Supplementary appendix

Supplement to: Rasella D, Aquino R, Santos CAT; Paes-Sousa R, Barreto ML. Effect of a conditional cash transfer programme on childhood mortality: analysis of nationwide data of Brazilian municipalities.

1. Multivariate Models for the Association of Under-five Mortality Rates with Different *Bolsa Familia* Program Coverage Indicators

Two different yearly coverages of the Bolsa Familia Program (BFP) were considered: BFP coverage of the target population (TP) and BFP coverage of the total population of the municipality. The TP coverage was calculated as the number of families enrolled in the BFP program in the municipality divided by the number of eligible families (according to the BFP criteria) in the same municipality.[1] The second was the BFP coverage over the total population, calculated as the number of individuals enrolled in the BFP (obtained multiplying the number of beneficiary families per the average family size) over the total population of the same municipality.

Different models have been fitted using these two indicators as continuous or categorized variables. While continuous variable allow to estimate the existence of an association along the entire range of values of a variable, categorized variables gives an easily interpretable measure of effect. Moreover the use of different levels of coverage allow to verify the existence of a gradient of effect, related to different degrees of implementation of the interventions.[2,3]

Multivariable negative binomial models with all variables expressed as continuous variables were fitted and the results are shown in Table S1.

In order to obtain models with categorized variables, the BFP coverage of the total population of the municipality, much like the FHP program, was grouped as follows: low coverage (coverage of <70.0%), intermediate coverage (coverage of 70,0% to 99.9%), and high coverage (coverage of ≥ 100.0%). BFP coverage of the total population of the municipality was calculated as the sum of people receiving benefits from the BFP over the total population of the same municipality, and – in the absence of any literature reference for this kind of coverage - was classified according to terciles of the distribution: low coverage (first tercile, from 0.0% to 17.1%), intermediate coverage (second tercile, from 17.2% to 32.0%) and high coverage (third tercile, higher than 32.0%).

Multivariable negative binomial models with categorized variables were fitted as described in the methods section of the article (Table S2).

The BFP has a high targeting accuracy compared to CCTs of others countries [4] and it has been shown that even families mistakenly included in the program are often poor or low income.[5] The fact that the municipality population coverage, controlled for the TP percentage, seems to have a slightly higher effect than the Bolsa Familia TP coverage could be explained by the inclusion in this indicator of low income, but not eligible, families, and by the effects of the program's externalities.[6]

TABLE S1. Fixed-Effect Negative Binomial Models for the Association Between Under-five Mortality Rates and BFP Coverage, Expressed as Continuous Variables: Brazil, 2004–2009

Variables	Under-fives mortality rate, RR (95%CI)			
	Crude	Adjusted	Crude	Adjusted
BFP coverage of TP	0.997 (0.997-0.998)*	0.999 (0.999-0.999)*	-	-
BFP municipality population coverage	-	-	0.992 (0.991-0.993)*	0.997 (0.996-0.999)*
FHP population coverage	-	0.999 (0.999-0.999)*	-	0.999 (0.999-0.999)*
Per capita income (monthly)	-	0.999 (0.999-0.999)*	-	0.999 (0.999-0.999)*
Percentage of target poor population	-	1.005 (1.002-1.009)*	-	1.007 (1.003-1.010)*
Percentage of individuals living in households with inadequate sanitation	-	1.010 (1.007-1.013)*	-	1.007 (1.004-1.011)*
Percentage of illiterates among individuals over 15 years old (inverse)**	-	1.191 (1.070-1.325)*	-	1.159 (1.040-1.291)*
Total fertility rate	-	1.048 (1.002-1.096)*	-	1.041 (0.996-1.088)
Hospitalization rate (per 100 inhabitants)	-	0.997 (0.991-1.003)	-	0.997 (0.991-1.003)
No. of observations	17118	17118	17118	17118
No. of counties	2853	2853	2853	2853

* P value < 0.05
** Illiteracy rate has been transformed into its scaled inverse because caused collinearity problems to the model (VIF>6)

