variables that best explained the municipality effects. These variables were included in the set of regressors and the models were re-estimated. This procedure is very similar to that used by Elbers et al (2003) to select location-specific regressors. These changes in variable selection led to inclusion of 24 regressors in both the reduced infant model and the reduced toddler model.

In Model 3, which used eight regressions based on infant/toddler, urban/rural, and NNE/CS divisions of the sample, the approach to variable selection was similar to that used in Model 2. This procedure, which uses fewer regressors and is more conservative with respect to concerns for overfitting, was chosen because the division of the sample into eight sub-samples led to a large reduction in degrees of freedom in each regression. We now present the results of these models and some tests of model performance.

Results and Predictive Performance of the Evaluation Models of Stunting Prevalence

Estimates from the infant and toddler HAZ regressions from Model 1 are presented in Appendix Tables B.1 and B.2. As noted above, the variable selection process kept 39 and 33 predictors in the infant and toddler models, out of 71 and 67 possible regressors, respectively. The coefficient of determination (R^2) for the infant model was 0.202 and for the toddler model was 0.264. A large share of the variation in HAZ remains unexplained by these models, despite having a very large set of regressors available. In similar poverty mapping exercises, the analogous first-stage regression is a model of consumption expenditure. These typically have an R^2 of 0.4 to 0.6. This provides the first evidence that predictive models of nutritional status are rather incomplete. It suggests that caution should be exercised when using these models for allocating program budgets.

Despite the shortcomings of these models, they appear to perform better than the original models that generated the malnutrition estimates used by BA. Using the 1996 PNDS data, we attempted to construct variables for mother's illiteracy, household sanitation, and household head income that were similar to the variables used in the original models in order to assess the performance of those models. Mother's illiteracy is provided in the PNDS, but sanitation variables are coded differently than in the 1989 PNSN, which somewhat limits the comparability of the sanitation variables. However, the greatest difficulty came in developing the income variables, since household head income is not directly measured in the PNDS. As a result, a crude proxy for household head income was developed based on asset ownership following an approach developed by the Brazilian research organization ANEP, referred to as "Critério Brasil."[15] There is considerable error in income measurement through this approach, although the scope of the problem is limited by the need only to classify households as above or below one or two income thresholds in each model. Using these variables constructed from the PNDS data, we estimated the original logit models for stunting from (1). The parameter estimates obtained differed considerably from those in the original 1997 study. This difference partially reflects the seven-year time span between the two data sets, though it also is determined by differences in variable definitions. Despite the shortcomings of our attempts to replicate the original models for the newer 1996 PNDS data, results from this exercise probably provide a reasonable picture of the explanatory power of a similar model with a small set of binary regressors. The outcome is that logit estimates of stunting prevalence for the original models on PNDS data (7 regional models in all) yielded a median pseudo-R^2 of 0.051 and no pseudo-R^2s larger than 0.092. For the sake of

[15] See Associação Nacional de Empresas de Pesquisa. January 2003. "Critério de Classificação Econômica Brasil Atualizado" at http://www.anep.org.br/m-arquivo.htm.

comparison, we estimated the infant and toddler models from Model 1 as logit model of stunting prevalence and obtained a pseudo-R^2 of 0.18 and 0.17, respectively. The evaluation model appears to have greater predictive ability than the model used to allocate the budget for *Bolsa Alimentação*.

Another indicator of the performance of these types of models for geographic targeting is their accuracy in predicting average malnutrition prevalence at higher levels of regional aggregation. A summary of stunting prevalence estimates in Table 3 provides further evidence on the performance of the evaluation model in Model 1.[16] The first column of means in Table 3 presents the observed stunting prevalence from the PNDS for the eight strata considered in specification tests above. Stunting is most severe in the Northeast, particularly in rural areas, and is generally worse for infants than for toddlers. The next column of means in Table 3 presents the average predicted stunting prevalence for each stratum based on our first-stage model of height-for-age Z-scores in the PNDS for Model 1. The evaluation models perform quite well at this stage. In each stratum, average predicted stunting rates are very close to the sample estimates. The lower panel of the same column contains average estimated stunting prevalence for the 8 strata based on the predicted probability of stunting for children in the second stage of Model 1 using the 2000 Census data. Estimated cohort stunting prevalence for both the infant and toddler models is very close to observed prevalence, with the exception of infants in the rural Center-South, which contribute relatively few observations to the first-stage parameter estimates. These estimates are reassuring, given that there are a number of factors that can contribute to a decline in accuracy in the second-stage models. These include the comparability of variable definitions between the PNDS and the 2000 census, sampling error in the PNDS, model error in the first-stage estimates, and idiosyncratic error captured in the household residuals of the first-stage model. Of course, the goal of this methodology is to correctly predict differences in malnutrition prevalence across municipalities, not average cohort prevalence. However, had these average prevalence estimates been very inaccurate, we would have had little confidence in the individual municipality estimates as well.

For Model 2 the R^2 for the infant model falls to 0.177 and that of the toddler model is 0.255. Table 3 shows that the reduced set of predictors in Model 2 perform comparably to Model 1 in estimating regional stunting prevalence in the first-stage model. Only the estimate for rural toddlers in the CS region is appreciably worse. In the second-stage estimates for the 2000 census, Model 2 is nearly identical to Model 1 for the toddler strata. For the infant strata, results are mixed, with Model 1 out-performing Model 2 for two of the four strata. If closer inspection of municipality stunting prevalence estimates for Model 1 reveals extreme estimates, Model 2 may be preferred, since it does about as well at estimating mean regional prevalence. These results also demonstrate that the evaluation model of stunting prevalence is reasonably robust to the choice of predictors. More evidence on this matter will be presented below.

For the eight strata regressions in Model 3, variable selection was based on the reduced set of candidate variables and included only 3-6 municipality variables. These limitations were driven

[16] The estimates of malnutrition rates available to us from BA do not permit calculation of stunting prevalence because of the adjustments made to the original stunting estimates, and our own estimates of the original stunting model are too different to rely on. Therefore, we do not include the "adjusted" malnutrition prevalence figures used by the MOH in this comparison of stunting prevalence.

by the need to avoid overfitting in sub-samples with relatively few degrees of freedom for estimation. For most of these eight models, the R^2 remains relatively stable compared to the R^2 for the reduced infant/toddler regressions in Model 2. However, the jump in R^2 to 0.335 for the CS rural toddler cohort suggests some overfitting for that model. The last two columns of Table 3 show the Model 3 estimates of stunting prevalence for the eight strata for the first-stage estimates from the PNDS and the second-stage estimates for the 2000 Census. The first-stage infant models in Model 3 perform better than Models 1 and 2, but the first-stage estimates for toddler are not better. From the second-stage estimates on the eight strata, half of the estimates are closer to observed stunting prevalence in Model 3 than in Model 1.

Table 3: Estimates of Regional Stunting Prevalence from the Evaluation Models

Stunting Prevalence			**Observed**		**Model 1**		**Model 2**		**Model 3**	
			1996 PNDS		1st-stage Prediction, 1996 PNDS		1st-stage Prediction, 1996 PNDS		1st-stage Prediction, 1996 PNDS	
		Obs	Mean	Std Err	Mean	Std Err	Mean	Std Err	Mean	Std Err
Infant	NNE Urban	484	0.172	0.018	0.157	0.007	0.156	0.006	0.174	0.008
	CS Urban	450	0.072	0.013	0.097	0.004	0.096	0.004	0.075	0.003
	NNE Rural	221	0.309	0.033	0.267	0.012	0.264	0.010	0.297	0.012
	CS Rural	85	0.105	0.037	0.147	0.011	0.140	0.011	0.135	0.014
Toddler	NNE Urban	883	0.139	0.012	0.138	0.007	0.136	0.006	0.147	0.007
	CS Urban	879	0.039	0.007	0.056	0.003	0.057	0.003	0.047	0.002
	NNE Rural	434	0.245	0.022	0.255	0.012	0.253	0.012	0.275	0.013
	CS Rural	174	0.107	0.026	0.103	0.010	0.099	0.009	0.114	0.012
					2nd-stage Prediction, 2000 Census		2nd-stage Prediction, 2000 Census		2nd-stage Prediction, 2000 Census	
					Mean		Mean		Mean	
Infant	NNE Urban	148,555			0.160		0.143		0.150	
	CS Urban	277,977			0.096		0.088		0.071	
	NNE Rural	98,527			0.292		0.268		0.328	
	CS Rural	59,963			0.168		0.159		0.146	
Toddler	NNE Urban	348,078			0.122		0.121		0.123	
	CS Urban	573,513			0.051		0.051		0.042	
	NNE Rural	256,053			0.245		0.244		0.260	
	CS Rural	129,308			0.081		0.081		0.098	

