Courant Research Centre
'Poverty, Equity and Growth in Developing and Transition Countries: Statistical Methods and Empirical Analysis'

Georg-August-Universität Göttingen
(founded in 1737)

Discussion Papers

No. 21

Matching Household Surveys with DHS Data to Create Nationally Representative Time Series of Poverty: An Application to Bolivia

Melanie Grosse, Stephan Klasen, Julius Spatz

December 2009

Platz der Göttinger Sieben 3 · 37073 Goettingen · Germany
Phone: +49-(0)551-3914066 · Fax: +49-(0)551-3914059

Email: crc-peg@uni-goettingen.de Web: http://www.uni-goettingen.de/crc-peg

Matching Household Surveys with DHS Data to Create Nationally Representative Time Series of Poverty: An Application to Bolivia

Melanie Grosse, Stephan Klasen, and Julius Spatz *

This version: December 19, 2009

Abstract

In many developing countries, there does not exist a time series of nationally representative household budget or income surveys, while there often are urban household surveys as well as nationally representative Demographic and Health Surveys (DHS) which lack information on incomes. This makes an analysis of trends and determinants of poverty and inequality over a longer time period impossible. This is also the situation in Bolivia where there exist urban household surveys and nationally representative DHS since 1989, while nationally representative household income surveys only exist since 1997. In this paper, we adjust a technique developed for poverty mapping exercises to link urban household income surveys with DHS data to generate a nationally representative time series of household income data from 1989 to 1999. Our technique performs well on validation tests, is superior to proxying welfare with asset ownership in the DHS, and is able to generate new information on poverty and inequality in Bolivia.

***JEL* Classification:** C81, D31, I31, I32, O54.

Key words: Microsimulation, survey matching, poverty, inequality, pro-poor growth, poverty profile, growth incidence curve, Bolivia

*University of Göttingen, Department of Economics, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany, email: mgrosse@uni-goettingen.de; sklasen@uni-goettingen.de; julius.spatz@gtz.de. An earlier version if this paper was produced as part of the World Bank-DFID-AfD-BMZ-GTZ-KfW coordinated process on operationalizing pro-poor growth (Klasen et al., 2007). We would like to thank Juan-Carlos Aguilar, Annette Langhammer, Louise Cord, Berk Özler, Omar Arias, Fernando Landa, Wilson Jimenez, and Sara Calvo. Furthermore we would like to thank participants and discussants at workshops at the World Bank, in Eschborn, La Paz, and Harvard, at the Symposium on Poverty, Inequality, and Policy in Latin America in Göttingen, at the 29th IARIW General Conference in Joensuu, the Spring Meeting of Young Economists in Geneva, and the European Society for Population Economics, Paris, for valuable comments and discussion. Special thanks go to Francisco H. G. Ferreira, Peter Lanjouw, Martin Ravaillon, Walter Zucchini, Stefan Sperlich, and Andrey Launov for detailed comments and discussion. Funding from the German Federal Ministry for Economic Cooperation and Development via the KfW Entwicklungsbank (KfW Development Bank) is gratefully acknowledged.

1 Introduction

In many developing countries, it is not possible to obtain a time series of household income surveys for poverty and inequality analyses. Nationally representative surveys often only started very recently (e.g., with the support of the World Bank living standard measurement survey (LSMS) program), and before there are often only regional, frequently urban, income surveys available. At the same time, many developing countries have participated in the program of Demographic and Health Surveys (DHS) since the late 1980s and often now have 2–4 such surveys available. This is, for example, the case in most Latin American countries before the mid-1990s, for example in Colombia with 5 DHS from 1986 onwards but household income surveys only since the mid-1990s or Peru with also 5 DHS since 1986 and national household surveys only from 1997 onwards. Other examples are Haiti with 3 DHS but only 1 income survey of 2001. This is also similar in several Sub-Saharan African countries where the 1-2-3 surveys are typically only urban. The great advantage of the DHS is the high degree of standardization over time (and countries) as well as that they are freely available. Unfortunately, these DHS data do not contain information on household incomes or expenditures. In order to use these data nevertheless for poverty analysis, asset indices have often been created and used to assess poverty differentially and poverty trends over time (Sahn and Stifel, 2000, 2003; Filmer and Pritchett, 2001). While these asset indices are often well-correlated with income, it is not clear how well they are able to reproduce poverty trends over time.

To explore the trends in the urban-rural divide as well as other dimensions of poverty in more depth and detail irrespective of the above mentioned data constraints, we set up a dynamic cross-survey microsimulation methodology. Our approach basically follows the poverty mapping literature based on Hentschel et al. (2000) and Elbers et al. (2003) who use household surveys and census data to generate detailed poverty maps at one point in time. A more recent application is done by Stifel and Christiaensen (2007) who use a single household survey and several DHS surveys to generate poverty data over time, i.e., for several years over one decade. Different to the first two studies is that we develop a dynamic component rather that the static poverty mapping within a single (or nearby) year. Different to the third study is that we explicitly model dynamics rather than assuming that there are none. In Section 2, we start by developing the methodology and describing the data used. The empirical application for the case of Bolivia in Section 3 is carried out in three steps. First, we generate an inter-temporally comparable microdata set of simulated incomes for total Bolivia (i.e., national-wide and separately for departmental capitals (short cities), other urban areas (short towns), and rural areas) between 1989 and 1999, and check the consistency between observed and simulated incomes where the former are available. Second, we use the simulated incomes to estimate detailed national poverty profiles by place of residence and by household characteristics to track the evolution of poverty for different subgroups of the population over time.[1] Third, we evaluate the "pro-poorness" of the simulated 1989–to–1999 income changes using growth incidence curves.[2] In Section 4, we perform sensitivity analyses to (a) check the robustness of our results to alternative specifications and assumptions and to (b) compare our results with those derived from the asset-index approach. In Section 6, we discuss the results.

2 Approach and Data

Our methodology to create a nationally representative time series of income data out of incomplete income or consumption expenditure data and to thereof derive poverty profiles and growth incidence curves builds upon the static cross-survey microsimulation methodology of Hentschel et al. (2000) and Elbers et al. (2003). Their objective is to analyze the spatial dimension of poverty in detailed poverty maps of national coverage for Ecuador. Their problem is that the Ecuadorian LSMS did not collect consumption expenditures for all households but only for a nationally representative sample of two-stage randomly selected households. The two-stage sample design, first selecting clusters and then households within the selected clusters, generates a sample in which the households are not randomly distributed over space, but are geographically grouped. Their solution to this problem is to combine the LSMS data with concurrent unit-record Census data of all Ecuadorian households and impute consumption expenditures for those municipalities which were not included in the LSMS sample. To this end, they estimate a consumption expenditure model in the LSMS data restricting the set of covariates to those which are also available in the Census data. Then they multiply for each household in the Census its covariates with the corresponding regression coefficient from the consumption expenditure model and add a randomly distributed error term.

[1] In a related study, Klasen et al. (2007) investigate also the effect of macroeconomic shocks and policies on poverty and inequality for a 10-years-period ahead. The authors use a dynamic computable general equilibrium model that is linked to the microdata also used in this study.

[2] In earlier versions of this paper (Grosse et al., 2005, 2007), we also presented the results of the Datt and Ravallion (1992) growth-inequality decomposition of poverty changes.

We have a similar objective but face a different data constraints. The pre–1997 LSMS of Bolivia are not nationally representative, but cover only the cities.[3] Additionally, the Bolivian rounds of Census are only available for 1992 and 2001. To overcome these data constraints, we extend the static cross-survey microsimulation methodology of Hentschel et al. (2000) and Elbers et al. (2003) by a dynamic component and use DHS data instead of Census data.

Stifel and Christiaensen (2007) apply the same technique, which is also based on Hentschel et al. (2000) and Elbers et al. (2003), to Kenyan data facing similar data constraints as we do. They use a household survey—the 1997 welfare monitoring survey (WMS)—and the three DHS rounds of 1993, 1998, and 2003. The difference to our paper is that their estimation procedure, despite predicting incomes to the past and future, remains stable concerning the modelling of the regression coefficients and the error terms over time. This means that they run a log-linear regression model in the WMS of 1997, and they use the coefficients (and error terms) obtained from the model in all three DHS surveys to simulate incomes. They argue that there are some parameters that are expected to be relatively stable over time (e.g., consumer durables or housing characteristics) and exclude others that are expected to be instable over time (e.g., education or employment). Testing if the parameters are stable or not, however, is not possible with their data set. Theoretical arguments on their selection strategy are scarce; instead their selection is based on stepwise regression models.

Our methodology takes dynamics explicitly into account and proceeds in three steps. First, we choose a base period t in which we have a nationally representative LSMS as well as a nationally representative DHS, and develop an empirical model of a monetary welfare indicator y (hereafter referred to as income) using the LSMS data. Similar to Stifel and Christiaensen (2007), Hentschel et al. (2000), and Elbers et al. (2003), we restrict the set of covariates X to those which are also available in the corresponding DHS. We choose the covariates to exhibit (a) the highest possible consistency between LSMS and DHS data as well as over time, and (b) the best possible fit of the regression model. We then construct a 3 x 3 block diagonal structure of the covariates by interacting them with three regional

[3] The 1997 LSMS is nationally representative but not comparable over time due to changes in the survey design.

dummies, and run a weighted standard log-linear OLS regression model

$$\begin{pmatrix} y_t^C \\ y_t^T \\ y_t^R \end{pmatrix} = \begin{pmatrix} X_t^C & 0 & 0 \\ 0 & X_t^T & 0 \\ 0 & 0 & X_t^R \end{pmatrix} \begin{pmatrix} \beta_t^C \\ \beta_t^T \\ \beta_t^R \end{pmatrix} + \begin{pmatrix} \varepsilon_t^C \\ \varepsilon_t^T \\ \varepsilon_t^R \end{pmatrix}, \tag{1}$$

where the indices C, T, and R stand for cities, towns, and rural areas, respectively, β are coefficient vectors, and ε is an independent error term. We account for heteroskedasticity using the covariance matrix estimator proposed by White (1980).[4] We predict incomes within the LSMS sample to detect problems that might arise from the modelling of the error term (see below).

Second, we check the consistency between the observed incomes of the LSMS and the simulated incomes of the DHS in period t. To this end, we apply the coefficient estimates $\hat{\beta}$ from regression model (Equation 1) to the DHS covariates $\tilde{X}$ and generate simulated incomes $\tilde{y}$

$$\begin{pmatrix} \tilde{y}_t^C \\ \tilde{y}_t^T \\ \tilde{y}_t^R \end{pmatrix} = \begin{pmatrix} \tilde{X}_t^C & 0 & 0 \\ 0 & \tilde{X}_t^T & 0 \\ 0 & 0 & \tilde{X}_t^R \end{pmatrix} \begin{pmatrix} \hat{\beta}_t^C \\ \hat{\beta}_t^T \\ \hat{\beta}_t^R \end{pmatrix} + \begin{pmatrix} u_t^C \\ u_t^T \\ u_t^R \end{pmatrix}. \tag{2}$$

Since the regression model in Equation (1) explains only a fraction (around 55 percent) of the variation in the data we add normally distributed random variables u^C, u^T, and u^R with mean 0 and a standard deviation equal to the standard deviation of the error term in the respective region. We repeat simulation procedure of Equation (2) for 200 times to simulate 200 nationally representative income samples. Letting $P(\tilde{y})$ be a poverty or inequality measure based on the simulated income distribution, we can generate the conditional distribution of $P(\tilde{y})$, in particular, its mean point estimate and its prediction error, from the 200 samples of simulated incomes. The fit of the simulation can be evaluated by comparing the poverty and inequality measures estimated from observed incomes of the LSMS, $P(y)$, with those from simulated incomes of the DHS, $P(\tilde{y})$.

Third, we choose an earlier period $t-1$ in which the LSMS covers only the cities and partially re-run our regression model

$$y_{t-1}^C = X_{t-1}^C \cdot \beta_{t-1}^C + \varepsilon_{t-1}^C \tag{3}$$

to obtain the coefficient estimates and the standard deviation of the error term for the cities in period $t-1$. As concerns the modelling of dynamics, we assume that the absolute

[4]Unfortunately, the primary sample units (or clusters) of the pre–1997 LSMS are not available in Bolivia so that we cannot split the error term into a spatial and an idiosyncratic component as in Elbers et al. (2003) and Stifel and Christiaensen (2007).

differences in the regression coefficients between cities on the one hand, and towns and rural areas on the other hand, remain constant between period $t-1$ and t, and get the coefficient estimates for towns and rural areas, respectively, in period $t-1$

$$\beta^T_{t-1} = \beta^C_{t-1} + (\beta^T_t - \beta^C_t) \quad \text{and} \quad \beta^R_{t-1} = \beta^C_{t-1} + (\beta^R_t - \beta^C_t). \tag{4}$$

We check the robustness of our results to alternative assumptions on the evolution of the regression coefficients between period $t-1$ and t in Section 4.2.

In a similar vein, we assume that the relative change in the standard deviations of the error terms between period $t-1$ and t is identical for all three regions, and get the standard deviations of the error terms for towns and rural areas, respectively, in period $t-1$

$$\sigma(\varepsilon^T_{t-1}) = \sigma(\varepsilon^C_{t-1}) \cdot \frac{\sigma(\varepsilon^T_t)}{\sigma(\varepsilon^C_t)} \quad \text{and} \quad \sigma(\varepsilon^R_{t-1}) = \sigma(\varepsilon^C_{t-1}) \cdot \frac{\sigma(\varepsilon^R_t)}{\sigma(\varepsilon^C_t)}. \tag{5}$$

Repeating the simulation exercise of Equation (2) with the estimated coefficients from Equations (3)–(5) and the DHS data in period $t-1$, we can create 200 nationally representative samples of simulated incomes in period $t-1$. Again, we can compare the poverty and inequality measures between the two household surveys. In contrast to above, however, this is only possible for the cities where observed incomes are available. After this consistency check, we use the simulated incomes to construct inter-temporally comparable poverty profiles of national coverage for Bolivia and to evaluate the "pro-poorness" of changes of simulated incomes over time using growth incidence curves.

Our data set of LSMS consists of three multi-purpose household surveys conducted by the Instituto Nacional de Estadísticas de Bolivia (National Statistical Office of Bolivia, INE): the 2$^{\text{nd}}$ round (Nov. 1989) and the 7$^{\text{th}}$ round (July to Dec. 1994) of the Encuesta Integrada de Hogares (EIH), and the 1$^{\text{st}}$ round (Nov. 1999) of the Encuesta Continua de Hogares (ECH). The EIH cover only the cities of Bolivia, while the ECH are nationally representative. Two-stage sampling techniques were used in selecting the sample of households, and sampling was done in a way to ensure self-weighting. The purpose of the LSMS is to collect individual, household, and community level data to measure the welfare level of the sampled population and its changes over time. In addition to income and/or expenditure data, the LSMS provide information on demographics, asset ownership, education, employment, and health.

In order to be able to compare our results with earlier empirical studies, we largely use household members as analysis unit. As welfare indicator, we use monthly consumption expenditures (including own consumption, but excluding annualized costs for durable consumer goods) for rural areas, and monthly labor income (excluding fringe benefits)[5] plus monthly capital income for urban areas. The choice of the mixed measurement unit can be justified by that (a) it is common for Bolivia (INE-UDAPE, 2002), (b) an all-expenditure specification is not possible since the EIHs collected only income but no expenditure data, and (c) an all-income specification is not preferable since incomes only poorly reflect the long-term welfare in rural areas due to large seasonal income fluctuations and a high degree of own consumption in agricultural households (Deaton and Zaidi, 2002). In order to account for non-declaration of incomes, we apply a statistical matching approach similar to Hernany (1999). By contrast, we do not adjust for sub-declaration (under-reporting) of incomes (i.e., we do not scale up the mean income and mean consumption expenditures in the LSMS to those in the national accounts) in our baseline scenario because (a) it is a priori not clear whether national account data or LSMS data are more accurate,[6] and (b) Bolivia does not report separate national account data for the cities, towns, and rural areas.[7]

To identify the poor, we use the two sets of poverty lines provided by the Unidad de Análisis de Políticas Sociales y Económicas (UDAPE) (Table 11 in the Appendix). The extreme poverty lines are given by the costs of food baskets which reflect the nutritional requirements of adults and the local eating habits of the middle quintile of the income distribution. The moderate poverty lines additionally include the costs of non-nutritional basic needs and are obtained by multiplying the extreme poverty lines by the inverse of local Engel coefficients. Since no rural poverty lines are available for 1989 and 1994, we extrapolate the relative difference between the rural poverty line and the weighted-average urban poverty line of 1999.

Our set of DHS consists of the first three Bolivian rounds which were conducted in 1989, 1994, and 1998.[8] Two-stage sampling techniques were used to select nationally representative samples of women aged between 15 and 49 who serve as eligible respondents

[5]Only if we exclude fringe benefits the measurement unit is inter-temporally comparable between 1989 and 1999. This is because the EIHs collected, if at all, only the incidence and type of fringe benefits but not their value. As a consequence, our poverty estimates for 1999 are somewhat higher than the official figures provided by INE (various issues).

[6]For a description and evaluation of, and an analysis of the sensitivity of poverty measures to, different adjustment methods, see Székely et al. (2000).

[7]In Section 4, we change this assumption and compare our results with the ones derived from an upscaling exercise using national account data which is available at the departmental level.

[8]The fourth Bolivian DHS round, which was conducted in 2003, is used by Branisa and Grosse (2009) for sensitivity analyses on the robustness of results using other models and error-specifications in the microsimulation, also focussing on the stability of the estimated coefficients and standard deviations.

of the DHS, i.e., women of reproductive age. The main objective of the DHS is to collect demographic data on health and fertility trends. Additionally, it includes some questions on the educational attainment and the employment situation of the respondent and her partner and on the asset ownership of the household.

The covariates taken from the two data sources and their sample means are listed in Tables 12 and 13 in the Appendix. They can be grouped into five categories: information on (a) demographics of the household, (b) asset ownership of the household, (c) educational attainment of adult men and women, (d) employment situation of adult men and women, and (e) health situation of children. By choosing suitable variables and dummy categories, we obtained a high degree of consistency both across surveys and over time.

We build our methodology around the base period 1998/9 and then apply it to the earlier periods 1989 and 1994. Additional data constraints impede our empirical analysis in three respects. First, to create inter-temporally comparable samples of simulated incomes for Bolivia it would be ideal to use a set of covariates which is available in all three pairs of concurrent surveys of 1989, 1994, and 1998/9. At the same time, however, the availability of covariates in the LSMS and the DHS changes over time due to changes in their questionnaires. In order to avoid a too small set of covariates we, thus, decided to use different sets of covariates for each period, i.e., different X enter for each three points in time t, to (a) check the consistency between the LSMS and the DHS data in 1999, (b) to create 200 samples of simulated incomes in the DHS 1989 data, and (c) to create 200 samples of simulated incomes in the DHS 1994 data.[9]

Second, since no Bolivian DHS round was conducted in 1999, we have to use the DHS 1998 data for our consistency check. That is, we compare the poverty and inequality measures based on observed incomes of the LSMS 1999 with those based on simulated incomes of the DHS 1998, assuming that the distribution of the covariates (and also of the returns to covariates) remained reasonably constant in between.[10] By contrast, for 1989 and 1994 we have concurrent rounds of LSMS and DHS. Third, due to its focus on health

[9] To put it more formally, we only require that the set of covariates is identical for the LSMS and the DHS in period $t-1$ as well as for the LSMS in period t. To check for robustness, we also performed our subsequent empirical analysis for the smaller set of common covariates. While, as expected, the consistency check performed worse, the empirical results did not change qualitatively. In Branisa and Grosse (2009), the authors use the smaller set of common covariates for their analysis.