TABLE S2. Fixed-Effect Negative Binomial Models for the Association Between Under-five Mortality Rates and BFP Coverage, Expressed as Categorized Variables: Brazil, 2004–2009

Variables	Under-fives mortality rate, RR (95%CI)			
	Crude	Adjusted	Crude	Adjusted
BFP coverage of TP				
Low (<70%)	1	1	-	-
Intermediate (70.0% to 99.9%)	0.92 (0.90-0.93)	0.93 (0.92-0.94)	-	-
High (>=100.0%)	0.88 (0.86-0.89)	0.90 (0.89-0.92)	-	-
BFP municipality population coverage				
Low (0.0% to 17.1%)	-	-	1	1
Intermediate (17.2% to 32.0%)	-	-	0.91 (0.89-0.93)	0.94 (0.92-0.96)
High (>32.0%)	-	-	0.82 (0.79-0.84)	0.87 (0.84-0.90)
FHP population coverage				
No FHP (0.0%)	-	1	-	1
Incipient (<30%)	-	0.99 (0.95-1.04)	-	0.99 (0.95-1.04)
Intermediate (>= 30%)	-	0.94 (0.90-0.99)	-	0.93 (0.89-0.98)
Consolidate (>= 70% and time of implementation in the municipality of 4 years or longer)	-	0.91 (0.87-0.96)		0.90 (0.86-0.95)
Per capita income (monthly) > 380 BR	-	0.96 (0.93-0.98)	-	0.94 (0.92-0.97)
Percentage of target poor population> 22.4		1.07 (1.02-1.11)		1.07 (1.03-1.12)
Percentage of individuals living in households with inadequate sanitation <16.7		1.10 (1.05-1.15)		1.11 (1.06-1.16)
Percentage of illiterates among individuals over 15 years old >11.1%		1.05 (1.00-1.09)		1.05 (1.00-1.09)
Total fertility rate > 2.32		1.05 (1.01-1.08)		1.06 (1.02-1.09)
Hospitalization rate (per 100 inhabitants) > 4.27		1.01 (0.98-1.04)		1.02 (0.99-1.05)
No. of observations	17118	17118	17118	17118
No. of counties	2853	2853	2853	2853

2. Estimating the percentage of deaths from vulnerable segments of the population

Considering the unit of analysis of an ecological study, for example a county, divided into two different population groups with different mortality rates (MR), with MR_p being the MR of the poorest part of the population, and MR_r being the MR of the rest of the population, the Rate Ratio (RR) is: $RR = MR_p / MR_r$. If we consider the deaths from the poorest group (D_p) over the population of the poorest group (P_p) and the deaths from the rest of the population (D_r) over the rest of the population (P_r), We can obtain the total deaths in the county (D_{tot}) and the total population (P_{tot}) from the following equations:

$$MR_p = RR \times MR_r$$

$$D_p / P_p = RR \times D_r / P_r$$

$$D_p = RR \times (D_{tot} - D_p) \times P_p / P_r$$

$$D_p \times (1 + RR \times P_p / P_r) = RR \times P_p \times D_{tot} / P_r$$

$$D_p = RR \times P_p \times D_{tot} / [P_p \times (RR-1) + P_{tot}]$$

$$D_p = K_p \times D_{tot}$$

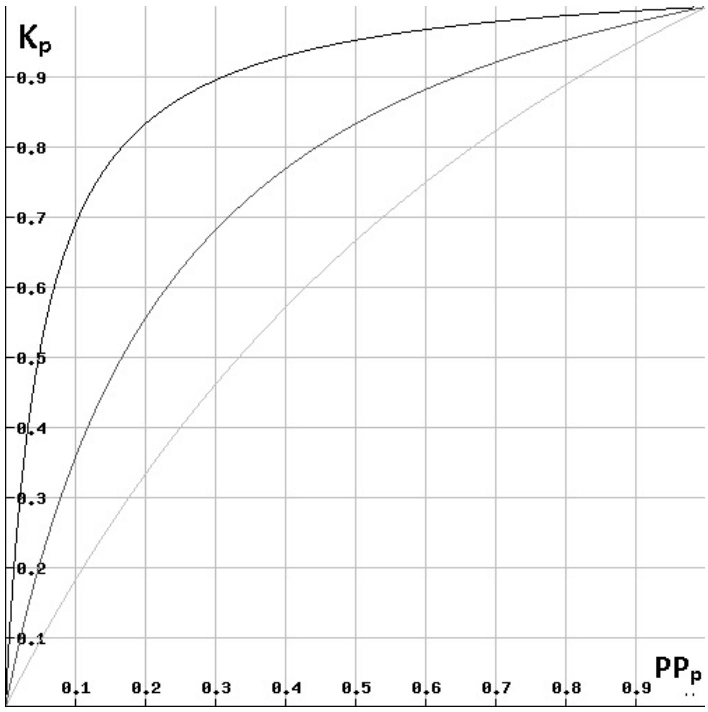
K_p represents the proportion of deaths that come from the poorest segment of the county's population, and depends on the Mortality Rate Ratio between the two population groups and the proportion of poor people over the total population of the county (PP_p) according to the following equation:

$$K_p = RR \times PP_p / [PP_p \times (RR-1) + 1]$$

Considering different values of Mortality Rate Ratios, the proportion of deaths that come from the poorest part of the population has a curvilinear relationship with the proportion of poor people in the county (FIGURE S1).

If we consider an under-five mortality rate ratio (RR) of 2.3,[7] in a county with 30% poor people the proportion of under-five deaths attributable to them will be 50%. In the case of segments of the population in extreme poverty the RR, and consequently the proportion of deaths attributable to them, will be considerably higher. The RR for specific causes,[8] especially poverty-related causes such as malnutrition or diarrhea, can be so high that the deaths attributable to extremely poor people reach almost the totality of the deaths for this specific cause in the county, as shown in the figure.

FIGURE S1: Proportion of deaths coming from the poorest part of the population (K_p) as function of the proportion of poor people in the county (PP_p) and according to different values of mortality rate ratio (RR).



Light grey: RR=2; Dark grey: RR=5; Black: RR=20

3. Negative binomial regression models with fixed effect specifications in impact evaluations

Negative Binomial Models

Negative binomial (NB) regression models are used when the outcome to be analyzed is a count data and the Poisson model assumption that the mean is equal to the variance does not hold, usually because the data are *overdispersed*.[9] The standard negative binomial regression model can be derived either as a Poisson-gamma mixture model or as a member of exponential distributions used as a basis for generalized linear models.

The NB regression can be used with longitudinal or panel data, where the same unit of analysis has repeated observations over a period of time.[10] In this case, in addition to the disturbance or error term, panel data models include a second term to control for unobserved time-invariant characteristics of the unit of analysis, or panel. According to how this term is estimated, the models can be distinguished in fixed effects or random effects models. From a statistical point of view, the choice between fixed-effects and random-effects models is based on the Hausman specification test.[10,11]

Fixed Effects Models in Impact Evaluations

In impact evaluations fixed effects (FE) models are usually preferred because they permit correlations between the unobserved time-invariant term and the explanatory variables.[12] In our case the time-invariant term could represent unobserved characteristics of the municipality such as geographical, historical, socio-cultural or socio-economic characteristics that did not change during the period of the study. In fixed effects models, but not in random ones, those characteristics could be correlated with the treatment variables, such as the BFP or FHP coverage. If for example these interventions were implemented with priority in remote and poor areas with higher mortality rates, and variables linked to those characteristics were not included in the model, the estimates of the intervention effects could suffer from selection bias. Fixed effects models allow to control for this selection bias because the fixed effect term of the equation represents these unobserved time-invariant characteristics of the panel.[12]

The Regression Model

The regression model to be estimated was as follows:

$$Y_{it} = \alpha_i + \beta_1 BFP_{it} + \beta_2 FHP_{it} + \beta_n X_{nit} + u_{it}$$

Where Y_{it} was the mortality rate for the municipality *i* in year *t*, α_i is the fixed effect for the municipality *i* that captures all unobserved time-invariant factors, BFP_{it} is the Bolsa Familia Program coverage for the municipality *i* in the year *t*, FHP_{it} the Family Health Program coverage for the municipality *i* in the year *t*, X_{nit} was the value of each *n* covariate of the model with in the municipality *i* in the year *t,* and u_{it} was the error.