4 Geographic Targeting Performance of *Bolsa Alimentação*

This section evaluates the accuracy of geographic targeting in *Bolsa Alimentação.* The implications of targeting failures for the program budget and for income redistribution are also considered. The analysis is based on a comparison of the municipality stunting prevalence estimates developed for this report to the estimated malnutrition rates used to allocate the program budget. Although we have demonstrated that our estimates of stunting prevalence are subject to non-negligible error, the three evaluation models presented here appear to have greater predictive ability than the approach used by BA. Thus, we will interpret differences in targeting outcomes between the evaluation model and the original estimates used by BA not as precise estimates of the size of targeting failures, but as suggestive of the spatial and socioeconomic distribution of these failures and as a rough estimate of their size. Throughout this comparison, we assume that household level targeting in the second stage of the targeting procedure is perfect in the sense that, once the number of stunted children in a municipality is determined through geographic targeting, identifying those children through household targeting is effortless.[17] Estimates in this section of the number of children to be included in the program from the evaluation models were constructed by multiplying the municipality estimated malnutrition prevalence times the number of children aged 6 months through 6 years in the 2000 census.

Comparing Distributions of Malnutrition Estimates

Table 4 presents summary statistics of estimated municipality malnutrition prevalence for BA and for the three evaluation models. Estimates of mean malnutrition prevalence for BA and the evaluation models are not directly comparable because the evaluation models measure stunting prevalence, whereas stunting estimates in the BA model were adjusted upward according to the regional average of the ratio of WAP10 to HAZM2. A comparison of the evaluation models shows that Model 1 performs best at estimating national stunting prevalence. Models 2 and 3 both underestimate stunting levels.

The other summary statistics in Table 4 demonstrate a primary shortcoming of the simple BA model raised in Section 2: reliance on a small set of predictors constrains the model's ability to differentiate between municipality malnutrition rates. The BA malnutrition prevalence figures have lower variance than the evaluation estimates and fall in a narrow range, from 4.0 to 29.9 percent. The evaluation models, on the other hand, have greater variance in estimated malnutrition prevalence and a much wider range.

Finally, Table 4 shows that the distributions of stunting prevalence estimates across the three evaluation models are quite similar. There are only minor differences in median prevalence and interquartile range, for example. This suggests that the evaluation models are fairly robust to model specification, including reductions in the number of predictors and differences in the

[17] We use the assumption of perfect household targeting within municipalities in order to focus exclusively on geographic targeting in this report. Of course, significant difficulties arise in designing effective household targeting as well. Preliminary field visits conducted by IFPRI show, for example, that there was considerable confusion within municipalities about responsibility for household targeting, which hampered the creation of the list of beneficiaries. We briefly discuss the consistency of geographic and household targeting objectives later in this report.

degree of regional aggregation for first-stage model estimation. These results are also consistent with the weak evidence for separate urban/rural and regional models provided by the Wald tests in Section 3. Notice also that stunting prevalence estimates in the tails of the distribution from Model 1 are not appreciably more extreme than those in Models 2 and 3. This provides further evidence that concerns about overfitting in Model 1, which can lead to extreme predictions, are not serious. Based on these results, we rely on Model 1 as our default set of estimates of stunting prevalence for the comparison with the original estimates undertaken in this section, in part because Model 1 provides the closest estimate of observed stunting prevalence of any of the evaluation models.

The last row of Table 4 presents Pearson correlation coefficients of estimates of municipality stunting prevalence from the three evaluation models with the adjusted stunting prevalence estimates from the original model used by BA. The correlation coefficient of 0.65 for the original model estimates with the preferred estimates from Model 1 suggests that the program achieved significant improvements in targeting over an untargeted program by using the original model, but that further improvements are available.

Table 4: Estimates of Malnutrition Prevalence:
***Bolsa Alimentação* and the Evaluation Model**

Observed Stunting Prevalence, 1996 PNDS: 11.3%				
	Bolsa Alimentação	**Evaluation**		
		Model 1	**Model 2**	**Model 3**
Mean	14.9	11.1	10.7	10.8
Std Dev	7.1	8.9	7.9	8.9
Min	4.0	1.2	0.0	0.0
Max	29.9	81.5	65.5	72.6
Percentiles				
1%	5.3	3.0	3.1	2.7
5%	6.1	4.1	4.3	3.8
25%	8.8	6.4	6.3	5.9
50%	12.3	10.6	10.3	10.0
75%	22.6	18.0	17.2	18.7
95%	26.2	29.5	27.1	29.6
99%	27.5	42.7	37.8	40.0
Pearson correlation with BA estimates		0.650	0.645	0.731

The HAZ Threshold and Stunting Estimates Implied by BA Program Size

As noted earlier, the upward adjustment of stunting prevalence estimates from the original model to bring them in line with average prevalence of WAP10 suggests that the government wanted to reach more children through the program than just those who were already stunted. A program expansion of this size would allow the government to reach many children who are not stunted by the "classical" definition, but whose height achievement is less than their potential or who may be at high risk of stunting.

Maintaining height-for-age as the preferred measure of nutritional status for geographic targeting, the government's desire to expand the program beyond stunted children suggests an implied HAZ threshold for program participation that is above -2. In order to calculate this implied threshold, we first fixed the number of children in the program at the level determined by the BA prevalence estimates. We then re-estimated modified "stunting" prevalence rates in Model 1 for infants and toddlers for various threshold levels of HAZ and calculated the corresponding numbers of children predicted below each threshold level. This procedure was repeated until an HAZ threshold was identified for which the predicted number of stunted children was equal to the number currently in the program. This procedure identified an implied threshold for inclusion in the program of HAZ<-1.823. By construction, this HAZ threshold creates a program of equal size to the current BA program, with 2.77 million child beneficiaries, or 12.4 percent of children age 6 months to 6 years 11 months. However, as with the Model 1 estimates of stunting prevalence based on the HAZ<-2 threshold, the evaluation model underestimates the mean of this "mild" stunting prevalence for HAZ<-1.823, which is 14.2 percent of children aged 6-59 months in the PNDS sample. We refer to stunting prevalence at this threshold consistent with BA program size as the "implied" stunting prevalence.

In a population such as Brazil's characterized by intermediate rates of stunting, it is important to adjust these prevalence estimates for baseline or expected prevalence. As noted in WHO (1995), when -2 Z-scores is used as a stunting threshold, 2.3 percent of the NCHS/WHO reference population will be classified as stunted even if their growth is not impaired. The correct measure of classic stunting prevalence for Brazil is obtained by subtracting this baseline prevalence from observed prevalence of children with HAZ below -2. Using the PNDS estimate of 11.3 percent of children age 6-59 months with HAZ < -2, this yields a stunting prevalence of 9.0 percent. The evaluation model estimates this figure at 8.8 percent of children. At a threshold of HAZ<-1.823, 3.4 percent of the reference population would be classified as malnourished. Subtracting this baseline prevalence yields a PNDS estimate for this "mild" incidence of malnutrition of 10.8 percent and an evaluation model estimate of 9.0 percent.