[10] Note that for Ecuador, Hentschel et al. (2000) and Elbers et al. (2003) use the LSMS from 1994 and the Census from 1990, so 4 years of distance of surveys, and assume that distance to be "reasonably" small. The same holds for Stifel and Christiaensen (2007), who face a 1 year difference for the base year. They also apply the same coefficients, similar as Hentschel et al. (2000) and Elbers et al. (2003), for predictions 4 years back and forth in time.

and fertility trends, the DHS data only include households with at least one woman of reproductive age (i.e., eligible women are those aged between 15 and 49). We, thus, have to replicate this implicit sample selection in the LSMS data.[11]

3 Empirical Results

3.1 Estimation Properties

Before comparing poverty and inequality indices of observed, predicted, and simulated incomes, we present some details on the regression results (Table 1) as well as on the properties of the predicted incomes (Tables 2 and 3). Table 1 presents the regression results (β coefficients and P-values) of regressing ln$y$ on the selected variables, run separately for the three regions (city, town, rural). One major concern might be that the simple log-linear OLS regression model is too simple or that the log-normality assumption of incomes does not hold,[12] but we take the above described estimation as a baseline estimation model.[13]

Table 2 shows predicted incomes and the logs ($\hat{y}$ and ln$\hat{y}$) using these regression results in the LSMS data set of 1999 itself by looking at predicted incomes $\hat{y}$ and their logarithm ln$\hat{y}$, and Figure 1 shows the cumulative distribution function and the kernel density estimator for the different income sets, exemplarily for total Bolivia only. What becomes clear is that predicting income without adding an error term gives too low values for $\hat{y}$ but not for ln$\hat{y}$ compared with observed values (y and ln$y$). This problem is due to the log-linear relation between y and ln$y$. By construction, the mean of ln$y$ and ln$\hat{y}$ is the same, even after adding an error term that is normally distributed and has mean 0. However, transforming ln$\hat{y}$ to $\hat{y}$ by taking the anti-log gives exponentially higher values to $\hat{y}$ the higher ln$\hat{y}$ was, so without error terms there are less larger values as in the observed

[11] For 1994 and 1998, but not for 1989, the DHS provide an additional data module on, and responded by, male adults. We opted against using this data module for two reasons: (a) the information was only collected for the husbands and partners of all women included in the main module (but not for men in households with no woman in reproductive age) so that we also would have had to reduce the sample size and possibly would have introduced another sample-selection bias, and (b) our microdata set of simulated incomes would no longer be inter-temporally comparable over the whole observation period.

[12] The visual inspection of the error terms in the three regions show no further signs of heteroskedasticity after using the White (1980) estimator. However, we have tried weighted least squares estimations as well, but the results are very similar, presented in Grosse et al. (2007). Kernel estimates and qqplots show that, besides the extremes, the log-normality assumption seems to hold.

[13] Problems might arise if there were some coefficients that drive the results—i.e., have a high regression coefficient strongly impacting the estimation—but which are insignificant. However, this is not the case. Of the 201 coefficients entering the estimation 120 are insignificant. Despite this being a high number, first of all, of the total 201 coefficients only 5 have a share of more than 10 percent of the total discrepancy of the mean of observed and simulated income in the LSMS compared to the DHS and only 6 coefficients have a share of more than 5 percent. Additionally, of the 120 insignificant coefficients, not a single one has a share of more than 10 or 5 percent of the total discrepancy. Overall, the by far highest coefficient, i.e., share of explanatory power, have the regional dummies for the cities, towns, and rural areas (Table 1).

Table 1: Regression Results, log-linear OLS, 1999

	City		Town		Rural	
	β	P	β	P	β	P
La Paz	0.09	0.39	0.13	0.81	0.19	0.04
Cochabamba	0.28	0.01	0.62	0.22	0.28	0.01
Oruro	0.04	0.75	-0.26	0.65	0.31	0.03
Potosi	0.10	0.45	0.14	0.78	0.04	0.65
Tarija	0.59	0.00	0.37	0.49	0.64	0.00
Santa Cruz	0.68	0.00	0.47	0.35	0.74	0.00
Beni & Pando	0.70	0.00	0.17	0.75	0.81	0.00
# elderly	-0.08	0.60	0.09	0.73	-0.08	0.34
# males	-0.07	0.02	0.10	0.22	-0.10	0.02
# females	-0.12	0.00	-0.10	0.09	-0.17	0.00
# youngsters	-0.03	0.62	-0.08	0.23	-0.01	0.79
# children	-0.11	0.16	-0.18	0.05	-0.08	0.10
# of working age / # all	1.02	0.01	0.22	0.66	0.74	0.01
gender hh head	0.03	0.73	0.25	0.15	-0.02	0.84
language of hh head	-0.01	0.86	-0.12	0.30	-0.06	0.32
hh head age $\leq$ 24	-0.21	0.31	0.01	0.98	0.01	0.98
hh head age 25–34	-0.25	0.22	0.03	0.94	0.05	0.74
hh head age 35–44	-0.39	0.05	0.01	0.99	0.08	0.62
hh head age 45–54	-0.45	0.03	0.13	0.77	-0.04	0.80
hh head age 55–65	-0.34	0.09	0.03	0.94	0.03	0.84
has house	0.07	0.20	-0.07	0.51	0.08	0.25
floor (cement)	0.17	0.21	0.03	0.86	0.24	0.00
floor (brick)	0.30	0.05	0.17	0.33	0.00	1.00
floor (other floor)	0.38	0.01	0.10	0.61	0.24	0.02
2-3 sleeping rooms	0.21	0.00	-0.18	0.11	0.07	0.24
$\geq$ 4 sleeping rooms	0.22	0.04	0.09	0.73	0.30	0.14
access to public water	-0.18	0.11	0.06	0.63	-0.07	0.22
has no toilet	-0.02	0.86	-0.22	0.10	-0.08	0.11
has electricity	-0.32	0.03	-0.19	0.46	0.13	0.05
cooking material	-0.26	0.02	-0.02	0.91	0.30	0.00
has phone	0.24	0.00	0.38	0.00	0.30	0.01
has radio	0.02	0.79	-0.11	0.29	0.10	0.07
has television	0.18	0.10	0.10	0.54	0.23	0.01
has fridge	0.23	0.00	0.03	0.77	-0.02	0.85
no partner in household	0.31	0.15	0.52	0.15	0.38	0.01
com. basic edu. (m.)	-0.12	0.35	-0.01	0.96	0.02	0.78
incom. secondary edu. (m.)	0.04	0.70	-0.20	0.25	-0.04	0.56
com. secondary edu. (m.)	-0.04	0.67	0.11	0.48	-0.02	0.83
tertiary edu. (m.)	0.24	0.03	-0.10	0.66	0.15	0.49
com. basic edu. (w.)	-0.02	0.89	0.04	0.81	0.20	0.00
incom. secondary edu. (w.)	0.05	0.64	0.12	0.41	0.27	0.00
com. secondary edu. (w.)	0.06	0.54	0.11	0.50	0.18	0.08
tertiary edu. (w.)	0.26	0.03	0.27	0.19	0.28	0.17
high skilled white collar (m.)	0.68	0.00	1.09	0.01	0.60	0.00
med. skilled white collar (m.)	0.41	0.03	1.02	0.01	0.45	0.00
skilled manual (m.)	0.44	0.02	0.69	0.07	0.54	0.00
unskilled manual (m.)	0.37	0.08	0.45	0.21	0.45	0.00
agr. employed (m.)	-0.19	0.60	0.47	0.28	0.48	0.00
agr. self-employed (m.)	0.88	0.01	0.07	0.88	0.31	0.01
sales and services (m.)	0.51	0.01	0.94	0.02	0.47	0.00

continued on next page

Table 1 continued

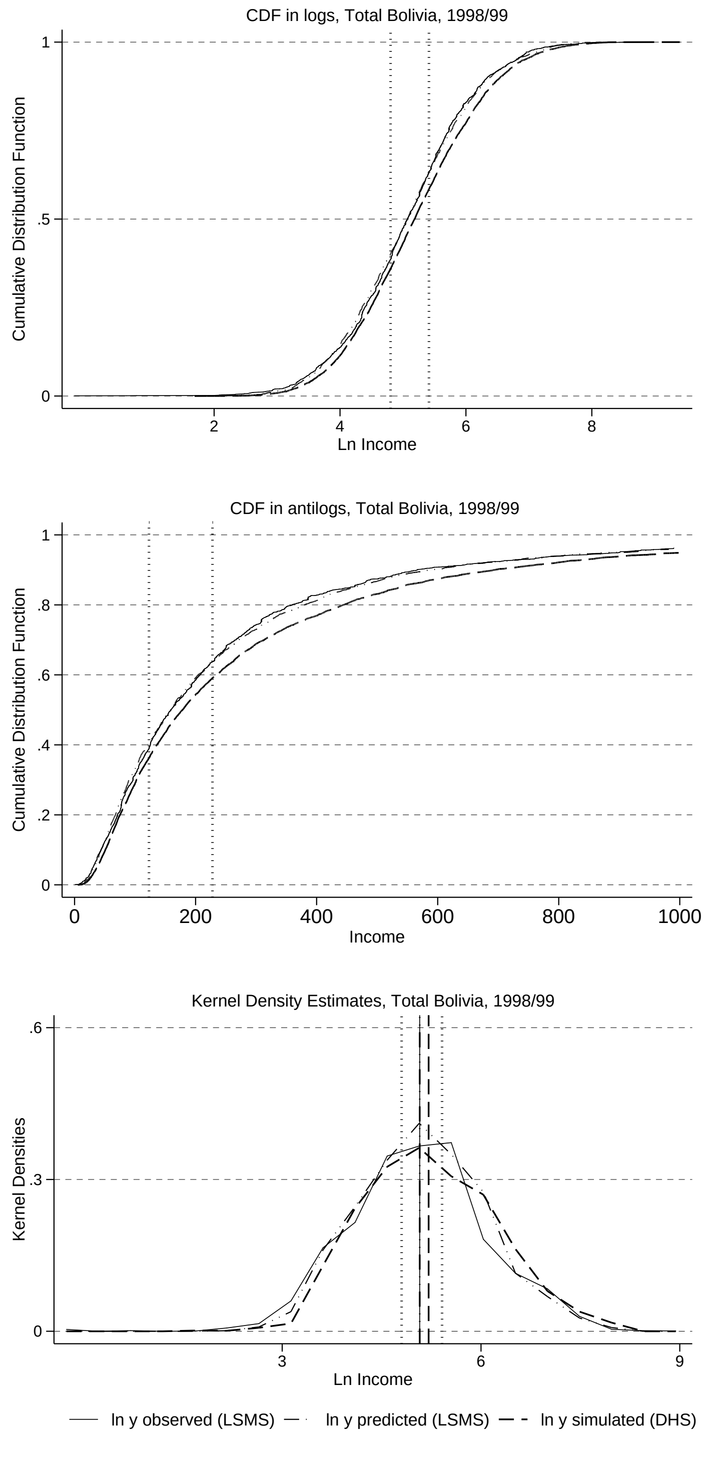
	City		Town		Rural	
	β	P	β	P	β	P
high skilled white collar (w.)	0.35	0.01	0.51	0.02	0.04	0.90
med. skilled white collar (w.)	0.24	0.01	0.77	0.00	0.26	0.07
skilled manual (w.)	0.03	0.78	0.37	0.02	-0.09	0.35
unskilled manual (w.)	0.32	0.00	0.61	0.00	-0.08	0.51
agr. employed (w.)	1.20	0.02	-0.81	0.17	0.07	0.45
agr. self-employed (w.)	0.53	0.00	-0.32	0.33	0.03	0.64
sales and services (w.)	0.30	0.00	0.67	0.00	0.20	0.06
has social security	0.09	0.09	0.08	0.48	0.16	0.05
birth in last 12 month	0.08	0.71	-0.32	0.30	-0.05	0.51
attended by doctor	-0.09	0.72	0.63	0.09	0.11	0.32
delivered in hospital	-0.08	0.64	-0.20	0.37	0.12	0.31
child under 4 years	0.02	0.86	0.14	0.57	0.13	0.29
has first polio vaccination	0.05	0.69	-0.04	0.84	-0.20	0.10
has triple dpt vaccination	0.06	0.61	-0.02	0.91	0.01	0.85
has had diarrhea	-0.14	0.14	0.04	0.79	0.03	0.60
has head cough/fever	0.03	0.67	0.08	0.54	0.02	0.71
c/t/r dummy/constant	4.83	0.00	4.21	0.00	3.79	0.00
# of observations	1037		332		922	
R^2	55.74		58.19		57.11	

Notes: β: regression coefficient, P: P-value. For details on the regression, see text. For details on the variables, see text and notes of Appendix Tables 12 and 13.
Source: Own Calculations.

case. This can be seen in Table 2 for total Bolivia, where the mean of the logarithm for observed $\ln y$, predicted without error $\ln\hat{y}$, and average $\ln\hat{y}$ are nearly exactly the same (columns 6–8), but the means of income for observed y, predicted without error $\hat{y}$, and average $\hat{y}$ are different.

In numbers, the mean of observed income for total Bolivia is at 344 Bolivianos compared to 292 for the prediction when adding no error. The within-sample prediction renders a different picture than the observed income because the prediction does not capture all the variation in the data set. Looking at the average of these 200 repetitions reveals that the mean (of 351 Bolivianos) comes very close to the observed mean y, however, the variation in the data set becomes too low because averaging partly eliminates the variation that had been added with the error term. Rather, when looking at the fourth column "one expl." (which shows the summary statistics of one example, i.e., of one simulated $\hat{y}$) we can see how close we predict incomes compared to observed incomes by looking at means or specific percentiles such as median (P50) or at the extremes of the distribution such as such as of the 5^{th} percentile (P5) or the 95^{th} percentile (P95). In Table 3, we repeat this exercise, and all results are also based on one prediction run (within the LSMS) and one simulation run (over to the DHS data set), but not on the average of the 200 replications.

Figure 1: Kernel Densities and Cumulative Distribution Functions, Total Bolivia, 1999



Notes: The dotted vertical lines mark the weighted poverty lines and the other vertical lines the means for the respective region.
Source : Own Calculations

We detect for cities that the prediction of the mean is better in 1989 and 1994,[14] but less so for 1998/9. For all regions holds that there is a tendency of overprediction with the mean being higher for the simulated data in the DHS compared to the predicted and observed data in the LSMS.

By construction, the mean of observed and predicted incomes in logarithms within the LSMS are the same, but the mean of simulated incomes in logarithms in the DHS is higher. The reason for this overprediction on the national level as well as in each region is the different endowment of the two data sets, i.e., on average higher endowment in the DHS with the covariates that have higher returns to income and lower endowment with those that have lower returns. In addition, the overprediction for the entire country comes from the different geographical allocation of the population (city, town, rural) with the DHS having more people living in cities and fewer living in towns and rural areas (Appendix Tables 12 and 13). When we combine this with the regression coefficient being very high for cities compared to other regression coefficients, we can explain the main part of the difference. Whether or not we over-, well-, or underpredict poverty measures mainly depends on where the poverty line is, as can be seen in the cumulative distribution functions (Figure 1). Interesting to note is that the study of Stifel and Christiaensen (2007) also find an underestimation of the poverty headcount (i.e., overprediction of income) in the DHS 1998 data of 1–2 percentage points which they do not investigate further. Instead, they adjust the poverty line in 1998 in the DHS to match the observed 1997 WMS levels and apply this poverty line back and forth in time.

[14] Even P5 and P95 as well as minima and maxima are relatively well reproduced when taking into account that they are most prone to being outliers or measurement error.

Table 2: LSMS: Observed and Predicted Income and Ln Income, 1999

	y				$\ln y$			
	Obs.	Predicted			Obs.	Predicted		
		no error	aver.	one expl.		no error	aver.	one expl.
	Total Bolivia							
Mean	344.66	292.03	351.29	351.41	5.3333	5.3333	5.3339	5.3411
Min	1.04	15.66	20.02	10.42	0.0408	2.7514	2.8186	2.3434
Max	9515.00	2726.87	3305.85	8218.22	9.1606	7.9109	7.9186	9.0141
P5	39.63	58.81	66.25	41.42	3.6795	4.0743	4.0743	3.7238
P25	104.96	111.24	129.90	100.27	4.6536	4.7117	4.7038	4.6078
P50	205.68	198.98	237.71	206.20	5.3263	5.2932	5.3002	5.3288
P75	398.63	373.66	447.82	412.44	5.9880	5.9233	5.9270	6.0221
P95	1167.32	850.62	1025.99	1185.64	7.0625	6.7460	6.7429	7.0780
SD	459.76	273.78	335.47	464.59	1.0103	0.8236	0.8249	1.0058
SK	5.35	2.53	2.55	4.66	-0.0712	0.1704	0.1769	0.1494
KUR	55.45	13.72	13.79	43.60	3.4922	2.4561	2.4528	2.8036
	City							
Mean	489.84	409.38	496.58	496.82	5.7782	5.7782	5.7769	5.7870
Min	1.04	52.21	63.81	22.22	0.0408	3.9552	3.9923	3.1011
Max	9515.00	2726.87	3305.85	8218.22	9.1606	7.9109	7.9186	9.0141
P5	85.71	114.99	136.82	84.21	4.4510	4.7449	4.7407	4.4334
P25	173.20	189.69	230.13	171.80	5.1544	5.2454	5.2458	5.1463
P50	320.42	311.21	376.09	317.44	5.7696	5.7405	5.7449	5.7603
P75	574.56	540.23	653.92	584.77	6.3536	6.2920	6.2861	6.3712
P95	1425.00	965.86	1162.01	1676.85	7.2619	6.8730	6.8777	7.4247
SD	572.68	313.64	383.81	572.41	0.9196	0.6788	0.6824	0.9012
SK	4.64	2.23	2.25	4.03	-0.2884	0.2175	0.2166	0.1666
KUR	40.78	11.43	11.48	32.69	5.0396	2.4841	2.4833	2.8947
	Town							
Mean	334.25	285.20	347.99	356.51	5.4219	5.4219	5.4276	5.4706
Min	2.38	15.66	20.02	10.42	0.8675	2.7514	2.8186	2.3434
Max	3500.00	1241.63	1599.05	3766.52	8.1605	7.1242	7.1897	8.2339
P5	50.26	70.41	87.22	40.49	3.9172	4.2544	4.2677	3.7011
P25	140.42	141.59	170.77	141.95	4.9446	4.9530	4.9523	4.9555
P50	216.50	230.13	279.36	244.98	5.3776	5.4386	5.4571	5.5012
P75	416.60	354.88	427.71	429.70	6.0321	5.8718	5.8588	6.0631
P95	938.17	723.74	906.44	1071.09	6.8439	6.5844	6.6043	6.9764
SD	346.29	209.03	257.62	375.58	0.9315	0.6924	0.6943	0.9390
SK	3.42	1.82	1.86	3.36	-0.6005	-0.1842	-0.1712	-0.3716
KUR	20.54	6.82	7.12	20.64	4.8209	3.4372	3.4051	3.6134
	Rural Areas							
Mean	145.52	130.36	148.84	145.06	4.6693	4.6693	4.6704	4.6577
Min	10.38	28.67	33.86	13.04	2.3394	3.3558	3.4081	2.5678
Max	1801.06	997.08	1080.70	1408.03	7.4961	6.9048	6.8659	7.2499
P5	31.15	46.87	53.39	33.59	3.4388	3.8474	3.8417	3.5143
P25	60.53	70.14	78.61	60.89	4.1031	4.2505	4.2274	4.1091
P50	104.96	95.93	110.16	100.65	4.6536	4.5636	4.5760	4.6116
P75	181.91	154.50	174.74	174.39	5.2035	5.0402	5.0414	5.1613
P95	384.38	339.65	388.67	399.43	5.9516	5.8279	5.8170	5.9900
SD	139.71	103.04	117.52	142.50	0.7837	0.5923	0.5929	0.7729
SK	4.16	2.90	2.82	3.16	0.0535	0.7246	0.7322	0.3041
KUR	34.76	14.84	13.68	17.83	2.9429	3.3683	3.3707	2.9578

Notes: P: percentile, SD: standard deviation, SK: skewness, KUR: kurtosis, y: nominal income (city, town) and consumption (rural), aver.: average over 200 y, one expl.: one simulated y.
Source: Own Calculations.