A variable representing time was not included in the model because the mortality rate ratio, comparing two or more groups of coverage exposed to the same mortality time trend, allowed us to control for secular trends.[2,3] Introduced in the models a time variable would have represented an over-specification problem, as confirmed by sensitivity analyses that have been performed. The fact that these models control for secular trends was confirmed by the estimates of the BFP and FHP effect on U5MR due to external causes: despite this group of causes presented a decreasing mortality in the studied period the coverage of the two programs, which had increased in the same period, did not show any effect on it.

Fixed Effects Negative Binomial Models

Fixed effects negative binomial (FENB) models may be estimated in two ways, unconditionally or conditionally.[13] Conditional models are usually preferred and implemented in the classical statistical software packages because they can adjust for a large number of panels without creating dummy slopes for each panel, that is extremely time and computing memory consuming if the number of panels is large. However it has been shown that the conditional maximum likelihood estimator of the FENB does not necessarily remove the individual fixed effects in count panel data, this happens only in specific conditions.[13,14]

Different solutions have been proposed. According to the literature the more appropriate - even if time-consuming - is the fitting of unconditional FENB scaling its standard error (SE) by the Person Chi2 or the deviance dispersion.[9,13,15]

As it is shown in TABLE S3, in order to verify the robustness of our analysis, we fitted the panel data models related to all-causes under-five mortality rate using three different model specifications: 1- Conditional FENB, 2- Unconditional FENB with scaled SE, 3- conditional FE Poisson with robust SE.

The estimated effects of BFP and FHP (and of the covariates) are almost identical in all these models. The values of the Akaike information criterion (AIC) and Bayesian information criterion (BIC), that due to their formula was possible to calculate only for the models 1 and 3, suggest that the conditional FENB is the models that better fits the data.

The same comparison of model specifications have been performed for all the other mortality outcomes of the study: conditional FENB models show similar effect estimates but better AIC and BIC that Poisson models with robust SE, on the other hand unconditional FENB models have problems of convergence in some outcomes - probably due to the high number of parameters calculated - but when convergent show similar values to the conditional FENB.

Considering that the negative binomial regression is the model that better fit our *overdispersed* mortality data, that the fixed effects is an important specification for impact evaluation analysis, and that the conditional FENB demonstrated to behave in our models - comparing its estimates with the unconditional FENB and Poisson regressions - as true fixed effects models, we decided to use for the analysis of our panel dataset models with conditional FENB specifications.

TABLE S3: Fixed effects Regression Models for the Association Between Under-five Mortality Rates (U5MR) and BFP Coverage with Different Model Specifications: Brazil, 2004–2009

Variables	U5MR, RR (95%CI)		
	1. Conditional FENB	2. Unconditional FENB with scaled SE*[a]	3. Conditional FE Poisson with Robust SE
BFP population coverage			
Low (0.0% to 17.1%)	1	1	1
Intermediate (17.2% to 32.0%)	0.94 (0.92-0.96)	0.95 (0.93-0.97)	0.94 (0.91-0.97)
High (>32.0%)	0.88 (0.85-0.91)	0.89 (0.85-0.92)	0.88 (0.84-0.92)
Consolidate (>32.0% and TP coverage>=100% for 4 years or longer)	0.83 (0.79-0.88)	0.84 (0.79-0.89)	0.84 (0.78-0.89)
FHP municipality population coverage			
No FHP (0.0%)	1	1	1
Incipient (<30%)	0.99 (0.94-1.04)	0.98 (0.93-1.03)	0.99 (0.94-1.04)
Intermediate (>= 30%)	0.93 (0.88-0.97)	0.94 (0.89-0.98)	0.93 (0.88-0.98)
Consolidate (>= 70% and time of implementation in the municipality of 4 years or longer)	0.88 (0.83-0.93)	0.88 (0.83-0.94)	0.87 (0.82-0.93)
Per capita income (monthly) > 380 BR$	0.95 (0.92-0.97)	0.94 (0.92-0.97)	0.94 (0.91-0.98)
Percentage of TP > 22.4%	1.07 (1.03-1.12)	1.07 (1.01-1.13)	1.07 (1.02-1.13)
Percentage of individuals living in households with inadequate sanitation <16.7%	1.10 (1.05-1.15)	1.10 (1.05-1.15)	1.09 (1.04-1.15)
Percentage of illiterates among individuals over 15 years old >11.1%	1.04 (1.00-1.08)	1.04 (1.00-1.09)	1.04 (0.99-1.09)
Total fertility rate > 2.32	1.07 (1.03-1.10)	1.07 (1.04-1.11)	1.07 (1.03-1.11)
Hospitalization rate (per 100 inhabitants) > 4.27	1.01 (0.99-1.04)	1.00 (0.97-1.03)	1.01 (0.97-1.06)
No. of observations	17118	17118	17118
No. of counties	2853	2853	2853
AIC	52,962	*[b]	53,070
BIC	53,063		53,163