Another indication of the accuracy of the estimates of malnutrition prevalence from the original model and evaluation Model 1 is available from new results of a school height census of children for Pelotas municipality in the state of Rio Grande do Sul.[18] As noted earlier, Morris and Flores (2002) show that a school height census of first graders in municipalities in Honduras provided reliable estimates of the height achievement of children under five in the same municipalities. They indicated that these results should provide a similar ranking of municipalities to a ranking based on stunting prevalence. For Pelotas, the figures available are for stunting rather than mean

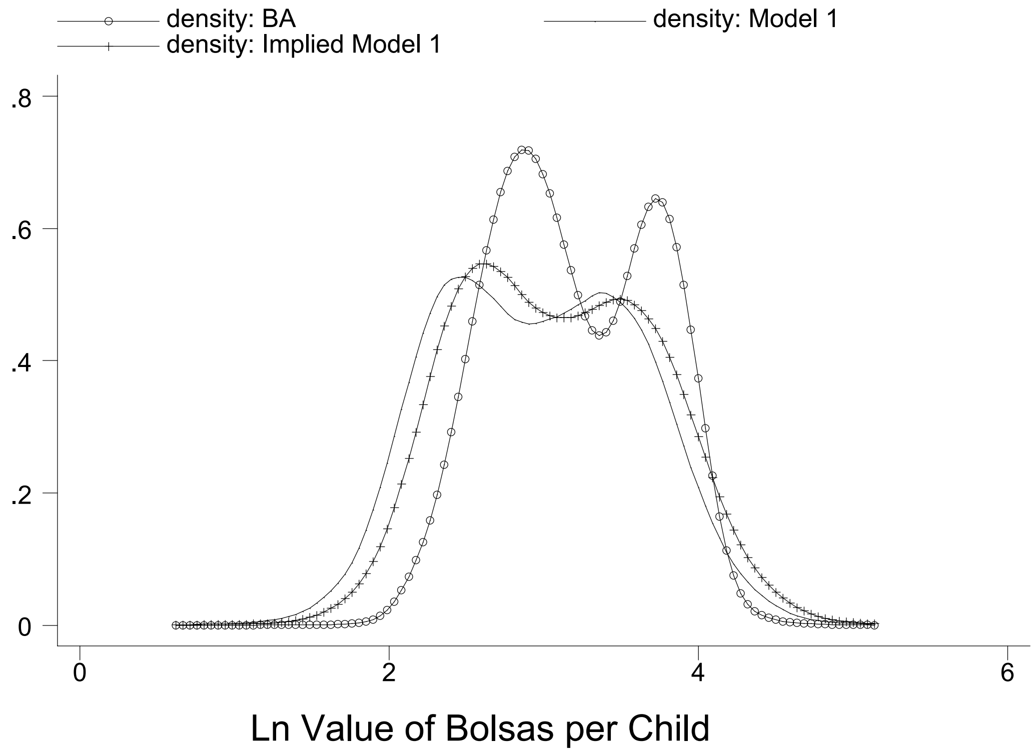
[18] We would like to thank Cora Luiza Araujo for sharing the results of the Pelotas school height census with us.

HAZ. The school height census showed stunting prevalence of 4.1% among first graders and 3.2% for school children aged 7-10. The stunting prevalence estimate for children aged 6 months to 5 years from evaluation Model 1 for Pelotas is 5.6%, and the estimate for toddlers (aged 2-5 years) is 4.3%. Although we are only able to make this comparison for one municipality, these results provide encouraging support for evaluation Model 1. The stunting estimate for toddlers is very close to the observed stunting prevalence in first graders. The adjusted stunting estimate from the model used by BA for Pelotas is 7.4%. We should expect this figure to be somewhat higher because the BA program has a larger target population than children who are classically stunted. However, using results from evaluation Model 1 adjusted upward for the BA-implied threshold of HAZ<-1.823, the implied stunting prevalence from Model 1 for Pelotas for children under 5 is 6.2% and the estimate for toddlers is 4.5%. Again, the evaluation model appears to provide more accurate estimates of stunting prevalence than the model used to allocate the program budget.

Having revised the estimates of municipality malnutrition prevalence to be consistent with the size of the BA program, we investigate the distribution of program resources in more detail. Figure 1 presents nonparametric estimates of the distribution of the BA program budget to municipalities based on three measures of malnutrition prevalence: the original estimates used by BA, the stunting estimates from evaluation Model 1, and the evaluation estimates for implied stunting prevalence from Model 1. The graph shows the distribution of the natural logarithm of the value of bolsas per child population given to municipalities for each measure.[19] This graph clearly demonstrates the bunching of the malnutrition prevalence estimates in the original model. The evaluation model of stunting, on the other hand, has many more municipalities receiving either a low or high budget per child, with a lower mean budget size. In the evaluation estimates revised for program size, this distribution shifts upward. Regarding the "Implied Model 1" distribution as the preferred distribution at this program size, the space below the peaks of the BA distribution and above the "Implied Model 1" distribution represents failures of geographic targeting including both leakage of program benefits to ineligible beneficiaries and undercoverage of eligible beneficiaries. That is, some of the municipalities represented there should have a smaller budget (leakage) and others should have a larger budget (undercoverage).

[19] This graph is comparable to a graph of the natural log of estimated malnutrition prevalence, since prevalence is simply the number of bolsas divided by the child population and the value of the transfer is uniform at R$15 per child per month.

Figure 1: Municipality Distribution of Stunting Prevalence



Malnutrition Maps for Brazil

The effect of the original model's difficulties in differentiating between low-, middle- and high-malnutrition-prevalence municipalities on the geographic distribution of malnutrition rates and program resources can be seen most clearly by mapping malnutrition prevalence for the various estimates being considered. Figures C.1-C.3 in Appendix C map the estimated malnutrition prevalence from the original "BA" model, Model 1 from the evaluation, and evaluation Model 1 revised for the BA-implied stunting threshold, respectively, for the 5507 municipalities in 2000. In each figure, estimates of malnutrition prevalence were grouped into seven categories of prevalence: six categories of five percentage points (from 0.00-4.99% up to 25.00-29.99%) and a top category of 30% and above. The comparison of Figures C.1-C.3 is striking. The BA program estimates in Figure C.1 have almost no municipalities in the lowest and highest categories of malnutrition rates. The North and Northeast regions clearly suffer from the worst malnutrition, but there is limited differentiation of estimates within regions. The stunting rates from the evaluation model presented in Figure C.2 stand in stark contrast to the results in Figure C.1. The stunting model shows considerable variability in malnutrition rates within regions as well as within states (demarcated by solid lines). Although the North and Northeast are obviously the worst affected, there are pockets of relatively low prevalence within these areas, as would be expected. Figure C.3 shows that, when the threshold stunting level is relaxed in the evaluation model to be consistent with the size of the BA program, many municipalities jump one or two categories of stunting prevalence for this relatively mild definition of stunting. It appears that more resources are being allocated to the interior areas of the North, the Northeast, and the Center-South of Brazil as a result of expanding the program beyond classically stunted children.

Maps of estimated stunting prevalence for Model 2 and Model 3 are provided for comparison in Figures C.4 and C.5, respectively. The map of stunting prevalence estimates for Model 2, which are based on a smaller set of predictors, show less variability in stunting as expected. Estimates from the eight age-regional models from Model 3 exhibit a more uniform distribution of stunting prevalence in the Center-South region and stronger regional divisions in prevalence than estimates from Model 1.

A striking feature of the maps from the evaluation models is the high estimated prevalence in the North region. The mean of the municipality stunting prevalence estimates for the North from Model 1 is 25.4 percent, which lies in the second highest prevalence category for these maps. More than 25 percent of municipalities in the North lie in the highest prevalence category (>30%) for estimates from Model 1. Although we have no means to validate these estimates, we are concerned that they are too high. The PNDS data from which the parameters of the stunting model are derived does not contain any observations on households in rural areas in the North region. If these households have very different distributions of X variables in the census than their distributions for other regions of Brazil in the PNDS, this could result in inaccurate stunting estimates. This high concentration of estimated stunting prevalence in the North region falls somewhat in Model 3 (Figure C.5), where prevalence estimates for the rural North are based on first-stage stunting models for households in the urban North and urban and rural areas of the Northeast. However, these stunting estimates are still quite high. Based on these concerns, we recommend that these estimates for the North region be used cautiously.

Figure C.3 also demonstrates that considerable geographic variation in the allocation of program resources would still be possible had the government grouped municipalities into categories of malnutrition prevalence and given each municipality a budget of bolsas per child equal to the average malnutrition prevalence for its category. Given the error in estimated stunting prevalence that was observed even with the very large set of potential predictors used for the evaluation model, the use of this type of classification might be justified. There is little reason to believe that prediction models using aggregate census data can reliably differentiate between fractions of percents in malnutrition prevalence in the manner applied by the BA administration.