Table 3: Income Properties from LSMS and DHS, 1989, 1994, and 1998/9

	LSMS 1989		DHS 1989	LSMS 1994		DHS 1994	LSMS 1999		DHS 1998
	Observed	Predicted	Simulated	Observed	Predicted	Simulated	Observed	Predicted	Simulated
Total Bolivia									
Mean Ln y	n.a.	n.a.	4.0888	n.a.	n.a.	4.4687	5.3333	5.3411	5.4703
Mean y	n.a.	n.a.	103.43	n.a.	n.a.	188.99	344.66	351.41	402.47
Min y	n.a.	n.a.	1.64	n.a.	n.a.	1.19	1.04	10.42	7.02
Max y	n.a.	n.a.	1793.15	n.a.	n.a.	4263.73	9515.00	8218.22	15610.24
P5 y	n.a.	n.a.	11.84	n.a.	n.a.	6.70	39.63	41.42	47.65
P50 y	n.a.	n.a.	57.34	n.a.	n.a.	109.97	205.68	206.20	230.71
P95 y	n.a.	n.a.	348.18	n.a.	n.a.	626.49	1167.32	1185.64	1324.23
SD y	n.a.	n.a.	139.34	n.a.	n.a.	265.70	459.76	464.59	530.38
City									
Mean Ln y	4.5630	4.5569	4.5860	5.2907	5.2929	5.3262	5.7782	5.7870	5.8800
Mean y	150.74	146.72	150.29	295.67	290.16	288.81	489.84	496.82	552.84
Min y	1.67	4.56	5.30	4.17	14.86	18.93	1.04	22.22	18.61
Max y	3884.94	3276.48	1793.15	7035.00	4708.07	4263.73	9515.00	8218.22	15610.24
P5 y	23.17	22.82	22.55	55.52	54.53	57.34	85.71	84.21	80.86
P50 y	92.17	92.45	97.99	188.51	187.09	198.96	320.42	317.44	352.52
P95 y	448.30	451.49	450.82	873.71	862.65	794.28	1425.00	1676.85	1693.05
SD y	216.07	183.25	169.70	369.92	325.21	292.29	572.68	572.41	634.44
Town									
Mean Ln y	n.a.	n.a.	4.0012	n.a.	n.a.	4.7134	5.4219	5.4706	5.4087
Mean y	n.a.	n.a.	94.67	n.a.	n.a.	211.74	334.25	356.51	359.09
Min y	n.a.	n.a.	2.44	n.a.	n.a.	1.51	2.38	10.42	7.02
Max y	n.a.	n.a.	1384.52	n.a.	n.a.	4128.74	3500.00	3766.52	4732.90
P5 y	n.a.	n.a.	9.19	n.a.	n.a.	14.49	50.26	40.49	41.43
P50 y	n.a.	n.a.	51.38	n.a.	n.a.	121.27	216.50	244.98	225.55
P95 y	n.a.	n.a.	284.99	n.a.	n.a.	690.79	938.17	1071.09	1052.25
SD y	n.a.	n.a.	130.98	n.a.	n.a.	348.60	346.29	375.58	432.48
Rural Areas									
Mean Ln y	n.a.	n.a.	3.5389	n.a.	n.a.	3.3661	4.6693	4.6577	4.8153
Mean y	n.a.	n.a.	51.74	n.a.	n.a.	62.37	145.52	145.06	171.40
Min y	n.a.	n.a.	1.64	n.a.	n.a.	1.19	10.38	13.04	12.24
Max y	n.a.	n.a.	756.14	n.a.	n.a.	1333.87	1801.06	1408.03	1942.00
P5 y	n.a.	n.a.	9.18	n.a.	n.a.	4.38	31.15	33.59	35.87
P50 y	n.a.	n.a.	32.95	n.a.	n.a.	27.23	104.96	100.65	120.01
P95 y	n.a.	n.a.	164.74	n.a.	n.a.	246.87	384.38	399.43	485.68
SD y	n.a.	n.a.	63.15	n.a.	n.a.	96.67	139.71	142.50	173.19

Notes: P: percentile, SD: standard deviation, y: nominal income (city, town) and consumption (rural). Simulated and Predicted is one simulated y.
Source: Own calculations.

In Table 4, we provide four sets of moderate poverty: (a) their point estimates from observed incomes of all households in the LSMS (column All HH), (b) their point estimates from observed incomes of households with at least one woman of reproductive age in the LSMS (column Sample), (c) their mean point estimates and standard deviation from 200 samples of predicted incomes in the LSMS (column Pred.), and (d) their mean point estimates and standard deviation from 200 samples of simulated incomes in the DHS (column Sim.). Results for extreme poverty and for inequality are shown in Appendix Tables 14 and 15. Note that, different from above, mean point estimates mean that we estimate the poverty and inequality indicators based on the 200 examples of predicted and simulated incomes and over them calculate the average of 200 poverty and inequality estimates. That is, poverty or inequality measures are not calculated using the mean income of the 200 prediction or simulation examples. Standard deviations for poverty estimates are very low and for inequality estimates even lower. This translates into ranges of about ±2 percentage points for P0, ±1 for P1, and even less for P2. The same magnitudes hold for the decimal places of inequality measures.

Taking differences between these columns enables us to decompose the overall difference between observed and simulated poverty and inequality measures into three components related to (a–b) the implicit sample selection of only households with at least one women in reproductive age, (b–c) the specification of the error term in the underlying regression model, and (c–d) differences in the distribution of the covariates between LSMS and DHS.

For 1989 and 1994, for which the consistency check is limited to the cities, the results are very encouraging, as they had also been for the income properties in Table 3. For 1999, the situation is somewhat less favorable. Restricting the sample to households with at least one eligible woman does not induce a serious bias in estimating poverty and inequality measures. Poverty indices are slightly higher and inequality indices slightly lower when comparing the first with the second column. Adding a normally distributed error term to create 200 samples of predicted incomes in the LSMS only slightly understates the poverty headcount P0 and slightly overstates the poverty gap P1 and the squared poverty gap P2. It also only slightly understates income inequality as evidenced by lower values of the Gini coefficient and the Atkinson indices in 1989 and 1994 and slightly overstates them in 1998.

Table 4: Poverty Indices Based on Observed, Predicted, and Simulated Incomes, 1989, 1994, 1998/9

	1989				1994				1998/9			
	LSMS			DHS	LSMS			DHS	LSMS			DHS
	All HH	Sample	Predicted	Simulated	All HH	Sample	Predicted	Simulated	All HH	Sample	Predicted	Simulated
Moderate Poverty Line												
Total Bolivia												
P^0	n.a.	n.a.	n.a.	76.10	n.a.	n.a.	n.a.	72.44	63.52	65.18	65.08	60.33
				(0.53)				(0.42)			(0.93)	(0.46)
P^1	n.a.	n.a.	n.a.	44.45	n.a.	n.a.	n.a.	45.28	31.53	32.45	33.63	30.04
				(0.35)				(0.22)			(0.57)	(0.24)
P^2	n.a.	n.a.	n.a.	30.48	n.a.	n.a.	n.a.	33.95	19.48	20.11	21.19	18.50
				(0.31)				(0.19)			(0.46)	(0.18)
City												
P^0	65.92	67.07	65.08	64.84	58.09	59.56	58.06	57.36	48.39	50.97	50.67	47.99
			(0.80)	(0.91)			(0.59)	(0.73)			(1.60)	(0.72)
P^1	31.96	32.64	32.79	32.92	25.15	25.87	25.89	25.26	19.75	20.90	22.47	21.22
			(0.48)	(0.53)			(0.33)	(0.39)			(0.86)	(0.37)
P^2	19.18	19.64	20.35	20.55	13.91	14.31	14.65	14.17	10.80	11.46	12.82	12.12
			(0.38)	(0.41)			(0.24)	(0.29)			(0.62)	(0.27)
Town												
P^0	n.a.	n.a.	n.a.	80.21	n.a.	n.a.	n.a.	73.42	66.60	69.03	67.49	64.26
				(1.26)				(1.16)			(2.48)	(1.18)
P^1	n.a.	n.a.	n.a.	49.66	n.a.	n.a.	n.a.	43.40	32.99	34.58	34.90	33.67
				(0.87)				(0.64)			(1.48)	(0.66)
P^2	n.a.	n.a.	n.a.	35.58	n.a.	n.a.	n.a.	30.66	19.94	20.97	22.21	21.76
				(0.79)				(0.55)			(1.22)	(0.53)
Rural Areas												
P^0	n.a.	n.a.	n.a.	87.96	n.a.	n.a.	n.a.	90.23	81.64	83.37	84.24	79.11
				(0.70)				(0.43)			(1.02)	(0.63)
P^1	n.a.	n.a.	n.a.	56.35	n.a.	n.a.	n.a.	69.86	46.02	47.71	48.74	43.11
				(0.53)				(0.28)			(0.90)	(0.40)
P^2	n.a.	n.a.	n.a.	40.54	n.a.	n.a.	n.a.	58.66	30.39	31.85	32.48	27.66
				(0.50)				(0.28)			(0.82)	(0.34)

Notes: Results for the extreme poverty line are shown in Appendix Table 14. Poverty indices are calculated using income data for departmental capitals and other urban areas, expenditure data for rural areas, and mixed income-expenditure data for total Bolivia. Standard deviations of the poverty indices in brackets (only applicable to those based on predicted and simulated incomes).
Source: Own Calculations.

The transition from LSMS data to DHS data does, as mentioned, reduce the poverty and does increase inequality measures, due to the better endowment in the DHS compared to the LSMS data sets, especially in 1998/9. In total, the underestimation of the poverty headcount is about 5 percentage points. Most of the underprediction is driven by rural areas (with the headcount being 5 percentage points lower) but also for cities and towns with the headcount also being 2 to 3 percentage points lower. For the extreme poverty line, the underprediction is less severe for cities and towns, but even worse for rural areas. In total, an additional problem is that the share of people living in (richer) cities is higher and of towns and especially rural areas is lower (Appendix Tables 12 and 13). The underlying economic reason of the underprediction is most probably the lack of consistency with respect to the collection period of the two underlying household surveys. The DHS 1998 data, the covariates of which were used to create the simulated incomes, were collected during an economic boom. By contrast, the observed incomes of the LSMS 1999 were collected after a sharp economic downturn when Bolivia experienced strongly negative growth in GDP per capita.[15]

These inconsistencies notwithstanding, we are confident that the conditions for applying our dynamic cross-survey microsimulation methodology are fulfilled for the case of Bolivia. First, the simulations can accurately reproduce the observed poverty trends in the cities, where we have observed incomes for comparison. The differences between observed and simulated poverty measures are small compared to their changes over time. Second, the DHS 1998 data, which are least consistent to those of the corresponding LSMS, are not used in the subsequent poverty and inequality analysis. Only the poverty profiles and growth incidence curves for 1989 and 1994 draw on simulated incomes of the DHS. Those for 1999 are based on observed incomes of the LSMS.

3.2 Poverty and Inequality Trends

To extend our illustration, we provide some analyses of poverty between 1989 and 1999.[16] We start our empirical analysis with a disaggregation of the poverty headcount by place of residence and household characteristics in Table 5. Between 1989 and 1999, total Bolivia experienced a significant decrease in the incidence of poverty. Moderate poverty decreased from three-quarters to two-thirds of the population. The reduction in extreme poverty

[15] Another reason besides the one year difference might be the relatively small size of the LSMS 1999. Furthermore, inofficial judgement of this data set is that it is of lower-than-average quality.

[16] For results on pro-poor growth using in addition the 4$^{\text{th}}$ round (Nov. 2002) of the Encuesta Continua de Hogares (ECH), see Grosse et al. (2007).

was even more spectacular; it decreased from 55 to less than 40 percent.[17]

As expected, rural households were more likely to be poor than those in the cities and towns even after controlling for local cost-of-living differences. What is more of concern here is that rural households did not fully participate in the reduction of moderate poverty between 1989 and 1999. The cities and towns could reduce the incidence of moderate poverty by 16 and 11 percentage points, respectively. In rural areas, this reduction was only 4 percentage points—despite starting from a higher level of poverty.[18] Furthermore, poverty in rural areas increased in between 1989 and 1994, quite contrary to the trends in cities and towns.[19] Taken together, the poverty trends suggest that rural areas were quite detached from improvements and deteriorations in the overall economic environment.

In Section 2, we assumed that the absolute difference in the regression coefficients between the cities on the one hand, and towns and rural areas on the other hand, remained constant between 1989 and 1999. If this assumption does not hold, i.e., if the coefficients in rural areas deteriorated relative to those in urban areas, the decline in poverty in rural areas shown in the subsequent analysis would be overstated. We address this potential bias in Section 4. Another factor that may contribute to overstating the decline in poverty—albeit in this case not limited to rural areas—is that the degree of underreporting, which is common to all income and expenditure surveys, may have fallen over time due to improvements in the questionnaire design.[20] Taken together, we, thus, treat the reduction in poverty as an upper bound, particularly so in rural areas.[21]

There are also substantial differences in the incidence of poverty across the nine departments of Bolivia. The moderate poverty headcount in 1989 ranged from 60 percent in Santa Cruz to 91 percent in Potosí. The corresponding figures for the extreme poverty headcount were 33 percent and 82 percent, respectively. The departmental distribution of the poverty headcount index was also very stable in Bolivia. While Santa Cruz, which is a major host of commercial agriculture and food-processing industry, had the lowest

[17] In the late 1990s, the poverty trend reversed and the incidence of moderate and extreme poverty in total Bolivia started to increase again (Grosse et al., 2007).

[18] That is, in relative terms, the performance of rural areas was even worse. As concerns extreme poverty, rural areas also experienced the lowest absolute (!) reduction of the poverty headcount index between 1989 and 1999.

[19] By contrast, households in the cities were most affected by the economic downturn in the late 1990s, leading to an increase of moderate and extreme poverty in total Bolivia between 1999 and 2002 (Grosse et al., 2007).

[20] Of course, and that is what the evidence mainly suggests, the degree of underreporting might have risen over time. Taking our data for Bolivia, underreporting seems to have fallen from 1989 to 1999, see Chapter 4.1 and especially Table 7 where the ratio of household survey to national accounts mean increases from 0.7 to 0.8 (LSMS) or even 0.9 (DHS) over time.

[21] For a literature overview of other studies on poverty in Bolivia, see Spatz (2006).

Table 5: Headcount, Spatial Disaggregation and Profile, by Income

	Moderate Poverty Line			Extreme Poverty Line		
	1989	1994	1998/9	1989	1994	1998/9
Total	76.10	72.44	65.21	54.92	51.99	38.35
	(0.53)	(0.42)		(0.62)	(0.40)	
By Region						
City	67.07	59.56	51.05	39.11	28.90	24.22
Town	80.21	73.42	69.09	59.43	50.97	34.31
	(1.26)	(1.16)		(1.44)	(1.14)	
Rural	87.96	90.23	83.37	71.87	80.85	59.98
	(0.70)	(0.43)		(0.92)	(0.47)	
By Department						
Chuquisaca	87.41	85.87	84.15	71.76	73.31	64.34
	(0.97)	(0.97)		(1.28)	(1.06)	
La Paz	77.73	69.96	68.55	55.90	48.59	46.33
	(1.07)	(0.82)		(1.22)	(0.89)	
Cochabamba	73.21	75.50	64.69	50.64	53.69	31.70
	(1.19)	(1.10)		(1.48)	(1.20)	
Oruro	82.13	81.35	68.64	63.33	65.46	47.63
	(1.16)	(1.19)		(1.41)	(1.27)	
Potosí	91.44	87.90	84.66	82.05	79.62	63.01
	(0.85)	(0.91)		(1.14)	(0.99)	
Tarija	81.26	81.49	61.68	60.00	58.95	26.39
	(1.18)	(1.12)		(1.46)	(1.19)	
Santa Cruz	60.30	57.20	50.59	33.28	30.79	21.66
	(1.22)	(1.10)		(1.38)	(0.90)	
Beni & Pando	78.43	77.95	53.00	54.83	55.49	14.73
	(1.16)	(1.32)		(1.48)	(1.59)	
By household size						
≤ 3	70.94	62.24	47.35	46.99	40.02	22.02
	(1.29)	(0.95)		(1.55)	(0.86)	
4–6	73.46	71.62	61.01	51.45	50.64	34.28
	(0.79)	(0.63)		(0.86)	(0.58)	
≥ 7	84.54	83.51	80.35	66.77	65.85	52.61
	(0.82)	(0.75)		(1.03)	(0.85)	
By percent of household members between 15 and 65 years						
≤ 50	82.31	81.52	74.93	63.00	62.00	48.79
	(0.65)	(0.54)		(0.80)	(0.56)	
> 50	67.59	60.90	53.64	43.86	39.27	25.91
	(0.82)	(0.64)		(0.95)	(0.56)	
By age of household head						
≤ 34	78.25	73.77	67.29	56.64	51.22	39.02
	(0.88)	(0.70)		(1.05)	(0.81)	
35–49	76.07	73.23	66.97	55.44	53.75	40.43
	(0.84)	(0.64)		(0.95)	(0.60)	
50–65	74.01	68.18	57.86	52.33	48.91	31.56
	(1.18)	(1.09)		(1.32)	(0.97)	
≥ 66	70.73	70.80	63.66	49.79	54.38	39.13
	(2.25)	(1.85)		(2.26)	(1.70)	
By language of household head						
Spanish	70.10	63.34	51.27	46.16	38.08	22.27
	(0.67)	(0.55)		(0.71)	(0.53)	
Indigenous	93.27	93.72	79.75	80.01	84.51	55.11
	(0.71)	(0.49)		(1.12)	(0.65)	

continued on next page

Table 5 continued

	Moderate Poverty Line			Extreme Poverty Line		
	1989	1994	1998/9	1989	1994	1998/9
By gender of household head						
Male	76.67	73.14	65.64	55.89	53.06	38.82
	(0.56)	(0.47)		(0.67)	(0.45)	
Female	73.17	69.07	62.82	49.98	46.83	35.73
	(1.49)	(1.11)		(1.63)	(1.10)	
By average years of schooling of adults						
≤ 5	89.70	89.20	86.04	72.49	75.63	61.53
	(0.60)	(0.51)		(0.92)	(0.60)	
6–12	68.50	67.56	63.14	42.10	40.78	32.01
	(0.97)	(0.70)		(1.01)	(0.68)	
≥ 13	33.82	28.92	20.11	13.41	10.19	4.65
	(1.94)	(1.47)		(1.45)	(1.03)	
By profession of principal wage earner						
White-Collar	49.47	37.30	33.84	26.49	16.18	14.82
Worker	(1.48)	(1.25)		(1.32)	(0.97)	
Blue-Collar	78.15	74.04	69.23	53.41	46.40	30.80
Worker	(1.03)	(0.85)		(1.22)	(0.97)	
Agriculture	92.53	94.15	88.11	79.45	87.69	65.56
	(0.68)	(0.38)		(1.07)	(0.50)	
Sales and	68.63	63.43	53.30	42.42	34.37	29.74
Services	(1.43)	(1.30)		(1.57)	(1.19)	
Not	80.61	72.86	53.82	58.31	46.66	32.02
Employed	(1.42)	(1.66)		(1.77)	(1.66)	
By percent of adult women in employment						
> 0	59.22	70.36	63.95	34.95	51.90	37.27
	(1.08)	(0.50)		(1.09)	(0.47)	
0	83.33	76.80	67.95	63.48	52.18	40.69
	(0.55)	(0.68)		(0.73)	(0.88)	

Notes: Poverty indices are calculated using mixed income-expenditure data. Standard deviations of the poverty indices in brackets (only applicable to those based on simulated data).—For the category schooling: Adult women aged between 15 and 49 and their husbands and partners.—For the category wage earner: In the case of DHS: Husband or partner of the oldest woman aged between 15 and 49. If she is single, this women herself. In case of LSMS. Household head.—For the category female employment: Women aged between 15 and 49.
Source: Own Calculations.

incidence of poverty throughout the entire observation period, it was highest in Potosí, followed by Chuquisaca, which are particularly dependent on subsistence agriculture.

