

Otter, Thomas

**Book — Digitized Version**

## Poverty, Income Growth and Inequality in Paraguay During the 1990s: Spatial Aspects, Growth Determinants and Inequality Decomposition

Göttinger Studien zur Entwicklungsökonomik / Göttingen Studies in Development Economics,  
No. 23

**Provided in Cooperation with:**

Peter Lang International Academic Publishers

*Suggested Citation:* Otter, Thomas (2008) : Poverty, Income Growth and Inequality in Paraguay During the 1990s: Spatial Aspects, Growth Determinants and Inequality Decomposition, Göttinger Studien zur Entwicklungsökonomik / Göttingen Studies in Development Economics, No. 23, ISBN 978-3-631-75367-5, Peter Lang International Academic Publishers, Berlin, <https://doi.org/10.3726/b13889>

This Version is available at:

<https://hdl.handle.net/10419/182896>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>

Thomas Otter

# Poverty, Income Growth and Inequality in Paraguay During the 1990s

Spatial Aspects, Growth Determinants and  
Inequality Decomposition



Thomas Otter

## **Poverty, Income Growth and Inequality in Paraguay During the 1990s**

The Paraguayan economy did not suffer debt crises in the eighties and had significant growth rates in the second half on the seventies, but poverty remained a problem. Understanding the performance and spatial distribution of poverty and inequality over a period of more than ten years can shed new light on structural causes behind what seems to be a low growth – high poverty – high inequality trap in Paraguay. How did poverty and inequality change during the 1990s. Did inequality reduce income growth? What were the growth determinants and what are the main forces driving inequality changes? These are the questions being answered in this book.

Thomas Otter is a researcher associated to the Ibero-America Institute for Economic Research of the University of Göttingen (Germany). He holds a doctorate in economics from the same university. The author has worked as a consultant for different development agencies in Latin America, Africa, and Asia. His research interests include pro-poor growth, inequality, and human development.

**Poverty, Income Growth and Inequality  
in Paraguay During the 1990s**



# **Göttinger Studien zur Entwicklungsökonomik Göttingen Studies in Development Economics**

Herausgegeben von/Edited by Hermann Sautter und/and Stephan Klasen

Bd./Vol. 23



**PETER LANG**

Frankfurt am Main · Berlin · Bern · Bruxelles · New York · Oxford · Wien

Thomas Otter

# Poverty, Income Growth and Inequality in Paraguay During the 1990s

Spatial Aspects, Growth Determinants  
and Inequality Decomposition



PETER LANG

Internationaler Verlag der Wissenschaften

**Bibliographic Information published by the Deutsche Nationalbibliothek**

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the internet at <<http://www.d-nb.de>>.

Open Access: The online version of this publication is published on [www.peterlang.com](http://www.peterlang.com) and [www.econstor.eu](http://www.econstor.eu) under the international Creative Commons License CC-BY 4.0. Learn more on how you can use and share this work: <http://creativecommons.org/licenses/by/4.0>.



This book is available Open Access thanks to the kind support of ZBW – Leibniz-Informationszentrum Wirtschaft.

Zugl.: Göttingen, Univ., Diss., 2007

Gratefully acknowledging the support of the  
Ibero-Amerika-Institut für Wirtschaftsforschung, Göttingen.

Cover illustration by Rolf Schinke

D 7

ISSN 1439-3395

ISBN 978-3-631-57201-6

ISBN 978-3-631-75367-5 (eBook)

© Peter Lang GmbH

Internationaler Verlag der Wissenschaften

Frankfurt am Main 2008

Printed in Germany 1 2 3 4 5 7

[www.peterlang.de](http://www.peterlang.de)

Reducing inequality and poverty at an individual level is not only a problem of the efforts being carried out by men and women, to improve their conditions of life, but of their opportunities to do so.