Quantifying Targeting Errors

A comparison of the allocation of the BA program budget based on the original model used by BA and the recommended allocation suggested by stunting estimates in Model 1 reveals considerable errors in geographic targeting. Using the municipality implied stunting estimates from evaluation Model 1 as a benchmark, 17.87 *centavos* out of every *real* spent on the BA program did not reach the intended municipality—a targeting accuracy rate of 82.13 percent. Equivalently, the government spent R$1.22 for every *real* transferred to intended beneficiaries. If geographic targeting was the only source of targeting errors, this targeting accuracy rate would represent a moderate level of leakage compared to similar programs (see Coady et al, 2002). However, these estimates of leakage represent a lower bound on targeting errors in the program because they only capture errors that occurred in geographic targeting. Errors in household targeting within municipalities can only result in further reductions in targeting accuracy. Nonetheless, geographic targeting errors alone were large enough to erase the surplus represented by the fact that the program targeted 2.77 million children when the number of stunted children in 2000 was 2.48 million according to the evaluation model. In fact, at these levels of leakage in geographic targeting, the BA budget would have to be increased from R$499mn to R$550 million in order to reach all stunted children. In fact, this scenario is actually too optimistic, since it assumes perfect household targeting.

Errors in household targeting may have been considerable in part because children at risk of malnutrition were identified only through the proxy of per capita household income below R$90 per month. However, if the target population for the program is stunted children, or those vulnerable to stunting, poverty is an imprecise selection criterion. The correlation between income poverty and child stunting may be high, but probably does not exceed 0.7 or 0.8. If the program was fully funded to cover all poor households with children, it would only reach the segment of the target population that were both stunted and poor—80 percent of stunted children would be a generous estimate. However, the program was only funded to cover stunted children, who number less than poor children for the poverty line chosen. This would cause many municipalities to run out of funds before reaching all poor households with children, and the share of stunted children reached would fall even further. For this reason, municipalities were given the opportunity to amend the roster of beneficiaries to improve targeting of children at risk of malnutrition. However, limited field visits suggest that confusion kept many municipalities from taking advantage of this opportunity. The resulting undercoverage of children at risk of stunting is impossible to quantify, but is potentially large.

These estimates of targeting errors are not as large if the target population is expanded to include all children in poor households. However, serious household targeting errors could still have occurred because income is self-reported for households entered in the *cadastro unico*, and these households have an incentive to under-report their earnings.

State-Level Malnutrition Estimates and the Distribution of Targeting Errors

As a method of summarizing the geographic distribution of errors in targeting, we consider malnutrition estimates and budgetary allocations at the state level. Table 5 presents estimates of malnutrition prevalence from the original model and the evaluation and the corresponding size of the budgetary misallocation for Brazil's 27 *Unidades de Federação*. Columns 1-3 show the estimated malnutrition rates for the BA program, stunting evaluation Model 1, and evaluation Model 1 for "implied" stunting, respectively. A comparison of Columns 1 and 2 shows the lower average malnutrition prevalence that results from the stunting model, but also uncovers substantial differences in malnutrition estimates across states for the two estimates. The figures in Column 3 reflect identical national malnutrition prevalence as Column 1, still with substantial differences in state-level estimates.

Columns 4 and 5 of Table 5 show the size of the budget for each state under the original BA and evaluation approaches for an identical total budget. These numbers indicate the state-level cost of budgetary misallocations under the program, under the assumption that the evaluation model estimates are preferred to the estimates used by BA. The state of Maranhao, which had among the very worst malnutrition rates under all three approaches in Columns 1-3 and is the poorest state, had the third largest *real*-value shortfall in its BA budget at R$8.7 million, according to the evaluation methodology. The states with the two greatest shortfalls in budget in terms of value are Amazonas and São Paulo, respectively. A useful way to quantify the budgetary value of targeting errors is to express the BA budget for each state as a fraction of the recommended budget under the evaluation. This performance indicator is presented in Column 6. In those states for which this indicator is less than 1, there was undercoverage of the target population of children according to the evaluation methodology. Where the indicator is greater than 1, the BA budget included leakage to non-target households. Of course, the total value of leakage and undercoverage for all states must be equal. The table shows that many states had substantial undercoverage: four states (Amazonas, Roraima, Amapa and Acre) received less than half of the budget they would have received under the evaluation methodology. One of the worst affected was the state of Amazonas, which had the second worst classic child stunting prevalence estimate at 25.2 percent, but received only 41.7 percent of the recommended budget, resulting in a shortfall of R$14.9 million.

The states suffering undercoverage show no clear pattern in terms of income distribution. Column 7 ranks the 27 states by per capita income calculated from reported monthly income of household heads from the 2000 census, adjusted for average household size and regional prices. States with the greatest undercoverage were ranked 16, 12, 17, and 14, respectively, by per capita income. The wealthiest UF, Distrito Federal, received only 66.2% of the budget recommended by the evaluation model. Also, the two states that benefited from the greatest leakage were ranked 24th and 15th in terms of per capita income.

Table 5: Malnutrition Estimates and Budgetary Allocations by State

STATE	Malnutrition Prevalence			Budget, R$mn		Targeting Accuracy	Income
	BA	Model 1	Implied	BA	Implied	BA/Impl.	Rank
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ACRE	14.4	26.1	29.0	2.6	5.2	0.498	17
ALAGOAS	21.1	16.8	18.6	16.7	14.7	1.137	25
AMAPA	11.3	20.4	22.6	1.8	3.6	0.498	14
AMAZONAS	11.8	25.2	28.2	10.7	25.6	0.417	16
BAHIA	21.4	16.7	18.5	67.2	58.1	1.156	21
CEARA	21.9	14.8	16.4	42.0	31.5	1.334	23
DISTRITO FED	6.4	8.3	9.6	3.0	4.5	0.662	1
ESP SANTO	9.8	7.3	8.2	6.6	5.5	1.193	9
GOIAIS	10.8	9.2	10.5	12.4	12.0	1.027	10
MARANHAO	22.4	24.7	27.7	36.2	44.9	0.806	27
MATO GR	11.3	11.3	12.5	7.0	7.8	0.905	7
MATO GR S.	10.1	8.4	9.3	5.0	4.6	1.089	8
MIN GERAIS	10.3	8.6	9.7	39.5	37.3	1.059	11
PARA	22.8	22.3	24.9	41.8	45.7	0.916	20
PARAIBA	20.6	12.0	13.2	16.6	10.7	1.556	24
PARANA	9.0	7.2	8.1	19.0	17.1	1.114	6
PERNAMB	19.2	12.3	13.6	35.9	25.5	1.407	15
PIAUI	22.1	19.2	21.3	15.8	15.2	1.037	26
RIO DE JAN	5.0	4.7	5.4	21.5	23.3	0.923	3
RIO GR N.	18.6	10.9	11.9	12.4	8.0	1.556	18
RIO GR S.	7.9	5.9	6.5	16.2	13.4	1.211	4
RONDONIA	13.1	12.3	13.7	4.9	5.1	0.957	13
RORAIMA	9.7	19.9	22.0	1.0	2.3	0.443	12
SAO PAULO	5.7	6.4	7.4	43.0	55.4	0.776	2
SERGIPE	17.6	13.6	15.1	8.2	7.0	1.163	22
ST CATARINA	7.5	7.0	8.0	8.5	9.1	0.939	5
TOCANTINS	11.9	18.2	20.2	3.8	6.4	0.591	19
Total					499.3		

5 Conclusions and Recommendations

The results of this evaluation show that, by using the original 1997 estimates of municipality malnutrition prevalence, the government of Brazil was able to target BA program resources toward localities with higher malnutrition rates. The correlation coefficient of *Bolsa Alimentação*'s estimates of malnutrition prevalence for Brazil's 5561 municipalities and the preferred estimates of stunting prevalence generated by the evaluation methodology was 0.65, suggesting sizeable gains from targeting.

However, data limitations and the simplicity of the original prediction model led to inaccuracies in estimating malnutrition prevalence. The main shortcomings of the MOH approach to geographic targeting include (i) use of outdated household sample and census data from 1989 and 1991, respectively; (ii) use of only three to five binary regressors to predict stunting probabilities in the sample, and (iii) failure to use newly available "mapping" techniques based on household record census data to assist in quantifying targeting errors.