When looking at household characteristics, the mayor determinants of poverty is household size with poverty increasing in line with increasing numbers of family members. The higher the share of working-aged members to overall members is, the lower is poverty. The relation of the age of the household head and poverty follows a u-shaped trend with the cohort of 50–65 years old being the ones with the lowest poverty incidence. Clearly to be seen is that indigenous households are much poorer than Spanish-speaking ones. As observed in several studies for Latin American countries (Marcoux, 1998), households with a female head seem to be less poor than those with a male head. Increasing education has a very strong poverty-decreasing effect. The same holds for the sector of employment of the principal wage earner where high-skilled professionals have very much lower poverty incidence than other groups. Working in agriculture is correlated with the highest poverty incidence. Female participation in the labor force reduces poverty.

3.3 Pro-Poor Growth

To evaluate whether the simulated income changes over time were "pro-poor" in the sense that the poor benefited more from economic growth than the rich, we apply the methodology of growth incidence curves (GIC) developed by Ravallion and Chen (2003). Comparing two periods, $t-1$ and t, the growth incidence curve plots the cumulative share of the population (depicted on the x-axis) against the income growth rate of the p^{th} quantile (depicted on the y-axis) when the population quantiles are ranked in ascending order of their income. It is given by

$$GIC := g_t(p) = \frac{y_t(p)}{y_{t-1}(p)} - 1 = \frac{\mu_t}{\mu_{t-1}} \cdot \frac{L'_t(p)}{L'_{t-1}(p)} - 1, \tag{6}$$

where $L'(p)$ is the slope of the Lorenz curve at the p^{th} quantile, and μ is mean income. It can be shown that the area under the GIC up to the poverty headcount index P^0 gives (minus one times) the rate of change of the Watts index over time

$$-\frac{dW_t}{dt} = \int_0^{P_t^0} \frac{d \log y_t(p)}{dt} \cdot dp = \int_0^{P_t^0} g_t(p) \cdot dp. \tag{7}$$

The desirable axiomatic properties of the Watts index motivate evaluating the "pro-poorness" of economic growth[22] by comparing the growth rate in mean income (GRIM) with the mean of the income growth rates of the poor in period $t-1$ which Ravallion and Chen (2003) coined the "pro-poor growth rate" (PPGR):

$$PPGR := \frac{1}{P_{t-1}^0} \cdot \int_0^{P_{t-1}^0} g_t(p) \cdot dp. \tag{8}$$

The comparison of the growth rates is shown in Table 6. Between 1989 and 1999, economic growth in Bolivia can be classified as pro-poor following the baseline scenario (first column labeled "base"). For both poverty lines and for all three regions, the PPGR exceeded the GRIM suggesting that economic growth was accompanied by falling inequality (Figure 2). For all regions, the income distribution of 1999 even first-order dominates the income distribution of 1989 as evidenced by that the GIC lies above 0 for all p,[23] except

[22] Alternative approaches of measuring pro-poor growth can be found in Klasen (2004) and Son (2003).

[23] For some regions only the first percentile shows a negative growth rate. This, however, is mainly a statistical problem since the results are sensitive to outliers which are likely to be found at the tails of the distribution.

for rural areas for which this condition is met at least for all poor. That is, abstracting from individual income mobility across quantiles, the welfare of all citizens in the cities and of all poor citizens in the rest of the country improved during the 1990s.[24]

Table 6: Annual Average Income Growth per Capita, 1989 to 1999

	1989–1998/9						
	base	a.dum	a.na	con01	div01	con05	div05
	Total Bolivia						
GRIM	2.16	1.61	0.80	2.01	1.92	2.10	1.65
PPGR mod.	2.91	1.86	1.14	2.65	2.12	3.76	1.09
PPGR extr.	3.05	1.85	1.19	2.79	2.10	4.25	0.82
	City						
GRIM	2.01	2.01	1.89	2.01	2.01	2.01	2.01
PPGR mod.	2.53	2.53	2.56	2.53	2.53	2.53	2.53
PPGR extr.	2.48	2.47	2.51	2.48	2.47	2.48	2.48
	Town						
GRIM	2.85	2.34	1.12	2.74	2.54	2.87	1.86
PPGR mod.	5.25	4.61	2.47	5.39	4.45	7.40	2.68
PPGR extr.	5.87	5.19	2.85	6.09	4.95	8.55	2.87
	Rural						
GRIM	0.46	-1.43	-1.26	-0.26	-0.56	0.09	-1.42
PPGR mod.	1.86	-0.05	0.10	1.51	0.53	3.56	-1.31
PPGR extr.	1.95	0.01	0.30	1.64	0.55	3.99	-1.48

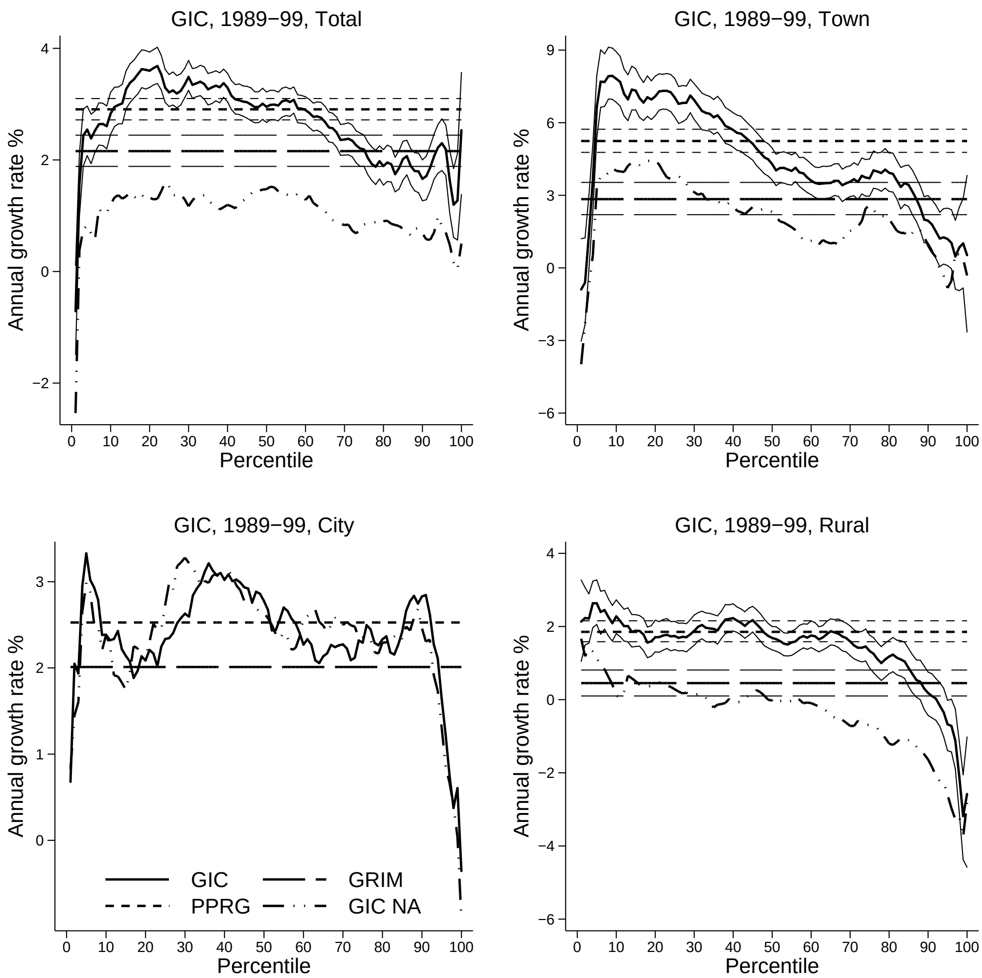
Notes: Annual average income growth rates are calculated using income data for cities and towns, expenditure data for rural areas, and mixed income-expenditure data for total Bolivia. For 1989, only the data for cities can be taken from the LSMS. All other growth rates are calculated using the DHS of 1989. The adjustment procedures are explained in Chapter 4. GRIM: growth rate in mean; PPGR (mod. and extr.): (moderate and extreme) pro-poor growth rate; base: baseline scenario; a.dum: adjustment of regional dummies; a.na: adjustment to national accounts; con01 (con05): convergence scenario(s) with range of $\phi = \pm 0.1(0.5)$; div01 (div05): divergence scenario(s) with range of $\phi = \pm 0.1(0.5)$.
Source: Own Calculations.

With the exception of the strongly anti-poor contraction in the cities in recent years, economic growth in Bolivia has been pro-poor since 1989, and particularly so in rural areas. This result seems to be at odds with Table 5 which shows only slowly falling poverty rates in rural areas since 1989. However, this puzzle resolves when taking into account that the depth of poverty in rural areas is so large that even substantial pro-poor growth did not lift the poor above the poverty line.[25] Hence, the prime concern is not that economic growth in the 1990s was anti-poor, but that it was so low and that the initial income inequality was so high that the poor remained poor despite some welfare improvements. For a country with such unfavorable initial conditions it would take another decade of such economic growth to make serious inroads into poverty.

[24] For results on pro-poor growth between 1999 and 2002, see Grosse et al. (2007).

[25] But it did reduce the poverty gap in rural areas.

Figure 2: Growth Incidence Curves



Notes: The 90% confidence intervals for GIC, GRIM, and PPGR are calculated using the 200 simulation runs. GIC: growth incidence curve, GRIM: growth rate in mean, PPGR: pro-poor growth rate (moderate poverty line). GIC NA: based on the adjustment to departmental national accounts as described in Section 4.1. For GIC NA, no CI are shown for better visuality of the graphs.
Source : Own Calculations.

4 Sensitivity Analyses

4.1 Disaggregated Data on GDP

One basic problem with the simulated data is that there are hardly any possibilities to cross-check the results with any other data source. National accounts are one option but, as mentioned before, not available for the urban-rural divide. The only data available is GDP per capita from the national accounts at the departmental level. To get an idea, however, about the plausibility of our results, we compare national account data with the results from the LSMS and DHS household surveys. Furthermore, we try to impute national account information also for the divide of cities, towns, and rural areas (Section 4.2).

The national account series available to us is compared to household survey data in Table 7 in the upper part ("original data"). Note that the household surveys are not meant to be representative at this level, but for a first check, it generates some intuition for the problems of the data. Furthermore, as mentioned above, it is not a priory clear if household survey data is inferior in quality compared to national account data. What becomes clear from the table is that, as expected, household survey data shows lower values compared to national account data. What also becomes obvious is that this difference is not stable over time and that it is not the same for all departments. For total Bolivia, the relation between DHS and national account data is 0.72 in 1989, goes down to 0.68 in 1994 and increases to 0.94 in 1998. Especially the latter value is pretty high, also compared to the value of 0.81 for the LSMS of 1999.

For the departments, the relation is in between 0.42 up to values close to 1. For the DHS, some values are even above 1. Obviously, there seem to be some differences between the three data sources. This becomes especially clear when looking at the ranking of departments and the difference of this ranking between household surveys and national accounts. There are 2 or 3 departments for which our simulated and observed data differ strongly from the national account data. First, La Paz appears to be richer when looking at household surveys compared to national accounts. The difference in ranking is very high, for example in 1994, La Paz is the third poorest department looking at national accounts but the second richest looking at DHS data. Another extreme case is Oruro, which is throughout the whole decade the second or third richest department on national account data but the third poorest on household survey data. Furthermore, the different dynamics of Beni and Pando cannot be taken into account correctly since their values cannot be

Table 7: Subnational Income from NA, LSMS, and DHS, 1989–1999

	1989		1994		1998/9		
	NA	DHS	NA	DHS	NA	DHS	LSMS
	Original data						
Total Bolivia	305.52	218.82	309.27	209.00	330.35	302.13	266.43
City		312.14		322.99		411.40	378.84
Town		196.56		210.10		273.21	258.58
Rural		117.22		71.91		133.91	112.37
Chuquisaca	280.72	133.85	252.11	124.59	299.99	186.84	145.10
La Paz	267.80	201.41	288.58	224.20	279.61	252.71	256.45
Cochabamba	318.05	241.03	329.06	195.40	359.50	319.48	261.88
Oruro	345.82	160.59	338.63	142.69	417.89	231.05	195.39
Potosi	196.05	101.77	164.93	90.26	191.82	162.05	127.02
Tarija	322.38	192.70	315.67	161.50	382.38	389.66	253.53
Santa Cruz	401.39	347.92	394.49	309.97	399.03	441.79	376.54
Beni	316.41	213.77	310.96	185.34	329.42	319.80	338.42
Pando	345.72	213.77	375.18	185.34	500.80	319.80	338.42
	Adjusted data						
Total Bolivia		299.48		307.97		330.56	328.19
City	385.24	420.62	418.78	462.97	442.74	453.48	461.19
Town	325.61	279.37	351.54	360.68	359.49	289.96	315.67
Rural	216.77	165.07	239.76	106.15	273.25	144.04	147.09
Chuquisaca		269.59		240.32		292.82	299.99
La Paz		256.86		285.43		273.98	279.61
Cochabamba		315.39		336.69		356.82	359.50
Oruro		344.08		352.26		426.69	417.89
Potosi		197.65		170.91		185.87	191.82
Tarija		317.36		338.91		375.80	382.38
Santa Cruz		407.70		383.50		423.22	399.03
Beni		336.15		295.12		347.35	329.42
Pando		336.15		295.12		347.35	329.42

Notes: Monthly per capita income, in constant Bolivianos (1995). Beni and Pando are not separated in the LSMS and DHS questionnaires, so the values hold for both departments. National accounts (NA) are not imputed for city/town/rural for the 1998 data, instead values of 1999 are shown.
Source: Own calculations.

separated in the household surveys. Pando seems to be richer and also more dynamic than Beni. However, both departments account for only less than 1 percent (Pando) and 4.5 percent (Beni) of the total population. For the other departments, our simulation is pretty close to the national accounts as concerns the ranking. Another general difference is, as mentioned above, that the DHS simulation for 1998 is higher in nearly all departments compared to the LSMS data of 1999 (except for Beni). The only strong difference between LSMS and DHS in ranking of departments is Tarija for which the DHS shows the second highest value and for the LSMS only the sixth. When looking at national accounts and DHS from earlier years, the lower rank seems to be more plausible, i.e., in the middle of the distribution rather than one of the richest departments. The overall poorest department according to all three data sets and showing hardly any growth is Potosi.

For a first sensitivity analysis, we simply adjust the LSMS and DHS data to the national accounts, however done at the level of the departments rather than to the overall national income mean (as normally done in the literature). Adjusting to the departmental level might be slightly less problematic than to overall national accounts because it takes some region-specific income dynamics and differences into account, but doubts remain about the correspondence of national accounts and participation of private households in GDP (Stifel and Christiaensen, 2007), and this is also true at the departmental level, but maybe to a lesser extent. Results on pro-poor growth of this exercise can be found in Table 6, column "a.na" (third column, the abbreviation stands for adjustment to national accounts) as well as in Figure 2. Growth remains pro-poor, however the growth rates are becoming smaller because the distance of households surveys to national accounts was wider in the earlier years, so closing this distance automatically shrinks the growth rates.[26]

4.2 Regional Differentials in Sectoral Participation

The differences in our results compared to national accounts motivate us to conduct one further sensitivity analyses with this data at hand since we want to focus somewhat more strongly on the urban-rural divide. For this, we use sectoral GDP and employment shares in sectors to break down the data to the urban-rural divide.

We have made a rather simple calculation to break down the data to the urban-rural divide. The data for the 3 points in time available is (i) sectoral GDP $y_{s_1,s_2,\dots s_n}$ (from the

[26]For the time from 1999 to 2002, household surveys more strongly underestimated the value compared to the national accounts, so that the negative growth during this time span would turn positive using the adjustment to national accounts (results not shown in the table).

national accounts), (ii) population shares in cities, towns, and rural areas $p_{c,t,r}$ (from the three DHS rounds and from the 1999 LSMS), and (iii) employment shares by sector of the population for all three regions $e_{s_{c,t,r}}$ (from the LSMS data, only available for 1999). So if we impute the per capita income in, for example, cities, we make the following simple calculation: $y_c = \frac{\sum_s y_s \cdot e_{s_c}}{p_c}$. This means to calculate the sum over all sectoral GDP per capita multiplied with total employment in the sector in cities, and dividing by total population in cities.