## **Acknowledgements**

I am grateful to the national statistics bureau of Paraguay (Dirección General de Estadística, Encuestas y Censos - DGEEC) and Pablo Sauma for their help in providing the survey and census data. I would also like to thank all my classmates and colleagues at the University of Göttingen for their numerous comments and suggestions that helped to improve this work considerably. I am particularly indebted to Peter Lanjouw and Jesko Hentschel from the World Bank. My very special thanks to Prof. Stephan Klasen and Prof. Michael Grimm, for their excellent supervision, patience and permanent support and encouragement and for the opportunity Prof. Stephan Klasen gave me, accepting me as one of his students.

Asuncion, March 2007

Thomas Otter



# Contents

Acronyms	15
Preface	17
<b>1. Micro level estimation of income</b>	<b>19</b>
1.1 Introduction	19
1.2 Methodology and Data	20
1.2.1 The Basic Methodology	20
1.2.2 The Income Model	21
1.2.3 The Implementation Procedure	21
1.2.4 Data Sources	26
1.3 Results	27
1.3.1 Regression Results	27
1.3.2 Poverty Estimates	37
1.3.3 Poverty and Inequality Maps	40
1.3.4 Pro-poor Growth Evidence	44
1.4 Discussion	46
1.5 Conclusions	48
<b>2. Does Inequality Harm Income Mobility and Growth?</b>	<b>51</b>
2.1 Introduction	51
2.2 Data: macro, micro and small area welfare estimates	53
2.3 Small area welfare estimation	54
2.4 The model	59
2.5 Estimation	67
2.6 Results	71
2.7 Discussion	87
2.8 Conclusions	90
<b>3. Characterization of inequality changes through microeconomic decomposition</b>	<b>93</b>
3.1 Introduction	93
3.2 Methodology	95
3.3 Income inequality in Paraguay: basic facts and sources of changes	100
3.3.1 Returns to education	102
3.3.2 Gender wage gap	108
3.3.3 Returns to experience	108
3.3.4 Unobservable Factors	109
3.3.5 Hours of work	110



3.3.6 Labour market participation	111
3.3.7 Education	115
3.4 Estimation strategy	116
3.5 Results	120
3.6 Discussion	126
3.7 Conclusions	128
<b>A: Annex to Chapter 1</b>	<b>131</b>
<b>B: Annex to Chapter 2</b>	<b>134</b>
<b>Bibliography</b>	<b>135</b>

## List of Tables

Table 1.1 Variable definitions for 1992 estimates	28
Table 1.2 Regression results Asuncion 1992	29
Table 1.3 Regression results Central Urban 1992	30
Table 1.4 Regression results Remaining Urban 1992	31
Table 1.5 Regression results Rural 1992	32
Table 1.6 Variable definitions 2002	33
Table 1.7 Regression results Asuncion 2002	34
Table 1.8 Regression results Central Urban 2002	34
Table 1.9 Regression results Remaining Urban 2002	35
Table 1.10 Regression results Rural 2002	36
Table 1.11 Percentage of Poverty – 1992	37
Table 1.12 Percentage of Poverty – 2002	38
Table 1.13 Inequality measures	39
Table 1.14 Mean inflation rates by deciles, Par.1997/98 – 2000/01 (%)	45
Table 2.1 Welfare estimates, Paraguay, Selected Years	56
Table 2.2 Variables and descriptive statistics – Asunción	63
Table 2.3 Variables and descriptive statistics – Central Urban	64
Table 2.4 Variables and descriptive statistics – Remaining Urban	65
Table 2.5 Variables and descriptive statistics – Rural	66
Table 2.6 Variables and descriptive statistics – Pro-Poor-Growth-Panels	67
Table 2.7 Variance adjustments – Asunción	71
Table 2.8 Variance adjustments – Central Urban	72
Table 2.9 Variance adjustments – Remaining Urban	72
Table 2.10 Variance adjustments – Rural	73
Table 2.11 Variance adjustments – Pro-Poor-Growth Panels	73
Table 2.12 Regression results – Asuncion	76
Table 2.13 Regression results – Central Urban	80
Table 2.14 Regression results – Remaining Urban	82
Table 2.15 Regression results – Rural	84
Table 2.16 Regression results – Pro-Poor-Growth Panels	86
Table 3.1 Income Distribution in Paraguay, Selected Years (Gini coefficient)	101
Table 3.2 Hourly Earnings by Educational Level in Paraguay, Selected Years	103
Table 3.3 Log-Hourly Earnings Equation Applied to Paraguay, Selected Years	104
Table 3.4 Hourly Earnings by Gender in Paraguay, Selected Years	108
Table 3.5 Hourly Earnings by Age Groups, Paraguay, Selected Years	109
Table 3.6 Weekly Hours of Work by Educational Levels in Paraguay, Selected Years	110