The original model is quite accurate at predicting mean stunting prevalence at high levels of regional aggregation. However, it does not perform as well at identifying differences in malnutrition prevalence between municipalities. As a result, the malnutrition prevalence estimates on which the Ministry of Health based the geographic allocation of bolsas were bunched around the mean. The evaluation model suggests that there were many municipalities with either higher or lower stunting prevalence than predicted in the BA model. These probable geographic targeting errors led to undercoverage of intended beneficiaries in some municipalities. A targeting accuracy analysis showed that, based on comparison with the evaluation model estimates, 17.87 *centavos* out of every *real* spent on the BA program did not reach the intended municipality. Equivalently, the government spent R$1.22 for every *real* transferred to intended beneficiaries. These errors in geographic targeting represent a lower bound on total targeting errors because they do not account for errors in household targeting.

The "malnutrition mapping" approach used in the evaluation models shows some promise as a technique for geographic targeting, but warrants caution as well. The models developed here were quite robust to changes in model specification including choice of predictors and level of regional aggregation for model estimation. However, the fairly large models estimated here could explain only 20-25% of the variation in observed child height achievement. This implies that important factors that determine local stunting prevalence rates have not been captured in these models, which could lead to meaningful errors in prevalence estimates and geographic targeting.

This cautionary note suggests that allocating bolsas to municipalities in direct proportion to estimates of malnutrition prevalence leads to targeting errors. For existing models, we do not know the size of the errors in the prevalence estimates, but we suspect that differences in estimates between municipalities equal to fractions of a percent are probably not meaningful. This suggests grouping municipalities by malnutrition prevalence levels, so as to average out the targeting errors. This procedure would still require arbitrary cutoffs between categories, but these would be fewer than the reversals in rankings that occur when prevalence estimates with three significant digits are treated as precise. Also, availability of standard errors for the estimates of stunting prevalence generated in these models would help to determine how wide to make the groupings of municipalities by prevalence. Subsequent drafts of this paper will include estimates of the standard errors of the stunting prevalence estimates.

References

Baker, Judy and Margaret Grosh. 1994. “Measuring the Effects of Geographic Targeting on Poverty Reduction.” *World Development*, 22(7):983-995.

Benício, Maria Helena D’Aquino and Carlos Monteiro. 1997. “Desnutrição Infantil nos Municípios Brasileiros”. Mimeo. University of São Paulo.

Coady, David, Margaret Grosh, and John Hoddinott. 2002. “The Targeting of Transfers in Developing Countries: Review of Experiences and Lessons.” Mimeo. The World Bank.

Elbers, Chris, Jean Olson Lanjouw, and Peter Lanjouw. 2003. “Micro-level Estimation of Poverty and Inequality.” *Econometrica* 71, no. 1:355-364.

Elbers, Chris, Jean Olson Lanjouw, Peter Lanjouw, and Philippe George Leite. 2001. “Poverty and Inequality in Brazil: New Estimates from Combined PPV-PNAD Data.” Mimeo. The World Bank.

Fujii, Tomoki. 2002. “Micro-level Estimation of Prevalence of Stunting and Underweight in Cambodia.” Mimeo. University of California, Berkeley.

Hentschel, Jesko, Jean Olson Lanjouw, Peter Lanjouw, and Javier Poggi. 1998. “Combining Census and Survey Data to Study Spatial Dimensions of Poverty.” World Bank Policy Research Working Paper, No 1928.

Kobayashi, Masahito. 1986. A Bounds Test of Equality Between Sets of Coefficients in Two Linear Regressions When Disturbance Variances are Unequal. *Journal of the American Statistical Association* 81, no. 394:510-513.

Minot, Nicholas. 2000. “Generating Disaggregated Poverty Maps: An Application to Vietnam.” *World Development* 28(2): 319-331.

Minot, Nicholas and Bob Baulch. 2002. “Poverty Mapping with Aggregate Census Data: What is the Loss in Precision?” MSSD Discussion Paper No. 49. International Food Policy Research Institute.

Mora, JO. 1989. “A New Method of Estimating a Standardized Prevalence of Child Malnutrition from Anthropometric Indicators.” *Bulletin of the World Health Organization* 67: 133-142.

Monteiro, Carlos. 1991. “Counting the Stunted Children in a Population: a Criticism of Old and New Approaches and a Conciliatory Proposal.” *Bulletin of the World Health Organization* 69: 761-766.

Morris, Saul S., Rafael Flores and Maricela Zúniga. 2000. "Geographic Targeting of Nutrition Programs Can Substantially Affect the Severity of Stunting in Honduras." *Journal of Nutrition* 130: 2514-2519.

Morris, Saul S. and Rafael Flores. 2002. "School Height Censuses Are Reliable and Valid Tools for Small-Area Targeting of Nutrition Interventions in Honduras." *Journal of Nutrition* 132: 1188-1193.

Purcell, N.J. and L. Kish. 1980. "Postcensla Estiamtes for Lacal Areas (or Domains)." *International Statistical Review* 48: 3-18.

Ravallion, Martin and Quentin Wodon. 1999. "Poor Areas or Only Poor People?" *Journal of Regional Science* 39(4): 689-711.

Simler, Kenneth and Virgulino Nhate. 2002. "Poverty, Inequality, and Geographic Targeting: Evidence from Small-area Estimates for Mozambique." Mimeo. International Food Policy Research Institute.

Skoufias, Emmanuel, Benjamin Davis and Sergio de la Vega. 2001. "Targeting the Poor in Mexico: An Evaluation of The Selection of Households for PROGRESA" FNCD Discussion Paper No. 103.

Theil, Henri. 1961. *Economic Forecasts and Policy*. (Amsterdam: North Holland).

World Health Organization. 1995. *Physical Status: The Use and Interpretation of Anthropometry* (Geneva: The World Health Organization).

Appendix A: Data Diagnostics

Table A.1: Comparison of the Distribution of Predictors Across the PNDS and Census Data Sets

		CENSUS		PNDS		K-Smirnov	T-Test
		Mean	SE	Mean	SE	P-value	P-value
age10m	**Dummy for age 10 moths**	**0.019**	**0.0001**	**0.023**	**0.0029**	**1.00**	**0.10**
age11m	**Dummy for age 11 moths**	**0.018**	**0.0001**	**0.022**	**0.0026**	**1.00**	**0.15**
age1223	**1 if child age 12-23 months**	**0.217**	**0.0003**	**0.222**	**0.0073**	**0.98**	**0.51**
age7m	**Dummy for age 7 moths**	**0.019**	**0.0001**	**0.015**	**0.0023**	**1.00**	**0.14**
age8m	**Dummy for age 8 moths**	**0.018**	**0.0001**	**0.017**	**0.0024**	**1.00**	**0.68**
age9m	Dummy for age 9 moths	0.018	0.0001	0.024	0.0027	1.00	0.02
agey_girl	**Interaction ageyrs and girl**	**1.104**	**0.0011**	**1.066**	**0.0259**	**0.61**	**0.14**
ageyrs	**Child's age in years**	**2.245**	**0.001**	**2.202**	**0.0235**	**0.11**	**0.07**
ageysq	**Child's age in years squared**	**6.764**	**0.0043**	**6.605**	**0.1019**	**0.11**	**0.12**
birthord	**Child's birth order for live births**	**2.553**	**0.002**	**2.621**	**0.0522**	**0.38**	**0.19**
car	**1 if HH has car**	**0.206**	**0.0004**	**0.22**	**0.0105**	**0.24**	**0.19**
electr	**1 if HH has electricity**	**0.9**	**0.0003**	**0.913**	**0.0085**	**0.21**	**0.13**
girl	**Child is a girl**	**0.492**	**0.0004**	**0.488**	**0.0089**	**1.00**	**0.69**
hhsize	Number of household members	5.22	0.0026	5.42	0.0579	0.00	0.00
hhsizesq	Number of household members squared	32.413	0.0447	34.455	0.9063	0.00	0.02
ped4ov	1 if husband completed at least 4 years of education	0.731	0.0004	0.701	0.0113	0.00	0.01
peopleroom	**No. HH members per room**	**1.24**	**0.001**	**1.27**	**0.0206**	**0.00**	**0.15**
rage	**Current age – respondent**	**28.287**	**0.007**	**28.177**	**0.1433**	**0.00**	**0.45**
ragesq	Respondent's (mother's) age squared	857.268	0.5294	834.654	8.5417	0.00	0.01
redy12m	**Respondent has 12 or more yrs**	**0.051**	**0.0002**	**0.046**	**0.0058**	**0.95**	**0.37**
redy4	Respondent has 4 yrs of educ	0.149	0.0003	0.178	0.0089	0.38	0.00
redy57	**Respondent has 5-7 yrs of educ**	**0.227**	**0.0004**	**0.219**	**0.0094**	**1.00**	**0.39**
redy8	**Respondent has 8 yrs of educ**	**0.1**	**0.0003**	**0.097**	**0.0062**	**0.97**	**0.58**
redy911	Respondent has 9-11 yrs of	0.206	0.0004	0.179	0.0089	0.52	0.00
refrig	1 if HH has refrigerator	0.738	0.0004	0.676	0.0118	0.00	0.00
tvset	**1 if HH has a television**	**0.597**	**0.0005**	**0.587**	**0.0126**	**0.00**	**0.44**
u5deprat	Under 5 dependency ratio	0.316	0.0001	0.328	0.0033	0.00	0.00
urban	**1 if HH located in Urban area**	**0.777**	**0.0004**	**0.759**	**0.0126**	**0.03**	**0.15**
vcr	1 if HH has a VCR	0.262	0.0004	0.183	0.0098	0.00	0.00
wellspring_in	1 if well or spring water in house for any use	0.19	0.0004	0.135	0.0092	0.00	0.00