What becomes evident from Table 7, lower part showing "adjusted data", is that the relation of national account data to household survey data higher in cities than it is in towns and way higher than in rural areas when comparing the original data in the upper part of Table 7 with the adjusted data for national accounts in the lower part. This problem holds for LSMS as well as DHS data. We address this problem in this and the following section. One of the basic assumptions of our dynamic cross-survey microsimulation methodology is that the absolute difference in the regression coefficients between the cities on the one hand and towns and rural areas on the other hand remained constant between 1989 and 1999. We present two ways to model the additional relative changes in returns to covariates in which the constancy assumption is relaxed. Explicitly testing these modelling exercises is only possible using data from the DHS 2003.[27]

The first very simple way does the following: The constancy-of-differences assumption of the basic model implies that the widening of the urban-rural divide during that time is, thus, entirely attributed (a) to changes in the endowment of covariates favoring urban areas, and (b) to nationwide changes in the return to covariates favoring those covariates which are relatively abundant in urban areas. If this assumption does not hold, i.e., if additionally (c) the returns to covariates in rural areas deteriorated relative to those in urban areas, the widening of the urban-rural divide would be understated. To get an idea of the possible size of this bias we have to simulate the opposite scenario where we assume that the widening of the urban–rural divide between 1989 and 1999 is entirely due to deteriorating returns to covariates in rural areas relative to those in urban areas. Adjusting Equation (4) leads to:

$$\beta^R_{t-1} = \beta^C_{t-1} + (\beta^R_t - \beta^C_t) + Adj_{growth} \tag{9}$$

where Adj_{growth} stands for the adjustment of the growth differential. Since it is a priori not clear which covariates are affected and to what extent, we take a rather simple approach

[27] This is done in Branisa and Grosse (2009).

and attribute the regional growth differentials in GDP per capita to growth differentials in the regression coefficients of the regional dummies.

This sensitivity analysis proceeds in three steps. First, we impute the 1989–to–1994 and the 1994–to–1999 cumulative growth differentials in GDP per capita between the cities on the one hand and towns and rural areas on the other hand.[28] We find that the economic growth performance was nearly identical across the three regions in the first half of observation period, but differed substantially thereafter.

Between 1989 and 1994, the cities (cumulatively) grew by only 0.1 and 0.2 percent faster than towns and rural areas, respectively. The corresponding figures for the period from 1994 to 1999 are about 2 and 9 percent, respectively. Second, we sterilize the growth differentials in GDP per capita by adding for towns and for rural areas the 1994–to–1999 growth differential in GDP per capita (relative to the cities) to the 1994 regression coefficient of the corresponding regional dummy, and sum of the 1989–to–1994 and the 1994–to–1999 growth differential in GDP per capita (relative to the cities) to the 1989 regression coefficient of the corresponding regional dummy. Third, we partially re-run our simulation with the adjusted coefficients to generate an adjusted spatial disaggregation of pro-poor growth in Bolivia (Table 6, second column "a.dum", which stands for adjustment via dummy correction).

Comparing the results with the corresponding entries of the baseline scenario in Table 6, column "base", reveals that the bias of neglecting a possible deterioration of the returns to covariates in rural areas relative to those in urban areas is evident when applying this simple way of modelling changes in relative returns. Sterilizing the regional growth differentials in GDP per capita decreases income in rural areas and less so in towns in 1989 compared to the baseline estimation, so that the GRIM and PPGR are lower. Due to lower growth in rural areas and towns, overall (mean) growth in Bolivia is now smaller between 1989 and 1999, and the growth is also less pro-poor as the rate of growth in rural areas, whose population predominates among the poor, is now estimated to have been lower. But the qualitative results from above do not change: We find that growth and pro-poor growth are somewhat smaller in total Bolivia and more significantly so in rural areas which even experienced negative mean income growth between 1989 and 1999; but

[28]We impute, as explained above, the separate growth rates of GDP per capita for the cities, towns, and rural areas by multiplying for each economic sector the average annual growth rate of value added per capita over the respective period (taken from the national accounts) by the employment shares of those sectors in the cities, towns, and rural areas, respectively (estimated from the LSMS 1999). Note that this is a constancy assumption as well. Here, employment shares do not change over time.

the PPGR remain higher, however very small, suggesting that the poor were able to make only few gains over the period.

4.3 Mobility Assumption

In the second sensitivity analysis for relaxing the assumption of constancy of the distance between urban and non-urban areas we do not make a priory assumptions about the changes in relative returns to covariates, but we generate a "mobility" scenario around the baseline scenario.[29] We again recall the constancy assumption in Equation (4) and rearrange it in the following way:

$$\beta^R_{t-1} = \beta^C_{t-1} + (\beta^R_t - \beta^C_t) \quad \Longrightarrow \quad (\beta^R_{t-1} - \beta^C_{t-1}) = \phi(\beta^R_t - \beta^C_t) \tag{10}$$

where ϕ is the "mobility parameter". In our baseline scenario, ϕ is equal to 1, thus absolute changes of the coefficients remain constant between the regions, here exemplarily only for the cities versus rural areas.

As an illustration, let us assume that we observe a coefficient β_E for secondary education of $\beta^R_{E,t} = 0.4$ and $\beta^C_{E,t} = 0.9$, which leads to an absolute difference of –0.5 in t, and that we observe $\beta^C_{E,t-1} = 0.8$ for $t-1$. What we have done in the baseline regression was to assume "no mobility" in the sense that the absolute difference stays constant over time which would be fulfilled for a coefficient of $\beta^R_{E,t-1} = 0.3$. If we assume that the difference decreases over time (which would for example be fulfilled for $\beta^R_{E,t-1} = 0.2$), we think of this as "mobility" (in the sense of converging or becoming more similar over time). This leads to a greater absolute difference of –0.6 in $t-1$, and the mobility parameter takes a value $\phi = 1.2 > 1$. If we assume the opposite (in the sense of divergence or dissimilarity) the absolute difference has to increase, from for example –0.2 in $t-1$ for $\beta^R_{E,t-1} = 0.6$ which leads to $\phi = 0.4 < 1$. As mentioned above, there is no way to know the exact structure of change of relative returns over time, especially because this change will be different of magnitude and even sign for each coefficient.

The results on moderate poverty of this exercise are shown in Table 8, for extreme poverty and inequality in Appendix Tables 16 and 17. It is only relevant for towns and rural areas in 1989 and 1994 (and also for the aggregate data for total Bolivia), and for comparison, the baseline scenario (no mobility) is copied from Table 4 and Appendix Tables 14 and 15. The general conclusion is that the results for poverty and inequality are pretty stable. We present two different assumptions, one of a weak mobility scenario of $\phi =$

[29] We thank Martin Ravallion for this suggestion.

± 0.1 and of a strong mobility scenario of $\phi = \pm 0.5$. The weak scenario generates a mobility band around the point estimate of about 1 percentage point or even less for all poverty measures. This holds for both poverty lines and for both years. Of course, the strong mobility scenario results in a broader band, and differences get larger in 1994, especially for P1 and P2. The deviations are not symmetric which is caused by the above explained non-linear relation between y and $\ln y$. Looking at inequality, the results are similar. Again, for the weak mobility scenario, the inequality indicators assuming no mobility do not differ too much from the mobility results. However, results are more sensitive to the strong assumption and also to the more sensitive Atkinson indices, especially to A(2.0).

In summary, the results are stable and convincing. Even with the stronger assumption of $\phi = \pm 0.5$ and the more sensitive indicators (P2 and A(2.0)), the tends in poverty and inequality remain. The same holds for the results on pro-poor growth. In Table 6, the columns labelled "con01–div05" show the results. As expected the "convergence" scenarios give stronger evidence of pro-poor growth and the "divergence" scenarios give lower growth rates compared to the baseline assumptions.

Table 8: Moderate Poverty: Convergence and Divergence Assumptions

	1989					1994				
	Convergence		Divergence		No mobility	Convergence		Divergence		No mobility
	ϕ=1.1	ϕ=1.5	ϕ=0.9	ϕ=0.5	ϕ=1.0	ϕ=1.1	ϕ=1.5	ϕ=0.9	ϕ=0.5	ϕ=1.0
Moderate Poverty Line										
Total Bolivia										
P^0	75.35	75.63	75.08	73.98	76.10	72.24	72.40	72.07	71.28	72.44
	(0.47)	(0.47)	(0.47)	(0.49)	(0.53)	(0.43)	(0.42)	(0.43)	(0.45)	(0.42)
P^1	43.68	45.75	42.49	39.81	44.45	46.22	48.37	44.75	40.57	45.28
	(0.33)	(0.32)	(0.34)	(0.35)	(0.35)	(0.24)	(0.24)	(0.25)	(0.26)	(0.22)
P^2	29.80	32.39	28.42	25.60	30.48	35.25	38.41	33.18	27.75	33.95
	(0.29)	(0.29)	(0.30)	(0.30)	(0.31)	(0.19)	(0.19)	(0.20)	(0.21)	(0.19)
Town										
P^0	79.74	80.44	79.17	77.17	80.21	73.44	73.60	73.20	71.91	73.42
	(1.24)	(1.19)	(1.32)	(1.40)	(1.26)	(1.28)	(1.18)	(1.35)	(1.42)	(1.16)
P^1	49.72	53.12	47.81	43.53	49.66	44.27	47.34	42.47	38.19	43.40
	(0.92)	(0.86)	(0.95)	(0.99)	(0.87)	(0.79)	(0.75)	(0.81)	(0.83)	(0.64)
P^2	35.85	40.02	33.63	29.06	35.58	31.78	35.71	29.58	24.69	30.66
	(0.82)	(0.79)	(0.83)	(0.84)	(0.79)	(0.63)	(0.61)	(0.63)	(0.64)	(0.55)
Rural Areas										
P^0	86.49	86.96	85.97	83.86	87.96	90.34	90.68	90.00	88.40	90.23
	(0.62)	(0.57)	(0.65)	(0.79)	(0.70)	(0.44)	(0.39)	(0.48)	(0.54)	(0.43)
P^1	54.90	59.01	52.54	47.22	56.35	71.35	75.82	68.21	59.06	69.86
	(0.51)	(0.46)	(0.53)	(0.58)	(0.53)	(0.26)	(0.23)	(0.28)	(0.34)	(0.28)
P^2	39.31	44.45	36.57	30.96	40.54	60.82	67.53	56.30	44.18	58.66
	(0.51)	(0.48)	(0.52)	(0.52)	(0.50)	(0.25)	(0.22)	(0.27)	(0.31)	(0.28)

Notes: See text for explanation. Results for the extreme poverty line are shown in the Appendix Table 16.
Source: Own Calculations.

5 The Asset Index Approach

The asset-index approach to construct national time series of basic poverty measures goes back to Filmer and Pritchett (2001) and Sahn and Stifel (2000, 2003). To proxy welfare in the absence of income or expenditure data, they assume that the asset ownership of households closely reflects their living standard. Using DHS data, we define a set of assets[30] and construct a metric asset index

$$AI_j = \frac{s_1(a_{j1} - \overline{a}_1)}{\sigma_1} + \ldots + \frac{s_k(a_{jk} - \overline{a}_k)}{\sigma_k} \tag{11}$$

where s_k is the "scoring factor" or the weight of the asset k, a_{jk} takes the value of 1 if household j owns asset k and 0 otherwise, $\overline{a}_k$ is the mean value of a_{jk} over all households, and σ_k is its standard deviation.

Following Filmer and Pritchett (2001), we use the principal component analysis (rather than the closely related factor analysis as in Sahn and Stifel (2000, 2003)) to determine the asset weights s_k. The underlying idea is to find a linear combination of the variables—the principal component or the asset index—which contains most of the common information of the variables and can be interpreted as a background variable contained in all of them.[31] Hence, the asset-index approach is valid if welfare is indeed the main determinant of asset variability among households. We apply the asset-index approach to track the evolution of poverty between period $t-1$ and t. Since the mean value of the asset index is zero by construction, we do not estimate Equation (11) for each period separately but over a pooled sample of the periods $t-1$ and t.

In contrast to our dynamic cross-survey microsimulation methodology, the creation of national poverty profiles on the basis of the asset index requires a common set of assets for all observation years. Unfortunately, there was a change in the DHS questionnaire design: the DHS 1994 and 1998 collected information on more and other assets than the DHS 1989.[32] The set of common assets over all Bolivian DHS rounds would have been very small so that we decided to restrict our empirical analysis to the years 1994–1998. The derivation of the asset index and the summary statistics of the assets included therein are

[30] Our asset definition is rather broad and includes not only real estate and financial assets, but also consumer durables and the household's endowment with human capital.

[31] A more recent method to construct asset indices is proposed by Kolenikov and Angeles (2009) and applied by Cardozo and Grosse (2009) for Colombia. The innovation is that it is possible to include categorical variables (with different parameter values) rather than only dummy variables. For example, Cardozo and Grosse (2009) include four or five different categories for wall and floor material rather than just a dummy for good and bad material.

[32] The lack of consistency applies especially to consumer durables (Table 13).

shown in Table 9. We use 25 assets—17 tangible assets and 8 human capital variables—to capture the welfare of households.[33] The eigenvalues of the principal component analysis suggest that the asset index is indeed an important determinant for the asset distribution among households. The first principal component explains about 22 percent of total asset variability.

Table 9: The Derivation of the Asset Index, 1994 and 1998

	pooled	1994			1998		
	s_k	$\bar{a}_k$	σ_k	$\frac{\sigma_k}{s_k}$	$\bar{a}_k$	σ_k	$\frac{\sigma_k}{s_k}$
Tangible Assets							
Telephone	0.254	0.106	0.308	0.826	0.250	0.433	0.587
Radio	0.180	0.852	0.355	0.508	0.881	0.324	0.557
Television	0.351	0.582	0.493	0.711	0.684	0.465	0.755
Fridge	0.285	0.297	0.457	0.625	0.377	0.485	0.589
House	-0.109	0.671	0.470	-0.233	0.650	0.477	-0.229
Family Land	-0.299	0.285	0.451	-0.662	0.213	0.409	-0.730
Electricity	0.342	0.676	0.468	0.731	0.757	0.429	0.798
Public Water	0.307	0.561	0.496	0.618	0.698	0.459	0.668
Other Water Source	-0.084	0.143	0.350	-0.239	0.109	0.312	-0.268
Cooking Material	0.335	0.641	0.480	0.699	0.718	0.450	0.745
Shared Toilet	-0.002	0.358	0.480	-0.005	0.194	0.396	-0.006
Private Toilet	0.243	0.240	0.427	0.570	0.483	0.500	0.487
Cement Floor	0.098	0.326	0.469	0.209	0.376	0.484	0.202
Brick Floor	0.055	0.117	0.322	0.171	0.076	0.265	0.208
Other (Non-earth) Floor	0.197	0.180	0.384	0.511	0.260	0.439	0.448
2-3 Sleeping Rooms	0.102	0.411	0.492	0.208	0.346	0.476	0.215
$\geq$ 4 Sleeping Rooms	0.113	0.057	0.232	0.487	0.062	0.240	0.470
Human Capital							
% of Men with							
Complete Basic	-0.084	0.119	0.321	-0.261	0.095	0.290	-0.289
Lower Secondary	-0.033	0.136	0.341	-0.098	0.115	0.316	-0.106
Higher Secondary	0.092	0.242	0.425	0.215	0.235	0.420	0.218
Tertiary Education	0.193	0.107	0.307	0.629	0.156	0.360	0.536
% of Women with							
Complete Basic	-0.075	0.125	0.315	-0.238	0.101	0.287	-0.261
Lower Secondary	-0.012	0.137	0.326	-0.036	0.133	0.317	-0.037
Higher Secondary	0.198	0.254	0.410	0.483	0.301	0.427	0.464
Tertiary Education	0.185	0.080	0.255	0.726	0.139	0.325	0.570
Asset Index		-0.371	2.281		0.383	2.317	

Notes: For the explanation of the variables, see Table 12. The left-out categories are: open water source, no toilet, earth floor, 0–1 sleeping rooms, no or incomplete basic schooling.
Source: Own Calculations.

Since all tangible assets are dummy variables, their scoring factors have a simple interpretation. Moving from "non-ownership" to "ownership" of one asset changes the asset index by $\frac{s_k}{\sigma_k}$. For example, having private telephone connection increases the asset index by 0.83 in 1994 and 0.59 in 1998.[34] In the case of the human capital variables, $\frac{s_k}{\sigma_k}$ gives the change in the asset index if the average education of adult household members switches from the reference state "less than complete basic schooling or unknown" to the respective

[33]To check the robustness of our empirical results, we also estimated the asset index without human capital variables. The empirical results (not shown here) do not change qualitatively.

[34]The reduction in the asset weight reflects the fact that private telephone connection has become more affordable and, thus, more widespread in Bolivia (Table 13).

schooling category.

As expected, consumer durables, such as telephone, radio, television, and fridge, have high scoring factors suggesting that they are powerful welfare predictors. By contrast, owning a house or of a plot of agricultural land indicates poverty which can mainly be explained by the widespread subsistence agriculture in rural areas of Bolivia. The quality of the dwelling also reflects the welfare of households. Access to public utilities, high-quality cooking materials, high quality toilet facilities, high-quality floor materials, and a large number of sleeping rooms all increase the asset index. The scoring factors of the human capital variables are more difficult to reconcile. We find negative returns to schooling up to lower secondary schooling (9 years of schooling)[35] which we attribute to that (a) our reference state includes "unknown" and that (b) the returns to basic and secondary schooling are indeed very small in Bolivia.

The asset-index value of the individual household is obtained by multiplying the deviation of the households asset endowment from the mean asset endowment with the vector of normalized scoring factors according to Equation (11). Aggregating the asset-index values over all households, we find the mean asset index increasing from -0.37 in 1994 to 0.38 in 1998, suggesting a favorable trend of the living standard in Bolivia in the observation period. Based on the estimates of the asset-index values at household level, we can check the consistency of poverty trends between our dynamic cross-survey microsimulation methodology and the asset-index approach.[36] We construct poverty profiles based on asset-index values and compare them to those in Section 3.2. To this end, we rank the households according to their asset-index values and calibrate the thresholds (i.e., poverty lines) between extremely poor, moderately poor, and non-poor so as to ensure that the incidence of poverty at the aggregated national level (i.e., in the first row of the poverty profile) in 1994 coincides with the one of the dynamic cross-survey microsimulation methodology, which is shown in Table 5.[37] We keep this threshold level for the asset-index poverty line of 1994 constant and apply it also to the 1998 data. The spatial poverty profile based on asset-index values is shown in Table 10.

[35] Comparing the results with the results for 1994–2003, we find a switching sign for lower secondary schooling for women which is negative for the period 1994–1998 but turned positive for the period including 2003 (results not shown here).

[36] When we rank the households according to (a) their simulated incomes and (b) their asset-index values and calculate the Spearman rank correlation coefficient between the two welfare indicators we find a close relationship between the simulated incomes and the asset-index values. The Spearman rank correlation coefficient is about 0.8.

[37] The distribution of the assets among extremely poor, moderately poor and non-poor is given in Appendix Table 18.