Table 3.7 Labour Status by Household Role, Paraguay, Selected Years	112
Table 3.8 Hours of Work Equation for Paraguay, Selected Years	113
Table 3.9 Labour Status and Education, Paraguay, Selected Years	114
Table 3.10 Composition of Sample by Educational Level in Paraguay, Selected Years	115
Table 3.11 Decomposition of the Change in Gini coefficient for Earnings and Equivalent Household Labour Income and Equivalent Labour Household Income Poverty, Par.1992 – 1997/98	121
Table 3.12 Decomposition of the Change in Gini coefficient for Earnings and Equivalent Household Labour Income and Equivalent Labour Household Income Poverty, Par. 1997/98 – 2005	122
Table 3.13 Decomposition of the Change in Gini coefficient for Earnings and Equivalent Household Labour Income and Equivalent Labour Household Income Poverty, Paraguay 1992 – 2005	123
Table 3.14 Decomposition of the Change in the Gini coefficient and Equivalent Household Income Poverty Rates Changing the Base Year, Paraguay, Selected Periods	124

## List of Figures

Figure 1.1 FGT0 Per capita income - 1992 at district level	40
Figure 1.2 Gini Per capita income - 1992 at district level	41
Figure 1.3 FGT0 Per capita income - 2002 at district level	41
Figure 1.4 Gini Per capita income - 2002 at district level	42
Figure 1.5 FGT0 Per capita income Itapua department - 1992 and 2002 at district level	42
Figure 1.6 Rural poverty estimates Itapua - 2002	43
Figure 1.7 Growth incidence curve of log per capita income	46
Figure 2.1 District means of education and education inequality of household head in rural Paraguay (1992)	88
Figure 3.1 Gini coefficient of Per Capita Household Income Distribution, Poverty and GDP per capita in Paraguay, 1992 – 2005	100
Figure 3.2 Hourly Earnings-Education Profiles for Men (Heads of Household and Other Family Members), Age 40	106
Figure 3.3 Hourly Earnings-Education Profiles for Women (Spouses), Age 40	107
Figure 3.4 Weekly Hours of Work by Educational Level for Men (Heads of Household), Age 40	111



## Acronyms

CEPAL	Economic Commission for Latin America (Comisión Económica para América Latina)
CNPV	Population Census (Censo Nacional de Población y Vivienda)
DGEEC	National Statistical Office (Dirección General de Estadística, Encuestas y Censo)
ECV	Household Survey (Encuesta de Condiciones de Vida)
EIH	Household Survey (Encuesta Integrada de Hogares)
EPH	Household Survey (Encuesta Permanente de Hogares)
GDP	Gross Domestic Product
i.i.d.	independent and identically distributed
IADB	Inter-American Development Bank
LHS	Left Hand Side
MERCOSUR	Common Southern Market (Mercado Común del Sur)
NVD	Number of Direct Neighbor Districts
OLS	Ordinary Least Squares
PCI	Per Capita Income
PPG	Pro-Poor-Growth
RHS	Right Hand Side
sqr	squared
UNDP	United Nations Development Program
2SLS	Two Stage Least Squares



## Preface

During the nineteen seventies, Latin America was a promising region and seemed to have prosperous development perspectives. However, the nineteen eighties went by as the “lost decade” for most countries, mainly due to external debt crises. Although overall development perspectives were reduced during this period, the region harbored good economic performers, such as Chile, or even within some countries more and less prosperous regions remained, such as the south of Brazil compared to the northern region. Nevertheless, this better performance did not benefit large parts of the population. Until today poverty remains a non-resolved problem, even in Chile and Brazil. Paraguayan economic history is similar and at the same time different to this general trend. Even if there were no debt crises in the eighties and important growth rates in the second half of the seventies, poverty remained a non-resolved problem in the eighties and an increasing problem during the nineties. Looking back in time, prior to 2005, it would seem that Paraguay, in a way, is locked in what is known as a “low growth high poverty trap”. Understanding the performance of poverty and inequality over a period of more than 10 years can throw new light on structural causes behind what seems to be a “low growth high poverty high inequality trap” in Paraguay.