Note: The Kolmogorov-Smirnov test is a non-parametric test for equality of the distributions. The T-Test results test the null hypothesis of equality of means across data sets for each variable.
Variables in bold italic when T-test p-value>0.05; Variables underlined when Kolmogorov-Smirnov p-value>0.05.

Appendix B: Estimates of First-stage Height-for-age Models[20]

Model 1: Infant and Toddler Cohorts Only with Full Regressor Set

Table B.1: HAZ Model for Infants

Number of obs	=	1237
F(39, 680)	=	7.40
Prob > F	=	0.0000
R-squared	=	0.2024

haz	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
age7m	.0031096	.2352661	0.01	0.989	-.458782	.4650013
age9m	-.134676	.2309357	-0.58	0.560	-.588066	.3187139
shgarbbuck00	1.065988	.4955849	2.15	0.032	.0930196	2.038957
age8m	-.3128967	.2323129	-1.35	0.178	-.7689904	.143197
dthr1519	-88.98376	86.94562	-1.02	0.306	-259.6818	81.71426
age10m	-.3361157	.2244244	-1.50	0.135	-.7767221	.1044908
age11m	-.3656496	.2254101	-1.62	0.105	-.8081912	.0768921
age1223	-.427469	.1964369	-2.18	0.030	-.8131283	-.0418098
girl	.2333992	.1204163	1.94	0.053	-.003011	.4698093
agey_girl	-.1913917	.1495444	-1.28	0.201	-.4849882	.1022049
shgarbpub~00	1.170127	.2690373	4.35	0.000	.6419337	1.698321
rage	.075847	.042739	1.77	0.076	-.0080614	.1597554
ragesq	-.0013248	.0007645	-1.73	0.084	-.0028258	.0001762
redy4	.1774863	.1039562	1.71	0.088	-.0266081	.3815806
redy57	.3346142	.1092052	3.06	0.002	.1202144	.5490139
redy8	.292159	.1551813	1.88	0.060	-.0125044	.5968224
redy911	.2725428	.1376436	1.98	0.048	.0023108	.5427748
redy12m	.3962685	.2328913	1.70	0.089	-.0609608	.8534978
ped4ov	.1291001	.0924078	1.40	0.163	-.0523217	.3105218
primestr	-.0124003	.0100055	-1.24	0.216	-.0320438	.0072432
hhsize	-.1223752	.0768308	-1.59	0.112	-.2732151	.0284647
hhsizesq	.0068916	.0051799	1.33	0.184	-.003278	.0170612
u5deprat	-1.349893	.3241462	-4.16	0.000	-1.986281	-.7135055
employ98	-.491193	.434291	-1.13	0.258	-1.343825	.3614389
postde99pc	845.625	248.8034	3.40	0.001	357.156	1334.094
peopleroom	-.1058729	.0464654	-2.28	0.023	-.1970972	-.0146486
orcacor97pc	-.0371317	.0282611	-1.31	0.189	-.092616	.0183525
wellspring~n	-.2152219	.1010115	-2.13	0.033	-.4135352	-.0169086
shhhhouse00	4.232899	2.10813	2.01	0.045	.0940638	8.371735
dthr2029	113.5525	60.17965	1.89	0.060	-4.59663	231.7016
refrig	.1581947	.0982926	1.61	0.108	-.0347806	.3511699
vcr	-.222028	.1326361	-1.67	0.095	-.482429	.0383731
ambus299pc	2276.991	1132.324	2.01	0.045	53.92977	4500.051
hospbed00pc	-36.68779	15.45311	-2.37	0.018	-67.02647	-6.349103
car	.2729624	.1116048	2.45	0.015	.0538517	.4920732
birthpcm	-4.848849	3.151954	-1.54	0.124	-11.037	1.339299
shhhapt00	3.846624	2.141556	1.80	0.073	-.3578357	8.051084
dthr14_red~s	8.192569	5.604241	1.46	0.144	-2.810088	19.19523
ambuis99pc	-391.2313	162.536	-2.41	0.016	-710.3339	-72.12865
_cons	-5.093573	2.276579	-2.24	0.026	-9.56312	-.624027

[20] All models were estimated to account for sample design through sampling weights, stratification and clustering. Reported standard errors are Huber-White heteroskedasticity robust standard errors.

Model 1: Infant and Toddler Cohorts Only with Full Regressor Set

Table B.2: HAZ Model for Toddlers

Number of obs	=	2363
F(33, 686)	=	17.75
Prob > F	=	0.0000
R-squared	=	0.2636

haz	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ageyrs	-.0166785	.2826364	-0.06	0.953	-.571571	.5382139
rage	-.0001722	.0304324	-0.01	0.995	-.0599194	.059575
cons99pc	-1061.151	223.9582	-4.74	0.000	-1500.842	-621.4593
ageysq	-.0394832	.0470478	-0.84	0.402	-.1318508	.0528844
girl	-.4495865	.1736045	-2.59	0.010	-.7904196	-.1087534
agey_girl	.1881371	.0563494	3.34	0.001	.0775079	.2987664
birthord	-.076394	.0192651	-3.97	0.000	-.1142167	-.0385713
shgarbbuck00	-.669874	.2820328	-2.38	0.018	-1.223582	-.1161664
ragesq	.0002645	.0004906	0.54	0.590	-.0006986	.0012275
redy4	.0875134	.089197	0.98	0.327	-.0876046	.2626314
redy57	.1879906	.0852518	2.21	0.028	.020618	.3553633
redy8	.4518498	.1009102	4.48	0.000	.2537354	.6499642
redy911	.3701898	.1135351	3.26	0.001	.1472893	.5930904
redy12m	.7170499	.1703012	4.21	0.000	.3827021	1.051398
ped4ov	.0979943	.0675436	1.45	0.147	-.0346122	.2306008
shhhapt00	2.354303	1.995757	1.18	0.239	-1.563914	6.27252
hhsize	-.1517244	.0438739	-3.46	0.001	-.2378609	-.0655879
hhsizesq	.007752	.0027127	2.86	0.004	.0024262	.0130779
u5deprat	-.5507253	.2178056	-2.53	0.012	-.9783373	-.1231133
hospbed00pc	23.9921	18.4192	1.30	0.193	-12.16983	60.15403
intesp00pc	-1.498231	.8713009	-1.72	0.086	-3.208833	.2123704
peopleroom	-.1094327	.0364581	-3.00	0.003	-.18101	-.0378554
dthr1519	62.65815	52.9485	1.18	0.237	-41.29423	166.6105
wellspring~n	.09029	.0757322	1.19	0.234	-.0583931	.2389731
electr	.1359447	.1108853	1.23	0.221	-.0817534	.3536428
dthrate	51.33245	19.98792	2.57	0.010	12.0907	90.5742
refrig	.3828001	.0722843	5.30	0.000	.2408863	.524714
vcr	-.1153503	.0795469	-1.45	0.147	-.2715226	.0408221
funFUN00	-.0033198	.0023102	-1.44	0.151	-.0078553	.0012157
tvset	.1263379	.0679333	1.86	0.063	-.0070337	.2597095
shhhouse00	2.31183	1.943492	1.19	0.235	-1.503777	6.127437
birthpcm	-11.04505	2.917468	-3.79	0.000	-16.77283	-5.317259
dthr01_red~s	-.6440162	.3178927	-2.03	0.043	-1.268126	-.019906
_cons	-2.015342	2.048312	-0.98	0.325	-6.036738	2.006055