Although the direction of change and determinants are qualitatively similar to our findings using the microsimulation methodology, there are some differences. The most striking difference between the asset index and the microsimulation methodology is that overall poverty reduction from 1994 to 1998 appears much stronger using the asset index. Keeping the threshold of 1994 constant yields a 5.1 percentage points higher poverty reduction using the moderate poverty line and 2.0 percentage points using the extreme poverty line compared to the results shown in Table 5. We suspect that this sharper reduction in poverty using the asset index is due to a combination of changes in preferences favoring some assets (e.g., televisions), relative price reductions of some assets (e.g., telephones), and public investment in infrastructure or education which have not (yet) translated into income gains. Thus, the sharper poverty reduction using the asset index says more about developments in preferences and in non-income dimensions of well-being than being the most reliable proxy for the income dimension.

Furthermore, taking the corresponding results of the dynamic cross-survey microsimulation methodology in Table 5 as reference point, we find that the asset-index approach strongly underpredicts poverty in the cities and towns and strongly overpredicts poverty in rural areas. In doing so, the results of the asset-index approach are closer to those of the unsatisfied-basic-needs approach[38] than those of the dynamic cross-survey microsimulation methodology. Additionally, not only the level but also the change in the incidence of poverty is more unevenly distributed across the three regions. While according to the dynamic cross-survey microsimulation methodology rural areas participated—albeit less than proportionately—in the overall poverty reduction, they experienced nearly no progress in reducing poverty according to the asset-index approach. These differences are partly due to that only the dynamic cross-survey microsimulation methodology accounts for differences in the local price levels (Table 11); they also show that progress in improving the asset base in rural areas have been much slower in the 1990s.

By contrast, Table 10 shows less variation in the incidence of poverty across departments. The 1994 moderate poverty headcount index ranged only from 66 percent in Santa Cruz and Tarija to 84 percent in Potosí. For comparison, the corresponding figures of

[38]The unsatisfied-basic-needs approach is very similar to the asset-index approach. It generates a weighted average of welfare indicators (e.g., educational attainment, housing quality, access to public utilities, and access to basic health services, in the case of Bolivia) and classifies households as poor if their weighted average indicator value is below a certain threshold. In contrast to the asset-index approach, the indicator weights are set arbitrarily. For a more detailed description of the unsatisfied-basic-needs approach and its application to Bolivia, see Hernany (1999).

Table 10: Poverty Profile, by Asset-Index, 1994–1998

	Moderate Poverty Line		Extreme Poverty Line	
	1994	1998	1994	1998
Total	72.57	60.07	50.45	36.92
By Region				
City	52.20	38.59	19.91	9.54
Town	70.03	57.25	35.27	23.59
Rural	97.76	97.14	91.66	88.55
By Department				
Chuquisaca	79.42	70.54	69.39	57.82
La Paz	71.45	60.97	47.43	33.62
Cochabamba	75.78	56.71	57.21	37.91
Oruro	72.65	60.55	41.30	29.66
Potosí	84.57	76.77	68.01	55.75
Tarija	67.88	54.86	45.48	35.02
Santa Cruz	63.60	50.88	37.71	26.38
Beni & Pando	81.82	69.41	62.86	50.06
By household size				
≤ 3	73.32	63.01	49.48	35.70
4–6	69.29	56.44	46.71	33.97
≥ 7	79.22	66.10	59.69	45.31
By percent of household members between 15 and 65 years				
≤ 50	79.49	71.52	58.25	47.72
> 50	63.77	47.53	40.54	25.07
By age of household head				
≤ 34	77.19	70.99	52.64	40.42
35–49	72.42	57.47	50.64	36.43
50–65	65.92	49.59	46.46	32.18
≥ 66	61.65	47.29	46.17	34.43
By language of household head				
Spanish	61.37	49.93	33.18	23.21
Indigenous	98.74	97.01	90.82	86.83
By gender of household head				
Male	73.09	61.04	51.42	38.22
Female	70.05	55.49	45.77	30.74
By average years of schooling of adults				
≤ 5	97.27	93.82	83.87	73.45
6–12	64.62	50.85	32.10	21.58
≥ 13	9.60	9.37	1.58	1.96
By percent of adult women in employment				
> 0	72.83	56.28	51.66	33.89
0	72.02	66.88	47.91	42.36

continued on next page

Table 10 continued

	Moderate Poverty Line		Extreme Poverty Line	
	1994	1998	1994	1998
By profession of principal wage earner				
White-collar admin.	27.57	18.39	10.90	6.82
Blue-collar admin.	80.65	68.47	46.27	29.28
Agriculture	98.89	96.75	94.52	91.20
Sales and services	64.27	48.85	29.66	15.92
Not employed / DK	54.01	44.13	26.93	19.94

Notes: Poverty indices are calculated using mixed income-expenditure data. Standard deviations of the poverty indices in brackets (only applicable to those based on simulated data).—For the category schooling: Adult women aged between 15 and 49 and their husbands and partners.—For the category wage earner: In the case of DHS: Husband or partner of the oldest woman aged between 15 and 49. If she is single, this women herself. In case of LSMS. Household head.—For the category female employment: Women aged between 15 and 49.
Source: Own Calculations.

the dynamic cross-survey microsimulation methodology were 58 percent and 88 percent, respectively. As concerns the departmental poverty ranking, we find greater consistency between the two approaches.[39] Santa Cruz is the richest department and Potosí and Chuquisaca are the poorest departments. The notable exception is Oruro which is relatively poor according to the dynamic cross-survey microsimulation methodology but relatively rich according to the asset-index approach. Another exception are Beni and Pando which are relatively rich according to the microsimulation but relatively poor according to the asset index.[40] As concerns household characteristics, some differences are observed compared to the income poverty profiles. For example are medium-sized households the richest compared to smaller or bigger ones. Furthermore, also the "oldest" households are the richest. However, this might be due to the fact that older households accumulate assets over time which constantly lose value but remain as an item in the household, irrespective of their value. Some characteristics are even more strongly indicating poverty, such as ethnicity, gender, schooling, or employment sector.

6 Discussion

In this paper, we developed a new methodology to create a national income time series out of incomplete income or expenditure data, and applied it to the case of Bolivia between 1989 and 1999. We show that our extension of the poverty mapping methodology is able to reproduce trends in differential in poverty well where we have comparable data. It also

[39]This result becomes even more obvious when we compare the departmental disaggregation of the poverty headcount by quintiles rather than only at the thresholds between extremely poor, moderately poor, and non-poor (results are not reported here).

[40]For more detailed poverty maps also at regional levels, see Spatz (2006).

appears superior to the use of asset indices for measuring trends in poverty which might more reflect changes in preferences, prices, and non-income indicators. As such it is of considerable use for situations where nationally representative income surveys are lacking, but DHS data are available. With this method it should be possible to generate longer time series of poverty and inequality than is currently possible for most Latin American and many African countries.

Further research should address the questions on how to judge the goodness of fit of the methodology by statistical procedures. The methodology presented here is based on the data constraint of having only one nationally representative pair of different household surveys (one having and the other not having income in the survey), and to have some urban income surveys for other years together with some national-wide other survey. Having a second pair of full surveys allows a backward and forward check of the approach described, in the sense of an out-of-sample prediction that can be compared to observed data.[41]

Our methodology is based on the idea that changes over time should be explicitly modelled. What is normally applied in the literature is to neglect dynamics. For example, the study of Stifel and Christiaensen (2007) uses a static prediction procedure for the regression coefficients and also tries to use variables that are "likely to remain stable over time", i.e., that are not sensitive to "economic or polity change" (Stifel and Christiaensen (2007), p. 323). However, this makes poverty trends over time somewhat slow: if regression coefficients are constant and variables are chosen to be nearly constant then changes are hardly to be observed. In this regard, such results hardly reflect income poverty but is much closer to looking at asset poverty (as measured by asset indices).

[41]As done in Branisa and Grosse (2009) for Bolivia using LSMS data from 2002 and DHS data from 2003 or in Mathiassen (2008) using several income surveys for Uganda.

References

Branisa, B. and M. Grosse (2009). Out of sample prediction of poverty mapping. Mimeo. University of Göttingen.

Cardozo, A. and M. Grosse (2009). Pro-poor growth using non-income indicators: An empirical illustration for Colombia. Ibero-America Institute for Economic Research Discussion Paper (Courant Research Centre 'Poverty, Equity and Growth' Discussion Paper) 192 (9), University of Göttingen.

Datt, G. and M. Ravallion (1992). Growth and redistribution components of changes in poverty measures: A decomposition with applications to Brazil and India in the 1980s. *Journal of Development Economics 38*(2), 275–295.

Deaton, A. and S. Zaidi (2002). Guidelines for constructing consumption aggregates for welfare analysis. LSMS Working Paper 135, World Bank.

Elbers, C., J. O. Lanjouw, and P. Lanjouw (2003). Micro-level estimation of poverty and inequality. *Econometrica 71*(1), 355–364.

Filmer, D. and L. Pritchett (2001). Estimating wealth effects without expenditure data – or tears: An application to educational enrollments in states of India. *Demography 38*(1), 115–132.

Grosse, M., S. Klasen, and J. Spatz (2005). Creating national poverty profiles and growth incidence curves with incomplete income or consumption expenditure data: An application to Bolivia. Ibero-America Institute for Economic Research Discussion Paper 129, University of Göttingen.

Grosse, M., S. Klasen, and J. Spatz (2007). Matching household surveys with DHS data to create nationally representative time series of poverty: An application to Bolivia. Centre for Statistics (ZfS) Discussion Papers 08.2007, University of Göttingen.

Hentschel, J., J. O. Lanjouw, P. Lanjouw, and J. Poggi (2000). Combining census and survey data to trace the spatial dimensions of poverty: A case study of Ecuador. *The World Bank Economic Review 14*(1), 147–165.

Hernany, W. (1999). La pobreza en el Área urbana de Bolivia: Período 1989-1997. Tesis para la obtención de la licenciatura en economía, Universidad Católica Boliviana, La Paz, Bolivia.

INE (various issues). Información estadística. Online data base. Instituto Nacional de Estadística. www.ine.gov.bo.

INE-UDAPE (2002). Mapa de pobreza 2001. Instituto Nacional de Estadística and Unidad de Análisis de Políticas Sociales y Económicas.

Klasen, S. (2004). In search of the holy grail: How to achieve pro-poor growth? In B. Tungodden, N. Stern, and I. Kolstad (Eds.), *Towards Pro-poor Policies – Aid, Institutions, and Globalization* (Proceedings from the Annual World Bank Conference on Development Economics – Europe 2003 ed.). Washington, D.C.: World Bank.

Klasen, S., M. Grosse, R. Thiele, J. Lay, J. Spatz, and M. Wiebelt (2007). Analyzing pro-poor growth in Bolivia: Addressing data gaps and modeling policy choices. In S. Klasen, M. Grimm, and A. McKay (Eds.), *Determinants of Pro-Poor Growth: Analytical Issues and Findings from Country Cases.* New York: Palgrave Macmillan.

Kolenikov, S. and G. Angeles (2009). Socioeconomic status measurement with discrete proxy variables: Is principal component analysis a reliable answer? *Review of Income and Wealth 55*(1), 128–165.

Marcoux, A. (1998). The feminization of poverty: Claims, facts, and data needs. *Population and Development Review 24*(1), 131–139.

Mathiassen, A. (2008). The predictive ability of poverty models. Emprical evidence from Uganda. Discussion Paper 560, Statistics Norway, Division for Development Cooperation.

Ravallion, M. and S. Chen (2003). Measuring pro-poor growth. *Economics Letters 78*(1), 93–99.

Sahn, D. E. and D. Stifel (2000). Poverty comparisons over time and across countries in Africa. *World Development 28*(12), 2123–2155.

Sahn, D. E. and D. Stifel (2003). Exploring alternative measures of welfare in the absence of expenditure data. *Review of Income and Wealth 49*(4), 463–489.

Son, H. H. (2003). Approaches to defining and measuring pro-poor growth. Mimeo. World Bank, Washington D.C.

Spatz, J. (2006). *Poverty and inequality in the era of structural reforms: The case of Bolivia*. Berlin: Springer. Kieler Studien 336.

Stifel, D. and L. Christiaensen (2007). Tracking poverty over time in the absence of comparable consumption data. *The World Bank Economic Review 2*(21), 317–341.

Székely, M., N. Lustig, M. Cumpa, and J. A. Mejía (2000). Do we know how much poverty there is? IDB Working Paper 437, Inter-American Development Bank.

White, H. (1980). A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica 48*(4), 817–838.

Appendix

Table 11: Poverty Lines for Bolivia for Various Years

	Moderate Poverty Line					Extreme Poverty Line				
	1989[a]	1994	1998	2002	2002[c]	1989[a]	1994	1998	2002	2002[c]
Urban Areas										
Chuquisaca	138.5	241.8	335.4	335.6	395.5	73.3	127.9	169.4	169.5	209.2
La Paz (City)	135.3	227.9	324.0	327.0	383.3	75.2	126.6	180.1	181.8	214.6
La Paz (El Alto)	116.6	192.6	270.4	272.6	332.9	70.7	116.7	164.1	165.5	201.8
Cochabamba	142.1	253.2	351.1	351.3	405.8	71.8	127.6	177.3	177.4	204.9
Oruro	123.0	207.1	294.7	297.4	351.1	75.2	126.6	163.9	165.3	214.6
Potosí	113.1	190.5	271.0	273.5	323.0	75.2	126.6	150.7	152.1	214.6
Tarija	144.3	257.3	356.8	351.3	412.1	71.8	127.9	178.6	177.4	204.9
Santa Cruz	141.8	237.8	354.7	343.9	404.8	72.0	120.7	180.2	174.7	205.5
Beni	141.8	237.8	354.7	343.9	404.8	72.0	120.7	180.2	174.7	205.5
Pando	141.8	237.8	354.7	343.9	404.8	72.0	120.7	180.2	174.7	205.5
Urban population weighted average	135.4	231.7	344.7	344.3	392.9	73.4	124.8	176.4	175.5	208.9
Rural Areas	96.9[b]	164.4[b]	233.6	233.4	276.6	55.2[b]	93.4[b]	131.2	133.0	157.6
Population weighted average	119.5	204.8	299.3	298.1	351.2	65.9	112.3	160.6	160.3	190.5

Notes: Numbers in current Bolivianos. [a]Since no poverty lines are available for the 2nd round (Nov. 1989) of the EIH, they are constructed as the arithmetic mean of the poverty lines for the 1st round (March 1989) and the 3rd round (Sept. 1990) of the EIH. — [b]Constructed by extrapolating the relative difference between the rural poverty line and the weighted average urban poverty line of 1999. — [c] 1989 poverty lines inflated with the CPI.
Source: Own compilation based on unpublished data of UDAPE.

Table 12: Sample Means from the Bolivian LSMS, 1989, 1994, 1999

	Total			City			Town			Rural		
	EIH89	EIH94	ECH99	EIH89	EIH94	ECH99	EIH89	EIH94	ECH99	EIH89	EIH94	ECH99
Demographics												
Place of Residence												
City	n.a.	n.a.	49.31	100	100	100	n.a.	n.a.	0	n.a.	n.a.	0
Town	n.a.	n.a.	15.70	0	0	0	n.a.	n.a.	100	n.a.	n.a.	0
Rural	n.a.	n.a.	34.99	0	0	0	n.a.	n.a.	0	n.a.	n.a.	100
Department												
Chuquisaca	n.a.	n.a.	6.95	4.59	4.59	5.01	n.a.	n.a.	0.92	n.a.	n.a.	12.39
La Paz	n.a.	n.a.	29.09	40.48	39.63	38.41	n.a.	n.a.	12.26	n.a.	n.a.	23.51
Cochabamba	n.a.	n.a.	18.06	14.70	14.22	15.23	n.a.	n.a.	18.77	n.a.	n.a.	21.74
Oruro	n.a.	n.a.	4.48	6.71	6.19	6.48	n.a.	n.a.	1.34	n.a.	n.a.	3.06
Potosí	n.a.	n.a.	8.95	4.30	3.81	4.55	n.a.	n.a.	6.40	n.a.	n.a.	16.3
Tarija	n.a.	n.a.	4.84	3.18	3.24	2.71	n.a.	n.a.	10.93	n.a.	n.a.	5.10
Santa Cruz	n.a.	n.a.	22.44	23.90	26.29	22.96	n.a.	n.a.	41.90	n.a.	n.a.	12.97
Beni and Pando	n.a.	n.a.	5.20	2.14	2.04	4.65	n.a.	n.a.	7.49	n.a.	n.a.	4.93
Number of												
Elderly (>65)	n.a.	n.a.	0.09	0.10	0.09	0.08	n.a.	n.a.	0.10	n.a.	n.a.	0.11
Men (15–65)	n.a.	n.a.	1.43	1.48	1.49	1.53	n.a.	n.a.	1.42	n.a.	n.a.	1.29
Women (15–65)	n.a.	n.a.	1.63	1.76	1.74	1.73	n.a.	n.a.	1.79	n.a.	n.a.	1.42
Youngsters (6–14)	n.a.	n.a.	1.58	1.55	1.40	1.37	n.a.	n.a.	1.59	n.a.	n.a.	1.88
Children (0–5)	n.a.	n.a.	0.96	0.95	0.98	0.71	n.a.	n.a.	1.04	n.a.	n.a.	1.29
All HH members	n.a.	n.a.	5.70	5.84	5.70	5.42	n.a.	n.a.	5.94	n.a.	n.a.	5.99
Dependency of HH	n.a.	n.a.	56.33	57.18	58.74	61.94	n.a.	n.a.	56.45	n.a.	n.a.	48.37
Language (Spanish)	n.a.	n.a.	51.06	58.00	55.75	67.07	n.a.	n.a.	65.36	n.a.	n.a.	22.10
Gender (Female)	n.a.	n.a.	15.14	12.38	13.85	17.32	n.a.	n.a.	16.01	n.a.	n.a.	11.66
Age of HH Head												
≤ 24	n.a.	n.a.	4.63	3.73	4.51	4.47	n.a.	n.a.	6.74	n.a.	n.a.	3.92
25–34	n.a.	n.a.	21.99	26.32	25.57	21.17	n.a.	n.a.	22.05	n.a.	n.a.	23.10
35–44	n.a.	n.a.	32.28	33.37	32.60	33.85	n.a.	n.a.	29.87	n.a.	n.a.	31.16
45–54	n.a.	n.a.	26.92	20.73	22.89	26.42	n.a.	n.a.	24.48	n.a.	n.a.	28.71
55–65	n.a.	n.a.	9.48	11.52	10.31	9.91	n.a.	n.a.	11.08	n.a.	n.a.	8.14
>65	n.a.	n.a.	4.70	4.33	4.12	4.17	n.a.	n.a.	5.78	n.a.	n.a.	4.97
Tangible Assets												
Water Source												
Inhouse Access	n.a.	n.a.	66.05	71.75	79.05	93.39	n.a.	n.a.	77.72	n.a.	n.a.	22.28
Open Water Source	n.a.	n.a.	27.12	7.62	4.93	2.02	n.a.	n.a.	18.07	n.a.	n.a.	66.55
Other Water Source	n.a.	n.a.	6.83	20.63	16.02	4.60	n.a.	n.a.	4.21	n.a.	n.a.	11.17