*[a] SE scaled by the Pearson chi-square statistic divided by the residual degrees of freedom, scaling by deviance statistics gave similar results
*[b] Not possible to be estimated according to the AIC and BIC formula;
FENB: Fixed Effects Negative Binomial,
FE: Fixed Effects

References

1. MI. Matriz de Informação Social. MDS http://aplicacoes.mds.gov.br/sagi/mi2007/tabelas/mi_social.php (accessed February 11, 2013).

2. Aquino R, Oliveira NF, Barreto ML. Impact of the Family Health Program on infant mortality in Brazilian municipalities. *Am J Public Health* 2009; **99**:87–93

3. Rasella D, Aquino R, Barreto ML. Reducing childhood mortality from diarrhea and lower respiratory tract infections in Brazil. *Pediatrics* 2010; **126**:e534-40.

4. Soares FV, Ribas RP, Osorio RG. Evaluating the impact of Brazil's Bolsa Família: cash transfer programmes in comparative perspective. Brasília: *International Poverty Centre*, 2007.

5. Lindert K, Linder A, Hobbs J, Briere B. The Nuts and Bolts of Brazil's Bolsa Família Program: Implementing Conditional Cash Transfers in a Decentralized Context. Discussion Paper n.0709. Brasília: *World Bank* ; 2007. http://siteresources.worldbank.org/INTLACREGTOPLABSOCPRO/Resources/BRBolsaFamiliaDiscussionPaper.pdf (acessed 11/05/2012).

6. Sudhanshu Handa and Benjamin Davis. The Experience of Conditional Cash Transfers in Latin America and the Caribbean. *Development Policy Review* 2006; **24**:513-536.

7. Rodwin VG, Neuberg LG. Infant mortality and income in 4 world cities: New York, London, Paris, and Tokyo. *Am J Public Health* 2005; **95**:86-90.

8. Nelson MD. Socioeconomic status and childhood mortality in North Carolina. *Am J Public Health* 1992; **82**:1131–1133.

9. Hilbe JM. Negative Binomial Regression. Cambridge: *Cambridge University Press*, 2007.

10. Frees, EW. Longitudinal and Panel Data. Cambridge: *Cambridge University Press*, 2004.

11. Wooldridge JM. Introductory Econometrics, a modern approach. 3rd edition. Cincinnati: *South-Western College Pub*; 2005.

12. Shahidur RK,. Koolwal GB, Samad HA. Handbook on Impact Evaluation: Quantitative Methods and Practices. Washington: *World Bank Publications*, 2010.

13. Allison PD, Waterman RP. Fixed–Effects Negative Binomial Regression Models. *Sociological Methodology* 2002; **32**: 247–65.

14. Guimarães P. The fixed effects negative binomial model revisited. *Economics Letters* 2008; **99**: 63–6.

15. Allison P. Beware of Software for Fixed Effects Negative Binomial Regression. http://www.statisticalhorizons.com/fe-nbreg. (accessed June 12, 2012)