Appendix C: The Geographic Distribution of Malnutrition Prevalence in Brazil

Figure C.1: Malnutrition Prevalence Estimates used by Bolsa Alimentação

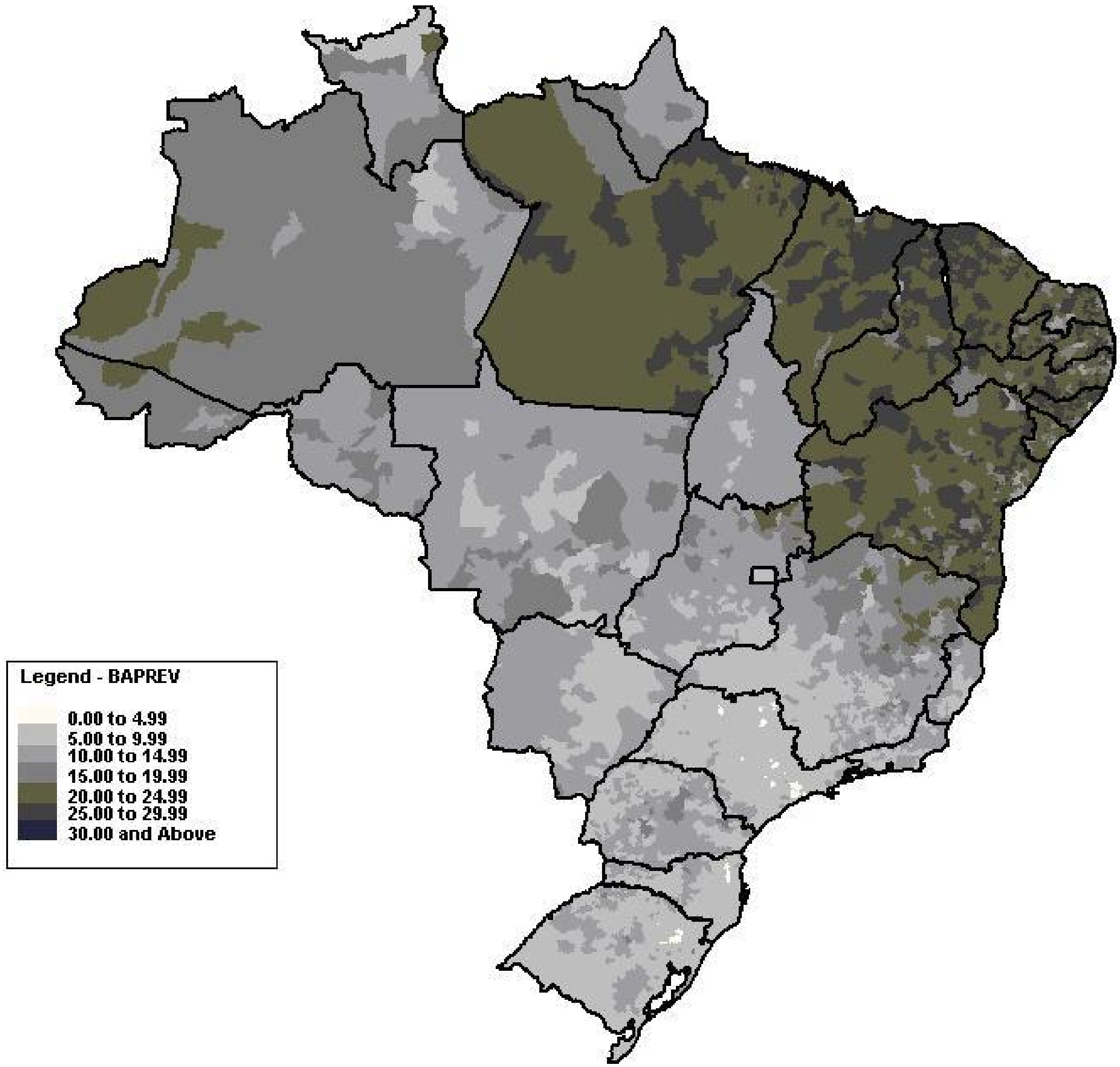


Figure C.2: Evaluation Model 1 Stunting Prevalence Estimates

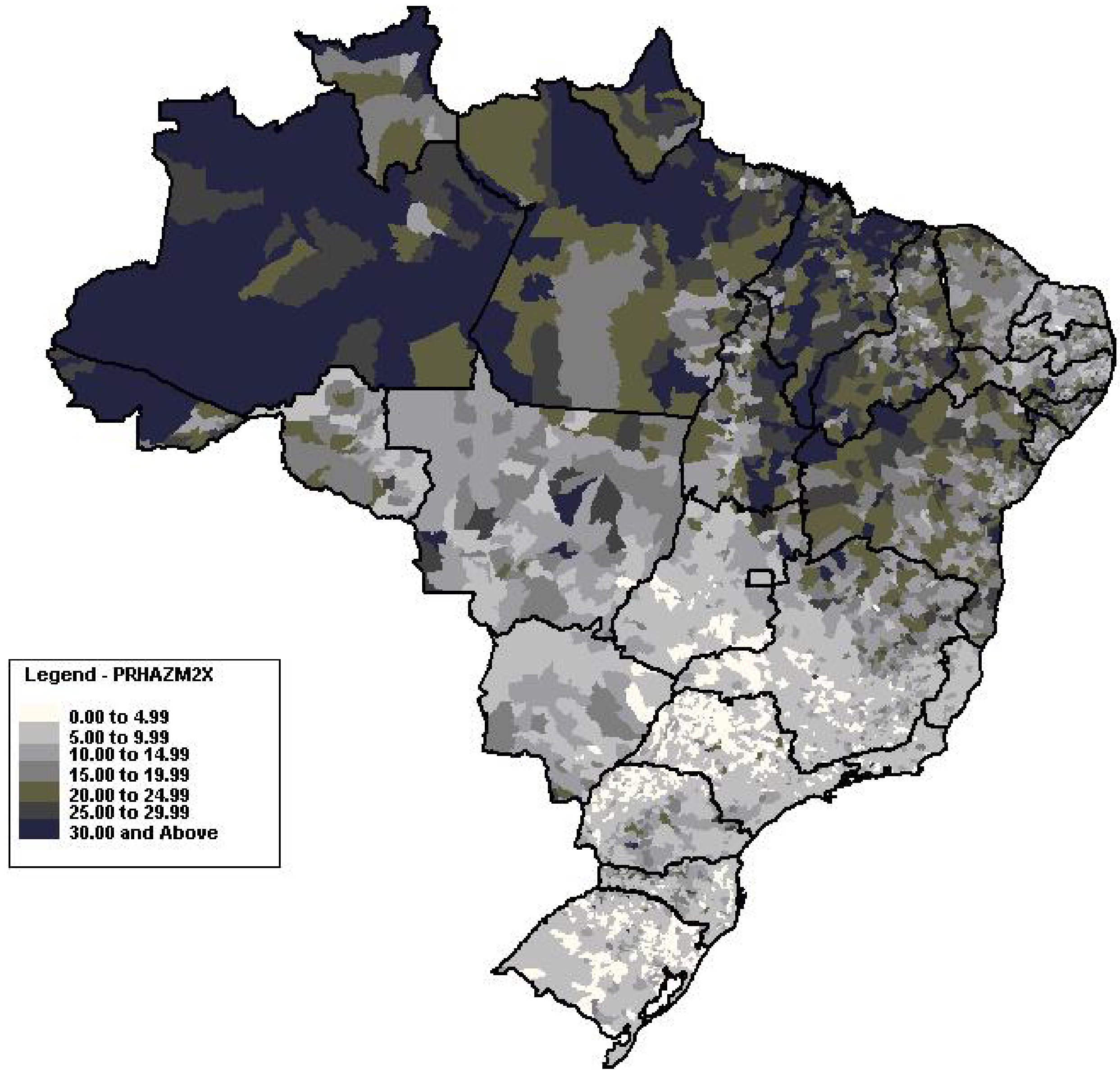


Figure C.3: Evaluation Model Implied Stunting Prevalence Estimates Conditioned on BA Program Size (HAZ<-1.823)