In recent years, there has been increasing empirical evidence worldwide that inequality levels and inequality changes are powerful determinants for poverty levels. Reducing poverty might be a tool for inequality reduction, if the effects of inequality on poverty are well understood, and vice versa.

This dissertation focuses on poverty and inequality issues in Paraguay during the nineties. In the first chapter, poverty levels and their spatial distribution are estimated. The second chapter searches for the effects of income and education inequality on growth, using the results of the first chapter as input. In chapter three, a decomposition of changes in inequality is carried out in order to better understand what the dynamics behind inequality changes are and what their impact on poverty is. The chapters are written to be read separately, consequently, some methodological repetitions were included.

Persistent poverty can be a serious impediment for growth. In the first chapter, a poverty and inequality mapping exercise shows that poverty levels and their spatial patterns were almost the same at the beginning of the 1990s and during the first few years of the 2000s. A small poverty decrease during the second half of the 1990s was not sustainable. So in a way, this is evidence that we have persistent poverty in Paraguay.



However, there may be some opportunities to reduce poverty without necessarily going through an important economic growth. In India, there are some states that are more efficient than others in reducing poverty through growth. Some of their strategies attack poverty in an attempt to reduce inequality even if there is not much economic growth. Consequently, there seems to be an interesting link between inequality and growth. To better understand which links existed between inequality and income growth during the nineties in Paraguay, chapter 2 takes the results from chapter 1 (estimated mean incomes by district) and examines the impact of initial income and education inequality on income growth. Since this exercise is based on poverty maps, this allows for the corroboration of spatial patterns and regional differences in the effects of income and education inequality on income growth.

Chapters 1 and 2 portray almost unchanged poverty levels at the beginning and at the end of a ten-year period, while also showing some reduction in income inequality. Questions on what drives inequality reduction during a period of almost inexistent economic growth and poverty reduction are answered halfway through chapter 3. These answers are the result of a microeconometric decomposition based on three different household surveys.

# Chapter 1

## Micro Level Estimation of Income

### Simulated welfare mapping (poverty maps) for Paraguay 1992 and 2002

#### 1.1. Introduction

Recent theoretical and empirical advances have brought income and wealth distributions back into a prominent position in growth and development theories, and as determinants of specific socio-economic outcomes, such as health or levels of violence and related phenomenon of inequality. To improve empirical investigation, new techniques were required for the simulation of small scale welfare indicators, such as income and its related distribution. Elbers, Lanjouw and Lanjouw (2003) designed a statistical procedure to combine different types of data and take advantage of the detail in household sample surveys and the comprehensive coverage of a census. The method extends the literature on small area statistics (Ghosh and Rao (1994), Rao (1999)) by developing estimators of population parameters which are non-linear functions of the underlying variable of interest (for example per capita income) by deriving them from the full unit level distribution of that variable. The most famous output of these exercises is known as “poverty maps”. The use of these poverty maps is an important poverty reduction policy implementation tool used for selecting the poorest villages in the country (or villages where the greatest number of poor people are), such as the programs Bolsa Escuela in Brasil, Progreso in Mexico, Puente in Chile, Bolsa Familia in Argentina, Bono de Desarrollo Humano in Ecuador or Tekopora in Paraguay; all of these conditional cash transfer programs, directly to extremely poor households.<sup>1</sup>