continued on next page

Table 12 continued

	Total			City			Town			Rural		
	EIH89	EIH94	ECH99	EIH89	EIH94	ECH99	EIH89	EIH94	ECH99	EIH89	EIH94	ECH99
Toilet Facility												
No Toilet	n.a.	n.a.	31.50	32.79	25.34	11.38	n.a.	n.a.	17.55	n.a.	n.a.	66.11
Shared Toilet	n.a.	n.a.	16.66	67.21	26.24	26.99	n.a.	n.a.	12.61	n.a.	n.a.	3.94
Private Toilet	n.a.	n.a.	51.84	n.a.	48.42	61.63	n.a.	n.a.	69.84	n.a.	n.a.	29.95
House	n.a.	n.a.	67.37	58.94	56.02	56.91	n.a.	n.a.	63.35	n.a.	n.a.	83.92
Electricity	n.a.	n.a.	72.94	n.a.	95.76	98.65	n.a.	n.a.	96.54	n.a.	n.a.	26.12
Telephone	n.a.	n.a.	25.30	n.a.	20.34	43.02	n.a.	n.a.	23.91	n.a.	n.a.	0.93
Radio	n.a.	n.a.	79.57	n.a.	89.19	86.91	n.a.	n.a.	78.97	n.a.	n.a.	69.51
Television	n.a.	n.a.	66.15	n.a.	91.59	94.86	n.a.	n.a.	84.42	n.a.	n.a.	17.47
Fridge	n.a.	n.a.	35.24	n.a.	46.36	52.79	n.a.	n.a.	45.33	n.a.	n.a.	5.97
Car	n.a.	n.a.	11.48	18.82	n.a.	18.24	n.a.	n.a.	9.13	n.a.	n.a.	3.00
Family Land	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Main Floor Material												
Earth	n.a.	n.a.	34.82	n.a.	11.41	7.59	n.a.	n.a.	24.76	n.a.	n.a.	77.69
Cement	n.a.	n.a.	37.67	n.a.	43.47	49.17	n.a.	n.a.	51.86	n.a.	n.a.	15.10
Brick	n.a.	n.a.	5.95	n.a.	10.79	6.80	n.a.	n.a.	10.81	n.a.	n.a.	2.56
Other Floor	n.a.	n.a.	21.57	n.a.	34.33	36.44	n.a.	n.a.	12.57	n.a.	n.a.	4.64
Cooking Material	n.a.	n.a.	66.56	n.a.	96.98	97.40	n.a.	n.a.	81.28	n.a.	n.a.	16.48
# Sleeping Rooms												
0–1	n.a.	n.a.	58.35	n.a.	43.28	47.18	n.a.	n.a.	57.19	n.a.	n.a.	74.61
2–3	n.a.	n.a.	35.58	n.a.	46.01	42.55	n.a.	n.a.	38.18	n.a.	n.a.	24.58
≥ 4	n.a.	n.a.	6.07	n.a.	10.71	10.27	n.a.	n.a.	4.63	n.a.	n.a.	0.81
Schooling of Adults												
Men (Partners)												
No Schooling / DK	n.a.	n.a.	5.18	2.72	1.27	0.67	n.a.	n.a.	3.75	n.a.	n.a.	11.96
Incomplete Basic	n.a.	n.a.	25.82	15.66	13.08	12.54	n.a.	n.a.	24.46	n.a.	n.a.	44.53
Complete Basic	n.a.	n.a.	11.41	11.86	10.88	8.98	n.a.	n.a.	10.15	n.a.	n.a.	15.27
Lower Secondary	n.a.	n.a.	15.33	16.60	17.55	14.39	n.a.	n.a.	15.07	n.a.	n.a.	16.74
Higher Secondary	n.a.	n.a.	28.36	32.28	35.75	39.28	n.a.	n.a.	36.01	n.a.	n.a.	10.14
Tertiary Education	n.a.	n.a.	13.90	20.89	21.47	24.14	n.a.	n.a.	10.56	n.a.	n.a.	1.36
Women (15–49)												
No Schooling / DK	n.a.	n.a.	12.52	6.35	4.52	3.82	n.a.	n.a.	4.89	n.a.	n.a.	28.22
Incomplete Basic	n.a.	n.a.	23.08	18.79	15.62	13.84	n.a.	n.a.	17.97	n.a.	n.a.	38.41
Complete Basic	n.a.	n.a.	9.43	9.36	10.24	7.62	n.a.	n.a.	9.27	n.a.	n.a.	12.04
Lower Secondary	n.a.	n.a.	14.65	14.37	15.37	15.42	n.a.	n.a.	19.27	n.a.	n.a.	11.50
Higher Secondary	n.a.	n.a.	28.52	35.79	39.89	38.57	n.a.	n.a.	39.70	n.a.	n.a.	9.35
Tertiary Education	n.a.	n.a.	11.80	15.34	14.36	20.74	n.a.	n.a.	8.90	n.a.	n.a.	0.49

continued on next page

Table 12 continued

	Total			City			Town			Rural		
	EIH89	EIH94	ECH99	EIH89	EIH94	ECH99	EIH89	EIH94	ECH99	EIH89	EIH94	ECH99
Employment												
Men (Partners)												
High-skilled Admin.	n.a.	n.a.	7.54	10.50	12.12	11.90	n.a.	n.a.	6.45	n.a.	n.a.	2.04
Medium-skilled Admin.	n.a.	n.a.	8.89	8.48	11.46	13.39	n.a.	n.a.	8.76	n.a.	n.a.	2.83
Skilled Manual	n.a.	n.a.	27.49	34.32	33.33	34.94	n.a.	n.a.	37.93	n.a.	n.a.	12.84
Unskilled Manual	n.a.	n.a.	5.10	2.71	7.95	5.99	n.a.	n.a.	5.62	n.a.	n.a.	3.64
Agric.: Employed	n.a.	n.a.	5.15	1.10	0.85	1.17	n.a.	n.a.	6.02	n.a.	n.a.	10.21
Agric.: Self	n.a.	n.a.	23.97	2.44	0.53	0.11	n.a.	n.a.	4.48	n.a.	n.a.	64.92
Sales / Services	n.a.	n.a.	17.49	24.38	26.91	26.26	n.a.	n.a.	25.53	n.a.	n.a.	2.07
Never Worked / DK	n.a.	n.a.	4.37	16.06	6.84	6.25	n.a.	n.a.	5.21	n.a.	n.a.	1.45
Women (15–49)												
High-skilled Admin.	n.a.	n.a.	3.39	1.83	2.31	5.15	n.a.	n.a.	3.55	n.a.	n.a.	0.83
Medium-skilled Admin.	n.a.	n.a.	5.13	8.77	9.12	7.93	n.a.	n.a.	4.30	n.a.	n.a.	1.55
Skilled Manual	n.a.	n.a.	6.92	5.08	7.40	7.22	n.a.	n.a.	11.64	n.a.	n.a.	4.36
Unskilled Manual	n.a.	n.a.	6.75	0.84	9.34	9.72	n.a.	n.a.	8.27	n.a.	n.a.	1.87
Agric.: Employed	n.a.	n.a.	3.42	0.23	0.30	0.34	n.a.	n.a.	0.57	n.a.	n.a.	9.04
Agric.: Own	n.a.	n.a.	18.53	0.36	0.13	0.33	n.a.	n.a.	2.65	n.a.	n.a.	51.31
Sales / Services	n.a.	n.a.	15.48	26.89	23.45	22.30	n.a.	n.a.	17.72	n.a.	n.a.	4.87
Never Worked / DK	n.a.	n.a.	40.39	55.99	47.95	47.00	n.a.	n.a.	51.29	n.a.	n.a.	26.17
Health												
Social Security	n.a.	n.a.	23.70	34.01	n.a.	34.05	n.a.	n.a.	28.02	n.a.	n.a.	7.19
Birth in Last 12 Months	n.a.	n.a.	15.72	15.63	15.25	10.40	n.a.	n.a.	16.22	n.a.	n.a.	23.00
Attended by Doctor	n.a.	n.a.	55.47	65.00	72.26	83.65	n.a.	n.a.	82.06	n.a.	n.a.	29.00
Delivered in Hospital	n.a.	n.a.	40.97	52.53	58.36	61.35	n.a.	n.a.	55.18	n.a.	n.a.	23.52
Child under 4 Years	n.a.	n.a.	46.56	48.06	46.03	37.28	n.a.	n.a.	49.21	n.a.	n.a.	58.47
First Polio Vacc.	n.a.	n.a.	89.22	88.60	n.a.	89.30	n.a.	n.a.	93.29	n.a.	n.a.	87.60
Triple DPT Vacc.	n.a.	n.a.	71.13	33.69	n.a.	75.19	n.a.	n.a.	67.85	n.a.	n.a.	68.74
Had Diarrhea	n.a.	n.a.	31.49	16.25	8.28	22.45	n.a.	n.a.	35.09	n.a.	n.a.	38.24
Had Cough/Fever	n.a.	n.a.	48.73	16.46	16.32	45.09	n.a.	n.a.	43.55	n.a.	n.a.	53.96

Notes: The explanation for some of the variables is the following. DK: Don't know or no answer. Dependency of HH: Ratio of household members aged between 15 and 65 to all household members. Language and Gender refer to the language spoken (dummy for Spanish) in the household and the gender of household head (dummy for female). In-house access: To publicly provided water supply. Family Land: Taken from the employment schedule whether one member works in agriculture on family owned land. Cooking Material: High-quality cooking material such as gas, kerosene, and electricity. Schooling and employment of men (partners) refers to only married household members in the LSMS; Furthermore it had to be taken from the memory of the women (aged 15–49) in the DHS since no male module is available in the DHS survey 1989. Employment: Admin. stands for white collar workers; Agric. stands for Agriculture. Social Security: Dummy if at least one member of the household is covered by social security. Child under 4 years: Variables reflect whether the child had the first polio and triple DPT vaccinations and whether it had diarrhea and cough/fever in the last 2 or 4 weeks. For the other variables, see main text.
Source: Own Calculations.

Table 13: Sample Means from the Bolivian DHS, 1989, 1994, 1998

	Total			City			Town			Rural		
	DHS89	DHS94	DHS98	DHS89	DHS94	DHS98	DHS89	DHS94	DHS98	DHS89	DHS94	DHS98
Demographics												
Place of Residence												
City	47.55	47.96	53.46	100	100	100	0	0	0	0	0	0
Town	11.24	12.06	14.46	0	0	0	100	100	100	0	0	0
Rural	41.21	39.98	32.08	0	0	0	0	0	0	100	100	100
Department												
Chuquisaca	5.68	5.96	6.61	3.25	4.16	5.20	3.25	1.34	1.97	9.15	9.50	11.05
La Paz	36.05	31.94	30.60	42.47	40.72	38.77	21.16	13.15	11.77	32.70	27.07	25.49
Cochabamba	17.20	17.55	17.31	16.45	14.30	14.14	11.77	11.89	23.49	19.55	23.15	19.81
Oruro	6.28	6.20	4.97	6.93	7.00	6.26	5.20	7.68	4.26	5.82	4.80	3.15
Potosí	9.79	9.72	9.01	3.87	4.50	4.35	18.88	10.37	10.93	14.13	15.80	15.92
Tarija	3.90	4.50	5.31	2.93	3.15	4.32	8.04	10.07	9.30	3.88	4.43	5.16
Santa Cruz	18.25	20.91	22.04	22.44	24.49	24.77	23.25	33.83	26.57	12.06	12.72	15.45
Beni and Pando	2.87	3.22	4.14	1.67	1.67	2.20	8.46	11.67	11.70	2.72	2.52	3.97
Number of												
Elderly (>65)	0.10	0.09	0.11	0.11	0.08	0.10	0.11	0.11	0.14	0.09	0.10	0.11
Men (15–65)	1.30	1.21	1.25	1.38	1.24	1.31	1.25	1.26	1.24	1.21	1.16	1.15
Women (15–65)	1.53	1.48	1.53	1.64	1.56	1.65	1.52	1.52	1.52	1.39	1.38	1.34
Youngsters (6–14)	1.42	1.32	1.29	1.35	1.17	1.09	1.50	1.48	1.46	1.49	1.46	1.55
Children (0–5)	1.00	1.02	0.93	0.84	0.88	0.76	0.99	1.01	0.95	1.18	1.19	1.20
All HH members	5.35	5.12	5.10	5.32	4.93	4.91	5.37	5.38	5.31	5.36	5.29	5.35
Dependency of HH	56.30	56.54	58.23	59.41	60.11	63.22	55.48	55.34	55.86	52.94	52.63	50.96
Language (Spanish)	74.13	70.04	78.46	93.30	92.79	96.27	88.14	89.50	91.05	48.18	36.88	43.09
Gender (Female)	15.14	17.15	17.45	18.17	18.38	19.08	16.02	19.78	19.71	11.40	14.89	13.71
Age of HH Head												
≤ 24	6.04	8.62	7.37	5.75	8.95	7.81	5.82	8.11	6.34	6.45	8.38	7.10
25–34	26.82	28.91	26.36	27.33	29.84	25.86	24.55	28.17	26.33	26.84	28.02	27.21
35–44	30.17	28.01	30.52	29.93	27.67	30.13	31.98	30.16	30.70	29.97	27.76	31.10
45–54	19.92	20.04	20.41	19.40	20.04	21.23	20.56	19.60	19.69	20.34	20.18	19.35
55-65	10.67	9.65	9.63	10.68	9.47	9.67	10.61	8.55	10.01	10.67	10.21	9.40
≥ 66	6.38	4.77	5.71	6.90	4.03	5.29	6.48	5.41	6.94	5.74	5.46	5.84
Tangible Assets												
Water Source												
Inhouse Access	47.36	56.08	69.75	67.63	77.03	88.48	58.86	79.09	84.05	20.84	24.01	32.09
Open Water Source	29.39	29.63	19.31	6.72	4.93	1.76	12.73	11.94	8.49	60.09	64.60	53.44
Other Water Source	23.24	14.29	10.94	25.64	18.04	9.76	28.40	8.98	7.47	19.07	11.39	14.48

continued on next page

Table 13 continued

	Total			City			Town			Rural		
	DHS89	DHS94	DHS98	DHS89	DHS94	DHS98	DHS89	DHS94	DHS98	DHS89	DHS94	DHS98
Toilet Facility												
No Toilet	49.72	40.19	32.25	26.51	26.32	16.27	40.46	26.29	22.60	79.02	61.02	63.22
Shared Toilet	50.28	35.83	19.41	73.49	30.03	28.04	59.54	53.57	21.37	20.98	37.43	4.15
Private Toilet		23.98	48.34	n.a.	43.65	55.69	n.a.	20.14	56.02	n.a.	1.55	32.63
House	63.83	67.06	64.98	53.02	52.54	54.55	59.58	60.65	60.97	77.46	86.43	84.18
Electricity	n.a.	67.61	75.73	n.a.	95.00	98.41	n.a.	86.17	90.43	n.a.	29.16	31.31
Telephone	n.a.	10.59	24.96	n.a.	20.20	40.87	n.a.	6.66	19.89	n.a.	0.25	0.74
Radio	n.a.	85.17	88.08	n.a.	94.74	95.64	n.a.	85.74	88.93	n.a.	73.53	75.11
Television	n.a.	58.19	68.39	n.a.	88.32	93.46	n.a.	72.15	81.03	n.a.	17.83	20.91
Fridge	n.a.	29.69	37.67	n.a.	45.56	53.36	n.a.	35.91	43.32	n.a.	8.78	8.96
Car	12.07	n.a.	n.a.	19.60	n.a.	n.a.	10.80	n.a.	n.a.	3.73	n.a.	n.a.
Family Land	n.a.	28.46	21.27	n.a.	0.95	0.55	n.a.	9.77	6.63	n.a.	67.10	62.40
Main Floor Material												
Earth	n.a.	37.63	28.84	n.a.	14.56	7.42	n.a.	26.30	19.89	n.a.	68.73	68.58
Cement	n.a.	32.64	37.57	n.a.	41.62	43.51	n.a.	39.76	51.01	n.a.	19.72	21.62
Brick	n.a.	11.72	7.58	n.a.	15.98	9.36	n.a.	21.61	11.08	n.a.	3.62	3.04
Other Floor	n.a.	18.01	26.01	n.a.	27.84	39.71	n.a.	12.33	18.02	n.a.	7.93	6.76
Cooking Material	n.a.	64.10	71.77	n.a.	96.22	98.29	n.a.	75.18	83.92	n.a.	22.22	22.09
# Sleeping Rooms												
0–1	n.a.	53.15	59.25	n.a.	47.39	50.19	n.a.	49.94	58.85	n.a.	61.02	74.55
2–3	n.a.	41.13	34.60	n.a.	44.48	40.11	n.a.	42.87	36.57	n.a.	36.58	24.52
≥ 4	n.a.	5.73	6.16	n.a.	8.13	9.70	n.a.	7.19	4.58	n.a.	2.40	0.97
Schooling of Adults												
Men (Partners)												
No Schooling / DK	14.21	5.48	4.24	9.55	2.27	1.92	11.98	4.23	2.69	19.99	9.74	8.64
Incomplete Basic	23.99	22.84	24.18	11.33	11.10	13.69	18.90	20.12	22.28	39.39	37.79	41.84
Complete Basic	17.67	14.12	11.29	14.68	10.67	7.25	18.89	15.23	11.58	20.62	17.91	17.64
Lower Secondary	13.34	16.16	13.71	16.22	14.66	14.12	13.97	15.58	16.40	9.99	18.15	11.85
Higher Secondary	17.77	28.74	28.03	25.58	39.58	34.81	27.08	34.97	30.67	6.58	13.72	16.01
Tertiary Education	13.02	12.67	18.55	22.64	21.72	28.21	9.18	9.87	16.39	3.45	2.69	4.02
Women (15–49)												
No Schooling / DK	18.69	13.43	9.32	8.03	4.65	3.13	12.22	9.94	4.94	32.74	25.01	21.62
Incomplete Basic	29.75	27.02	23.33	21.17	18.09	14.70	26.22	22.95	18.15	40.60	38.97	40.05
Complete Basic	13.87	12.49	10.10	13.54	9.63	7.02	15.60	12.11	9.61	13.77	16.04	15.46
Lower Secondary	14.12	13.74	13.29	18.63	14.65	12.72	19.11	17.35	16.98	7.57	11.55	12.58
Higher Secondary	16.38	25.36	30.09	25.94	38.84	40.82	20.66	31.54	38.33	4.19	7.33	8.50
Tertiary Education	7.19	7.96	13.86	12.68	14.14	21.61	6.19	6.11	11.98	1.12	1.10	1.80