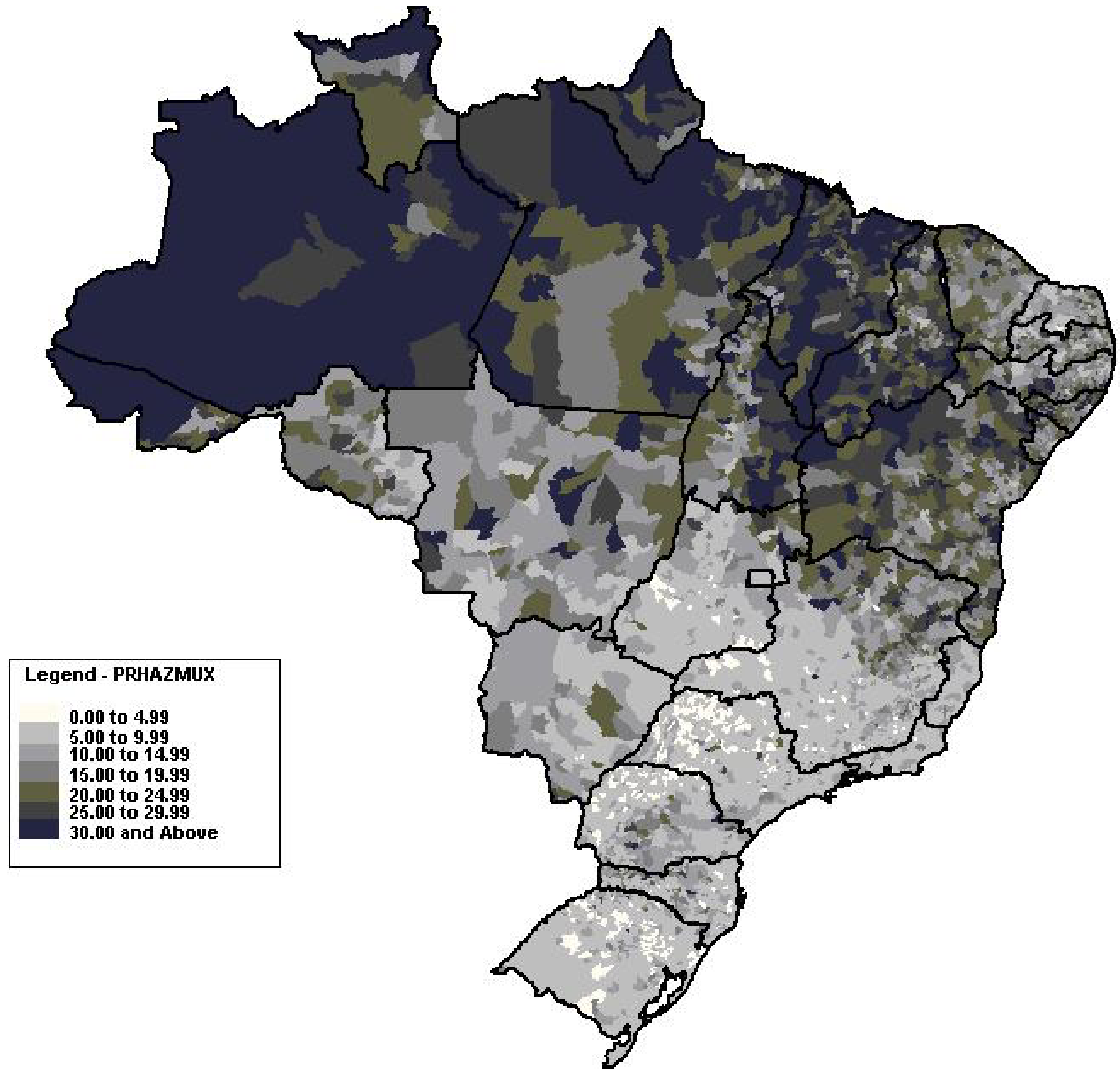


Figure C.4: Evaluation Stunting Model 2, based on Reduced X Matrix

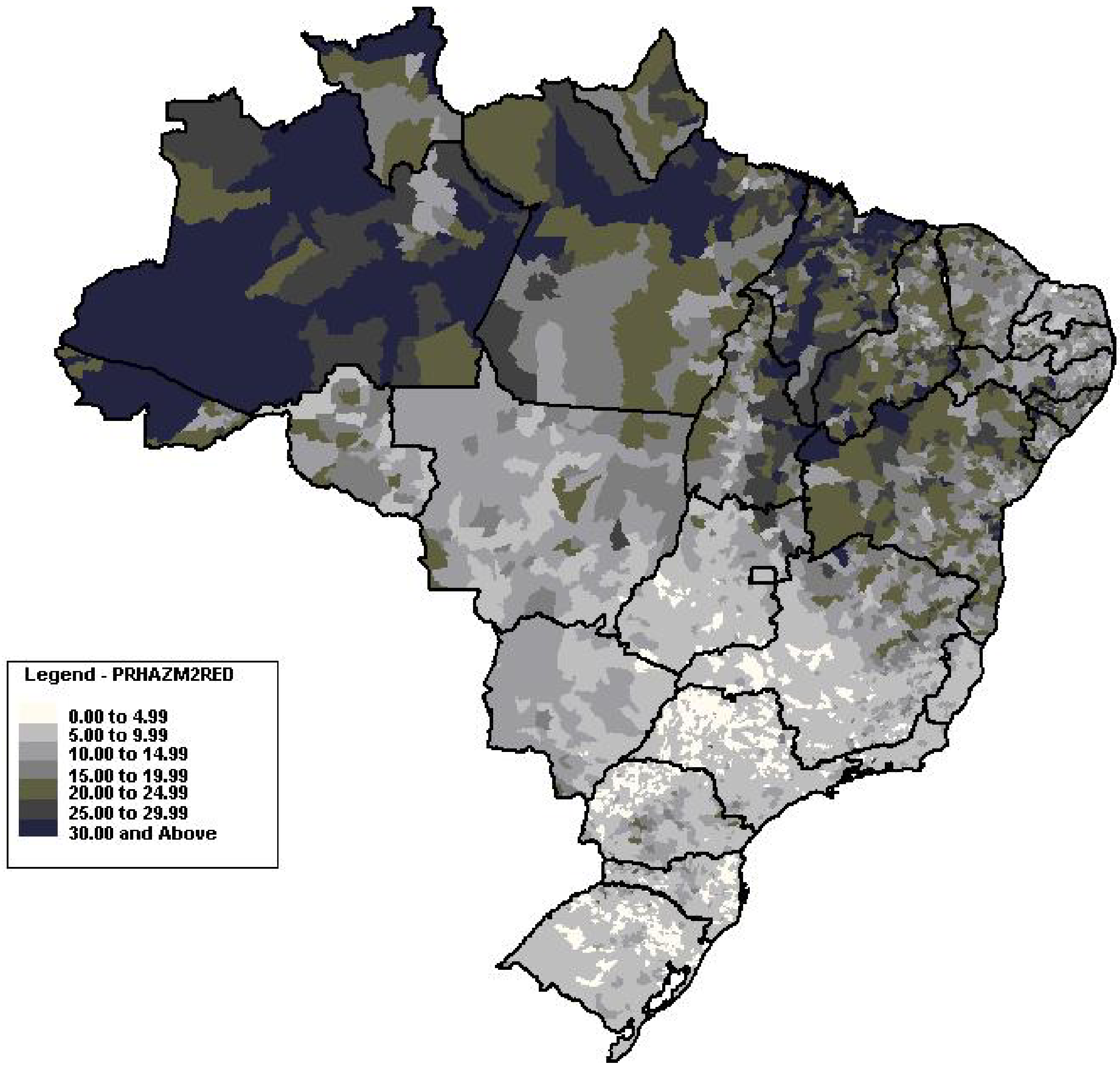


Figure C.5: Evaluation Stunting Model 3, based on Regional Models