---

1 The first poverty map for Paraguay was built by Marcos Robles (2000) combining Population Census 1992 with household survey 1997/98 data using the methodology proposed in Hentschel et al (2000), although the Government did not start using this kind of tool until 2003. In 2003, the Government needed to update poverty maps urgently with census and survey data from 2002. The author of this paper carried out this update for the Social Ministry using Elbers et al (2003) methodology based on a 10% sample of census data (the only census sub-sample available by the end of 2003) and the 2002 household survey. The attained results were the input for the “Indice de Priorizacion de Gasto” IPG, a geographic targeting tool for household cash transfer programs. In 2004, the IPG ranking was updated by Marcos Robles and Horacio Santander with the entire census data from 2002 and 2003 household surveys. Although a number of poverty maps in Paraguay already exist, the results shown in this paper are the only ones that combine the entire

In this paper, the method of Elbers, Lanjouw and Lanjouw (2003) is applied using Paraguayan data from 1992 and 2002 producing estimates with levels of precision comparable to those of commonly used survey based welfare estimates - but for populations down to less than 1,000 people living within the same village. This is an enormous improvement over survey based estimates, which are typically only consistent for areas encompassing hundreds of thousands, or even millions of households. Experience using the method in South Africa, Brazil, Panama, Madagascar, and Nicaragua suggest that the method is reliable (Alderman, *et. al.* (2002), and Elbers, Lanjouw, Lanjouw, and Leite (2004)).

## 1.2 Methodology and Data

### 1.2.1 The Basic Methodology

Paraguayan household surveys collect very detailed information on household characteristics, including its income level;<sup>2</sup> however, coverage is limited and only representative at a relatively large geographical unit. Then again, Paraguayan population census has a complete coverage of all households, but collects very limited information on household characteristics and no information on income. The methodology developed by Elbers, Lanjouw and Lanjouw (2003) attempts to combine the advantage of detailed information on household characteristics obtained from a household survey with the complete coverage of a population census.

By combining the respective strengths of survey and census data, the simulated-welfare mapping method aims to estimate welfare indicators for small administrative areas. The approach uses household survey data to estimate a model of per capita income (or any other household or individual-level indicator of well-being) as a function of variables that are available in both the household survey and the population census.

The resulting parameter estimates from this estimation procedure are then used in a simulation to predict per capita income for each household in the census. Using the predicted per capita income, household level measures of poverty and inequality are then calculated and aggregated for small areas, such as districts,

---

2002 census with 2002 survey data, and what is more, show the only research results on poverty maps using 1992 census and 1992 survey data.

2 Poverty estimates by income is the official poverty measurement in Paraguay, carried out and updated periodically by National Statistical Office (DGEEC). Official poverty lines (caloric consumption line for extreme poverty and basic family basket for moderate poverty line) are updated by inflation for 4 different regions in the country; Asuncion, Central Urban, Remaining Urban and Rural.

sub-districts, or villages. This explains the origin of the name ‘simulated-welfare mapping’ for the method.

Importantly, the method allows for the calculation of standard errors for either welfare measure estimated. This feature is critical in that it offers a means of assessing the statistical reliability of the estimates as well as of comparisons across estimates for different communities.

### 1.2.2 The Income Model

Following *Elbers et al.* (2001, 2003), the empirical model of household income is defined as:

$$\ln y_{vh} = E(y_{vh} | x_{vh}) + u_{vh} \quad (1.1)$$

where  $\ln y_{vh}$  is the logarithm of per capita income of household  $h$  in village  $v$ ,  $x_{vh}$  is a vector of observed characteristics of this household (including village level variables), and  $u_{vh}$  is the error term. Note that we assume  $u_{vh}$  is uncorrelated with  $x_{vh}$ . This model is simplified by using a linear approximation to the conditional expectation  $E(y_{vh} | x_{vh})$  and decomposing  $u_{vh}$  into uncorrelated terms:

$$u_{vh} = \eta_v + \varepsilon_{vh} \quad (1.2)$$

where  $\eta_v$  represents a village level error term common to all households within the village, and  $\varepsilon_{vh}$  is a household specific error term. It is further assumed that  $\eta_v$  is uncorrelated across villages and  $\varepsilon_{vh}$  is uncorrelated across households. With these assumptions, equation (1.1) reduces to

$$\ln y_{vh} = x_{vh}\beta + \eta_v + \varepsilon_{vh}. \quad (1.3)$$

Estimation of the parameters underlying this equation, in particular the vector of parameters  $\beta$  and the distributional characteristics of the error terms, can be done by using standard tools from econometric analysis (*Elbers et al.*, 2003).