continued on next page

Table 13 continued

	Total			City			Town			Rural		
	DHS89	DHS94	DHS98	DHS89	DHS94	DHS98	DHS89	DHS94	DHS98	DHS89	DHS94	DHS98
Employment												
Men (Partners)												
High-skilled Admin.	9.56	6.70	8.68	16.82	12.18	13.89	8.31	4.96	7.19	1.89	0.66	0.98
Medium-skilled Admin.	8.45	9.11	8.63	12.54	13.41	11.16	10.95	12.09	9.45	3.24	2.98	4.20
Skilled Manual	25.04	25.79	24.91	32.95	35.13	31.12	33.97	28.75	27.82	13.86	13.65	13.67
Unskilled Manual	5.06	4.29	4.16	6.91	5.67	5.75	6.35	4.59	4.88	2.64	2.54	1.27
Agric.: Employed	4.37	6.01	4.33	0.48	0.98	0.77	4.10	8.95	6.91	8.77	11.14	8.91
Agric.: Self	27.55	25.12	22.26	2.15	0.76	0.99	9.92	9.62	8.32	60.47	59.31	62.58
Sales / Services	16.85	19.34	20.29	23.50	26.54	27.71	24.87	27.21	26.11	7.32	8.19	5.81
Never Worked / DK	3.11	3.64	6.73	4.65	5.33	8.61	1.52	3.83	9.31	1.83	1.53	2.59
Women (15– 49)												
High-skilled Admin.	1.43	1.42	3.07	2.58	2.39	4.93	0.67	1.34	2.40	0.31	0.30	0.28
Medium-skilled Admin.	5.39	7.14	8.17	8.38	11.30	11.29	8.29	8.90	9.37	1.16	1.61	2.41
Skilled Manual	3.58	6.53	6.99	3.93	8.25	8.18	3.43	7.10	7.53	3.22	4.30	4.76
Unskilled Manual	0.42	9.47	7.95	0.23	14.18	11.10	1.94	11.69	8.19	0.23	3.15	2.60
Agric.: Employed	0.50	6.32	0.92	0.13	0.42	0.01	0.25	1.54	0.91	1.01	14.85	2.43
Agric.: Own	0.80	15.01	12.18	0.04	0.13	0.10	0.10	2.26	1.40	1.86	36.70	37.15
Sales / Services	13.59	17.21	19.09	18.75	21.96	25.06	18.81	24.69	24.77	6.22	9.26	6.59
Never Worked / DK	74.28	36.89	41.64	65.97	41.37	39.33	66.52	42.48	45.42	85.99	29.83	43.80
Health												
Social Security	21.44	n.a.	21.31	29.19	n.a.	31.11	30.19	n.a.	23.18	10.10	n.a.	4.12
Birth in Last 12 Months	19.83	18.64	17.08	16.30	15.57	14.15	20.22	18.61	15.58	23.80	22.34	22.63
Attended by Doctor	40.29	42.06	56.73	63.31	63.20	76.54	49.36	57.50	72.66	20.00	20.50	31.24
Delivered in Hospital	36.86	31.17	42.62	56.56	46.37	51.45	50.79	40.73	60.59	17.98	16.03	27.79
Child under 4 Years	51.02	50.08	47.31	43.90	44.75	41.27	50.64	49.26	45.08	59.34	56.73	58.39
First Polio Vacc.	70.64	56.13	76.16	76.67	62.39	79.23	72.35	56.31	76.86	65.07	50.13	72.31
Triple DPT Vacc.	30.22	26.32	44.09	39.50	32.54	48.46	30.65	27.49	46.58	22.19	20.13	38.07
Had Diarrhea	29.26	21.45	20.84	28.38	21.34	19.02	30.98	24.08	19.92	29.61	20.84	23.29
Had Cough/Fever	40.93	30.35	48.17	37.31	31.80	47.13	39.71	31.85	46.78	44.29	28.56	49.89

Notes: For the explanation of the variables, see Appendix Table 12.
Source: Own Calculations.

Table 14: Extreme Poverty Indices Based on Observed and Simulated Incomes, 1989, 1994, 1998/9

	1989				1994				1998/9			
	LSMS			DHS	LSMS			DHS	LSMS			DHS
	All HH	Sample	Predicted	Simulated	All HH	Sample	Predicted	Simulated	All HH	Sample	Predicted	Simulated
Extreme Poverty Line												
Total Bolivia												
P^0	n.a.	n.a.	n.a.	54.92	n.a.	n.a.	n.a.	51.99	37.19	38.29	40.53	35.43
				(0.62)				(0.40)			(1.01)	(0.42)
P^1	n.a.	n.a.	n.a.	26.58	n.a.	n.a.	n.a.	30.73	15.13	15.64	16.79	14.13
				(0.34)				(0.21)			(0.49)	(0.19)
P^2	n.a.	n.a.	n.a.	16.12	n.a.	n.a.	n.a.	22.44	8.24	8.59	9.08	7.48
				(0.28)				(0.16)			(0.33)	(0.12)
City												
P^0	38.20	39.11	39.13	38.79	28.04	28.90	29.58	28.18	22.50	24.10	25.30	23.03
			(0.77)	(0.89)			(0.58)	(0.67)			(1.51)	(0.61)
P^1	14.58	14.95	15.95	15.87	9.47	9.74	10.24	9.57	7.39	7.86	9.03	8.19
			(0.39)	(0.42)			(0.25)	(0.30)			(0.61)	(0.26)
P^2	7.50	7.71	8.61	8.62	4.57	4.67	4.89	4.50	3.57	3.79	4.43	4.03
			(0.27)	(0.29)			(0.15)	(0.18)			(0.37)	(0.16)
Town												
P^0	n.a.	n.a.	n.a.	61.02	n.a.	n.a.	n.a.	50.97	32.45	34.19	38.88	38.17
				(1.44)				(1.14)			(2.52)	(1.08)
P^1	n.a.	n.a.	n.a.	32.27	n.a.	n.a.	n.a.	26.36	13.09	13.81	16.28	16.64
				(0.88)				(0.60)			(1.26)	(0.55)
P^2	n.a.	n.a.	n.a.	20.99	n.a.	n.a.	n.a.	17.34	7.41	7.85	9.11	9.57
				(0.71)				(0.46)			(0.88)	(0.38)
Rural Areas												
P^0	n.a.	n.a.	n.a.	71.87	n.a.	n.a.	n.a.	80.85	57.93	59.98	62.64	54.88
				(0.92)				(0.47)			(1.68)	(0.70)
P^1	n.a.	n.a.	n.a.	37.39	n.a.	n.a.	n.a.	57.43	25.88	27.37	27.90	22.92
				(0.57)				(0.31)			(0.93)	(0.37)
P^2	n.a.	n.a.	n.a.	23.45	n.a.	n.a.	n.a.	45.50	14.55	15.65	15.58	12.29
				(0.48)				(0.30)			(0.68)	(0.26)

Notes: Poverty indices are calculated using income data for departmental capitals and other urban areas, expenditure data for rural areas, and mixed income-expenditure data for total Bolivia. Standard deviations of the poverty indices in brackets (only applicable to those based on predicted and simulated incomes).
Source: Own Calculations.

Table 15: Comparison of Inequality Indices Based on Observed and Simulated Incomes, 1989, 1994, 1998/9

	1989				1994				1998/9			
	LSMS			DHS	LSMS			DHS	LSMS			DHS
	All HH	Sample	Predicted	Simulated	All HH	Sample	Predicted	Simulated	All HH	Sample	Predicted	Simulated
Total Bolivia												
Gini	n.a.	n.a.	n.a.	0.550	n.a.	n.a.	n.a.	0.583	0.528	0.524	0.537	0.531
				(0.006)				(0.004)			(0.007)	(0.004)
A(0.5)	n.a.	n.a.	n.a.	0.246	n.a.	n.a.	n.a.	0.291	0.229	0.225	0.234	0.229
				(0.006)				(0.004)			(0.007)	(0.004)
A(1.0)	n.a.	n.a.	n.a.	0.427	n.a.	n.a.	n.a.	0.541	0.401	0.397	0.410	0.404
				(0.007)				(0.004)			(0.009)	(0.005)
A(2.0)	n.a.	n.a.	n.a.	0.653	n.a.	n.a.	n.a.	0.836	0.634	0.632	0.632	0.629
				(0.007)				(0.003)			(0.010)	(0.005)
City												
Gini	0.502	0.503	0.492	0.496	0.493	0.482	0.470	0.454	0.483	0.480	0.490	0.488
			(0.006)	(0.009)			(0.005)	(0.006)			(0.011)	(0.006)
A(0.5)	0.208	0.209	0.196	0.199	0.202	0.191	0.178	0.166	0.192	0.187	0.194	0.193
			(0.006)	(0.007)			(0.004)	(0.005)			(0.009)	(0.005)
A(1.0)	0.358	0.358	0.350	0.355	0.341	0.327	0.317	0.299	0.341	0.336	0.350	0.348
			(0.008)	(0.010)			(0.006)	(0.007)			(0.013)	(0.007)
A(2.0)	0.559	0.558	0.566	0.573	0.537	0.522	0.513	0.493	0.565	0.561	0.567	0.570
			(0.009)	(0.011)			(0.006)	(0.008)			(0.015)	(0.008)
Town												
Gini	n.a.	n.a.	n.a.	0.543	n.a.	n.a.	n.a.	0.531	0.451	0.454	0.481	0.499
				(0.016)				(0.013)			(0.020)	(0.011)
A(0.5)	n.a.	n.a.	n.a.	0.241	n.a.	n.a.	n.a.	0.234	0.168	0.170	0.189	0.203
				(0.016)				(0.012)			(0.016)	(0.009)
A(1.0)	n.a.	n.a.	n.a.	0.423	n.a.	n.a.	n.a.	0.432	0.315	0.318	0.345	0.370
				(0.019)				(0.014)			(0.024)	(0.013)
A(2.0)	n.a.	n.a.	n.a.	0.661	n.a.	n.a.	n.a.	0.736	0.584	0.587	0.579	0.614
				(0.018)				(0.014)			(0.029)	(0.013)

continued on next page

Table 15 continued

	1989				1994				1998/9			
	LSMS			DHS	LSMS			DHS	LSMS			DHS
	All HH	Sample	Predicted	Simulated	All HH	Sample	Predicted	Simulated	All HH	Sample	Predicted	Simulated
	Rural Areas											
Gini	n.a.	n.a.	n.a.	0.484	n.a.	n.a.	n.a.	0.632	0.436	0.423	0.443	0.443
				(0.009)				(0.007)			(0.012)	(0.006)
A(0.5)	n.a.	n.a.	n.a.	0.191	n.a.	n.a.	n.a.	0.327	0.155	0.145	0.158	0.158
				(0.008)				(0.008)			(0.009)	(0.005)
A(1.0)	n.a.	n.a.	n.a.	0.334	n.a.	n.a.	n.a.	0.548	0.281	0.267	0.282	0.284
				(0.011)				(0.008)			(0.013)	(0.007)
A(2.0)	n.a.	n.a.	n.a.	0.529	n.a.	n.a.	n.a.	0.765	0.471	0.458	0.459	0.465
				(0.011)				(0.006)			(0.016)	(0.008)

Notes: Inequality indices are calculated using income data for the cities and towns, expenditure data for rural areas, and mixed income-expenditure data for total Bolivia. Standard deviations of the inequality indices in brackets (only applicable to those based on predicted and simulated incomes).
Source: Own Calculations.

Table 16: Extreme Poverty: Convergence and Divergence Assumptions

	1989					1994				
	Convergence		divergence		No mobility	Convergence		divergence		No mobility
	ϕ=1.1	ϕ=1.5	ϕ=0.9	ϕ=0.5	ϕ=1.0	ϕ=1.1	ϕ=1.5	ϕ=0.9	ϕ=0.5	ϕ=1.0
Extreme Poverty Line										
Total Bolivia										
P^0	53.87	55.70	52.63	49.43	54.92	52.87	54.06	51.95	48.73	51.99
	(0.57)	(0.54)	(0.58)	(0.63)	(0.62)	(0.41)	(0.41)	(0.41)	(0.46)	(0.40)
P^1	25.90	28.85	24.33	21.16	26.58	32.18	35.54	29.94	23.94	30.73
	(0.33)	(0.31)	(0.33)	(0.33)	(0.34)	(0.20)	(0.19)	(0.20)	(0.21)	(0.21)
P^2	15.64	18.49	14.25	11.72	16.12	24.01	28.25	21.33	14.89	22.44
	(0.25)	(0.25)	(0.25)	(0.23)	(0.28)	(0.15)	(0.14)	(0.15)	(0.15)	(0.16)
Town										
P^0	61.02	64.09	59.02	54.20	61.02	51.96	54.96	50.03	45.17	50.97
	(1.50)	(1.40)	(1.53)	(1.65)	(1.44)	(1.28)	(1.29)	(1.29)	(1.37)	(1.14)
P^1	32.62	37.21	30.19	25.25	32.27	27.64	32.09	25.17	19.79	26.36
	(0.88)	(0.85)	(0.89)	(0.90)	(0.88)	(0.64)	(0.63)	(0.65)	(0.67)	(0.60)
P^2	21.40	25.93	19.16	14.95	20.99	18.56	23.01	16.20	11.38	17.34
	(0.69)	(0.70)	(0.68)	(0.64)	(0.71)	(0.47)	(0.46)	(0.47)	(0.48)	(0.46)
Rural Areas										
P^0	70.00	73.61	67.54	61.07	71.87	81.67	83.73	79.93	73.36	80.85
	(0.82)	(0.76)	(0.85)	(1.05)	(0.92)	(0.47)	(0.41)	(0.49)	(0.65)	(0.47)
P^1	36.13	42.04	32.99	26.65	37.39	59.75	66.82	54.89	41.51	57.43
	(0.59)	(0.55)	(0.60)	(0.60)	(0.57)	(0.29)	(0.24)	(0.32)	(0.36)	(0.31)
P^2	22.55	28.22	19.80	14.79	23.45	48.38	57.66	42.40	27.75	45.50
	(0.48)	(0.48)	(0.47)	(0.43)	(0.48)	(0.27)	(0.24)	(0.29)	(0.30)	(0.30)

Notes: See Chapter 4 for explanation.
Source: Own Calculations.

Table 17: Inequality: Convergence and Divergence Assumptions

	1989					1994				
	Convergence		Divergence		No mobility	Convergence		Divergence		No mobility
	ϕ=1.1	ϕ=1.5	ϕ=0.9	ϕ=0.5	ϕ=1.0	ϕ=1.1	ϕ=1.5	ϕ=0.9	ϕ=0.5	ϕ=1.0
Total Bolivia										
Gini	0.546	0.568	0.535	0.514	0.550	0.602	0.626	0.587	0.550	0.583
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
A(0.5)	0.242	0.264	0.232	0.214	0.246	0.311	0.348	0.291	0.248	0.291
	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
A(1.0)	0.423	0.462	0.405	0.373	0.427	0.574	0.652	0.531	0.442	0.541
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
A(2.0)	0.651	0.707	0.625	0.579	0.653	0.864	0.933	0.815	0.692	0.836
	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.002)	(0.001)	(0.003)	(0.004)	(0.003)
Town										
Gini	0.549	0.594	0.529	0.498	0.543	0.545	0.601	0.518	0.468	0.499
	(0.015)	(0.016)	(0.015)	(0.015)	(0.016)	(0.013)	(0.014)	(0.012)	(0.011)	(0.011)
A(0.5)	0.246	0.291	0.228	0.201	0.241	0.247	0.306	0.221	0.178	0.203
	(0.015)	(0.017)	(0.014)	(0.013)	(0.016)	(0.013)	(0.016)	(0.011)	(0.009)	(0.009)
A(1.0)	0.432	0.500	0.403	0.358	0.423	0.456	0.553	0.409	0.328	0.370
	(0.018)	(0.019)	(0.018)	(0.017)	(0.019)	(0.014)	(0.015)	(0.014)	(0.013)	(0.013)
A(2.0)	0.674	0.753	0.636	0.575	0.661	0.771	0.887	0.700	0.561	0.614
	(0.017)	(0.015)	(0.018)	(0.019)	(0.018)	(0.015)	(0.010)	(0.016)	(0.016)	(0.013)
Rural Areas										
Gini	0.499	0.556	0.474	0.433	0.484	0.657	0.742	0.626	0.500	0.443
	(0.010)	(0.010)	(0.010)	(0.009)	(0.009)	(0.006)	(0.006)	(0.004)	(0.006)	(0.006)
A(0.5)	0.203	0.253	0.183	0.153	0.191	0.356	0.467	0.348	0.201	0.158
	(0.009)	(0.011)	(0.009)	(0.007)	(0.008)	(0.008)	(0.009)	(0.004)	(0.006)	(0.005)
A(1.0)	0.351	0.424	0.321	0.274	0.334	0.587	0.726	0.652	0.357	0.284
	(0.012)	(0.013)	(0.011)	(0.010)	(0.011)	(0.007)	(0.007)	(0.004)	(0.008)	(0.007)
A(2.0)	0.547	0.627	0.510	0.452	0.529	0.799	0.901	0.933	0.564	0.465
	(0.012)	(0.011)	(0.012)	(0.011)	(0.011)	(0.005)	(0.003)	(0.001)	(0.008)	(0.008)

Notes: See Chapter 4 for explanation.
Source: Own Calculations.

Table 18: Asset Endowment Among Poor and Non-Poor

	1994			1998		
	Ext. Poor	Mod. Poor	Non-Poor	Ext. Poor	Mod. Poor	Non-Poor
Tangible Assets						
Telephone	0.02	0.29	37.58	0.40	2.13	67.81
Radio	72.93	79.67	99.59	73.19	82.63	98.31
Television	21.55	42.39	99.57	21.85	51.90	99.31
Fridge	4.57	11.58	77.13	4.46	12.82	84.28
House	79.91	72.17	53.70	77.45	66.27	62.57
Family Land	54.95	39.07	0.67	53.36	32.34	0.50
Electricity	36.51	55.28	99.93	37.04	62.82	99.95
Public Water	22.72	40.27	97.49	31.25	54.53	98.30
Other (Non-open) Water Source	22.19	19.18	1.46	21.36	16.12	1.22
Cooking Material	31.59	50.76	99.06	30.04	56.99	99.50
Shared Toilet	36.73	39.73	25.61	9.58	21.27	15.92
Private Toilet	0.71	6.80	69.00	26.91	30.64	81.54
Cement Floor	20.43	30.14	39.19	22.20	37.85	37.06
Brick Floor	5.10	9.24	18.23	4.88	8.20	6.42
Other (Non-earth) Floor	6.83	9.15	41.23	6.05	10.29	55.50
2-3 Sleeping Rooms	32.74	36.04	54.44	20.10	23.45	55.51
$\geq$ 4 Sleeping Rooms	1.93	2.00	15.49	0.84	1.12	15.62
Human Capital						
% of Men with						
Complete Basic	16.85	15.49	2.43	16.25	13.31	2.29
Lower Secondary	16.69	17.14	4.34	11.49	14.00	6.84
Higher Secondary	12.71	19.51	36.44	13.55	20.84	28.61
Tertiary Education	1.27	2.15	32.98	1.86	3.65	37.95
% of Women with						
Complete Basic	17.14	16.00	3.28	16.71	14.08	2.64
Lower Secondary	13.24	16.01	7.78	14.63	16.41	7.45
Higher Secondary	6.33	13.80	55.66	7.22	19.76	49.48
Tertiary Education	0.23	1.09	25.96	0.86	2.99	34.25
Number of Observations	3382	4848	1792	3571	5439	3005

Notes: For the explanation of the variables, see Appendix Table 12. The left-out categories are: open water source, no toilet, earth floor, 0–1 sleeping rooms, no or incomplete basic schooling. The category moderately poor includes the category extremely poor, so that the number of observations of each year is the sum of moderately poor and non-poor. Numbers are in percent.
Source: Own Calculations.