### 1.2.3 The Implementation Procedure

The standard procedure to implement the simulated-welfare mapping method for creating a map of mean income by sub-national administrative unit consists of five steps:

*Step 1: Matching Variables in the Survey and the Census*

In order to obtain rigorous estimates of income levels of the households in the census, the explanatory variables selected in the income determination model have to exist and be measured in the same way in both the household survey and in the census. If the sample of the household survey was randomly selected and is nationally representative, the distribution of each explanatory variable in the household survey can be expected to be the same as its distribution in the census.<sup>3</sup>

*Step 2: Selecting Explanatory Variables for the Income Model*

The selection of the explanatory variables in the income model starts by running a regression of log per capita income, using the survey data base, on the matched variables identified in Step 1, as well as some variables that can be created from other variables such as the square and cube of household size. In order to obtain a robust specification, variables are only selected for inclusion in the model if they contribute significantly to the explanation of per capita income. Hence variables with low statistical significance are dropped from the model.<sup>4</sup>

After a promising set of variables has been selected in this way, the regression is run again and the residuals of this regression are saved. These residuals need to be scrutinized to check if there are some outliers in the observation. If indeed there are some residual values which are far out of the range of most residual values, then these observations must be checked for coding or other errors. Ultimately, it may be necessary to delete them from the data.

The village level variables are obtained from either the population census data aggregated at the village level (for example the total population or age means of household heads in each village) or from other administrative data sources. These survey and census data can be completed with other data sources, mostly administrative data, such as the existence of public schooling (number of schools in a district) or infrastructure (kilometers of asphalt roads). These variables are then grouped into several sets such as demographic variables, village infrastructure variables, and village economic variables.

---

3 As a matter of fact, only variables that have the same distribution in census and survey are selected for inclusion in the income prediction models.

4 There are two kinds of dropped variables. First there are dummy variables whose frequencies are  $< 0.03$  or  $> 0.97$ , to be dropped (even if most of them are expected to be insignificant since they would show low variance). This is carried out in order to make sure that the values of the variables included in the model show some variance which can influence in the variance of predicted income. Second, all other variables which are not significant in regression are dropped in order to make the models as robust as possible.

The residuals of the last regression are then aggregated at the village level to calculate the mean of these residuals for each village. The variable selection is then carried out by running separate regressions of the village-level mean of residuals on each set of the village-level variables. The variables with significant  $t$ -values are selected as the candidates for inclusion in the income model.

The feasibility of including these candidate village-level variables in the income model is tested by running regressions of village dummy variable on these variables. One regression is run for each village dummy variable. If the coefficient of a certain variable in a regression is one, it shows that there is a perfect multicollinearity between this variable and the village dummy variable. This will happen if, for example, a village has a certain infrastructure while no other villages have, or on the other hand, all villages except one have a certain infrastructure. Such variables are necessarily excluded from the model.

### *Step 3: Estimating the Income Model*

The result of step 2 is a complete specification of the income model, incorporating both household-level and village-level independent variables of the model. The next step is to test whether there is heteroscedasticity in the data. This will determine the method to be employed to estimate the model. The first step to accomplish this is to estimate the model of equation (1.3) using Ordinary Least Squares (OLS) and save the residuals as variable  $\hat{u}_{vh}$ .

Based on equation (1.2) the residuals  $\hat{u}_{vh}$  are then decomposed into uncorrelated components as:

$$\hat{u}_{vh} = \hat{u}_{v*} + \left( \hat{u}_{vh} - \hat{u}_{v*} \right) = \hat{\eta}_v + e_{vh} \quad (1.4)$$

To investigate the presence of heteroscedasticity in the data, a set of potential variables that best explain the variations in  $e_{vh}^2$  are used to estimate the following logistic model:

$$\ln \left[ \frac{e_{vh}^2}{A - e_{vh}^2} \right] = z_{vh}^T \hat{\alpha} + r_{vh} \quad (1.5)$$

where we take  $A$  as being equal to  $1.05 * \max \{e_{vh}^2\}$ , as in Elbers *et al.*, (2003). This specification puts bounds on the predicted variance of  $\varepsilon_{vh}^2$ .

In the case where homoscedasticity is rejected, a household specific variance estimator for  $\varepsilon_{vh}$  is calculated as:

$$\hat{\sigma}_{\varepsilon, vh}^2 = \left[ \frac{AB}{1+B} \right] + \frac{1}{2} \hat{\text{Var}}(r) \left[ \frac{AB(1-B)}{(1+B)^3} \right] \quad (1.6)$$

where  $B = \exp \left\{ z_{vh}^T \hat{\alpha} \right\}$ .

The income model is then re-estimated using the Generalized Least Squares (GLS) method, employing the estimated variance-covariance matrix,  $\hat{\Sigma}$  with a structure shown in (1.7), resulting from equation (1.6) and weighted by the population weight,  $l_{vh}$ . The estimated parameters,  $\hat{\beta}_{GLS}$ , and their variance,

$$\text{Var} \left( \hat{\beta}_{GLS} \right),$$

are saved for use in the simulation.

$$\begin{bmatrix} \text{var}(n_s) + \text{var}(e_{sh}) & \text{var}(e_{sh}) & \text{var}(e_{sh}) & \text{var}(e_{sh}) \\ \text{var}(e_{sh}) & \text{var}(n_s) + \text{var}(e_{sh}) & \text{var}(e_{sh}) & \text{var}(e_{sh}) \\ \text{var}(e_{sh}) & \text{var}(e_{sh}) & \text{var}(n_s) + \text{var}(e_{sh}) & \text{var}(e_{sh}) \\ \text{var}(e_{sh}) & \text{var}(e_{sh}) & \text{var}(e_{sh}) & \text{var}(n_s) + \text{var}(e_{sh}) \end{bmatrix} \quad (1.7)$$

#### Step 4: Simulations on Census Data

The purpose of this step is to apply the parameters estimated in the previous step to the census data. However, since the values of these parameters are obtained through estimations, they are not the precise values of these parameters and subject to sampling error. This needs to be taken into account when applying the parameters to the census data by taking into account the sampling error of the coefficient estimates. For a start, recall that the purpose is to calculate the simulated version of equation (1.3):

$$\ln y_{vh}^s = x_{vh} \beta^s + \eta_v^s + \varepsilon_{vh}^s \quad (1.8)$$

where the superscript  $s$  refers to simulated version of each parameter or variable and now  $x_{vh}$  refers to characteristics of the households in the population census data.

#### Simulation of $\beta$

The simulated value of  $\beta$  is attained through a random draw, assuming

$$\beta \sim N \left( \hat{\beta}_{GLS}, \text{Var} \left( \hat{\beta}_{GLS} \right) \right).$$

Note that the draw has to take into account the covariance across  $\beta$ 's. The randomly drawn parameter is defined as  $\beta^s$ . The next step is to then apply this



simulated parameter to each household in the census data to calculate the value of  $x_{vh}\beta^s$ .

#### *Simulation of $\eta_v$*

The process of obtaining the simulated value of  $\eta_v$  requires two steps of simulations. This is because the variance of  $\eta$  itself is estimated with error. Hence, the first step is to obtain the simulated variance of  $\eta$ ,  $\sigma_\eta^{2s}$ . Elbers *et al.* (2003) propose to draw  $\sigma_\eta^{2s}$  from a gamma distribution:

$$\sigma_\eta^2 \sim G\left(\hat{\sigma}_\eta^2, \hat{\text{Var}}(\hat{\sigma}_\eta^2)\right).$$

Consequently, a random draw of the variance for the whole sample is exercised and its mean is defined as  $\sigma_\eta^{2s}$ . Then the second step is to randomly draw  $\eta_v^s$  for each village in the census data, assuming  $\eta_v \sim N(0, \sigma_v^{2s})$ .

#### *Simulation of $\varepsilon_{vh}$*

The process of obtaining the simulated value of  $\varepsilon_{vh}$  requires the use of the estimation results of equation (1.5). Assuming

$$\alpha \sim N\left(\hat{\alpha}, \text{Var}\left(\hat{\alpha}\right)\right),$$

a random draw of  $\alpha$  is made and defined as  $\alpha^s$ . As in the case of  $\beta$ , the draw has to take into account the covariance across  $\alpha$ 's. The simulated parameter is then used to simulate the household specific variance estimator for  $\varepsilon_{vh}$  as defined in equation (1.6) for each household in the census data. Finally, the simulated value of household specific idiosyncratic error,  $\varepsilon_{vh}^s$ , for every household in the census data is obtained by taking a random draw, assuming  $\varepsilon_{vh} \sim N(0, \sigma_{vh}^{2s})$ .<sup>5</sup>

#### *Collecting*

Now all three components of equation (1.7) have been simulated, the value of  $\ln y_{vh}^s$  for all households in the census data can be calculated by summing up the values of  $x_{vh}\beta^s$ ,  $\eta_v^s$ , and  $\varepsilon_{vh}^s$  that have been obtained. The whole set of simulations is then repeated a number (150 in our case) of times, so that in the end a database of 150 simulated values of (log) per capita household income of all the households in the census data is created. This is mainly to see if there variance within these 150 simulations in this fixed effects exercise is acceptably small.

---

5 Elbers *et al.* (2003) mention alternatives for the assumption that the error component terms follow normal distributions. In separate sets of simulations we have experimented with these alternative assumptions. In no case did this lead to significantly different results.



### *Step 5: Calculation of Poverty and Inequality Indicators*

The final output of Step 4 is a database of 150 simulated values of household income of all households in the census data. This database is used as the basis for calculating point estimates and standard errors of various poverty and inequality measures at the department, district and village levels. The point estimate of each measure is the mean of the calculated measure over the 150 simulated household incomes. Meanwhile, the standard error of this estimate is equal to the standard deviation of the calculated measure over the 150 simulated household incomes. The welfare indicators of a region – at any level – are calculated directly from the data of all individual households residing in that region.<sup>6</sup>

### **1.2.4 Data Sources**

Four sources of data were used: (i) Encuesta de Condiciones de Vida (ECV) 1992 (ii) Censo Nacional de Población y Vivienda (CNPV) 1992, (iii) Encuesta Permanente de Hogares (EPH) 2002, and (iv) CNPV 2002. Both census and EPH 2002 were carried out by the Paraguayan National Statistical Office DGEEC (Dirección General de Estadística, Encuestas y Censos), while the 1992 household survey was carried out by National University and the Inter American Development Bank.

Both surveys are representative household surveys, covering all areas of the country, with representative results for four different regions, Asuncion, Central Urban, Remaining Urban and Rural.<sup>7</sup> In the 1992 survey, 5,059 households (22,257 individuals) were interviewed, while 3,789 households (17,600 individuals) were interviewed in 2002. In general, both surveys follow the general format of a World Bank Living Standards Measurement Survey. The population censuses of 2002 and 1992 are respectively the sixth and fifth population census

---

6 The application of this poverty mapping exercise from step 3 to 5 is implemented using a computer program called PovMap (Version 1.2.4, February 2005), developed by Qinghua Zhao at the World Bank. All other steps were carried out using SPSS 13.0.

7 The Asuncion region only includes the city of Asuncion, while Central Urban covers the urban areas of the most populated department of Paraguay, called “Departamento Central”. Most of these urban areas are direct neighbors of Asuncion, together forming a kind of metropolitan region, excluding Asuncion. Remaining Urban include all other urban areas except Asuncion and Central Urban. Their common characteristic is to be urban, although not building a continuous geographic area. Rural includes all the rural areas. 1992 and 2002 household survey exclude in their sampling the departments Boqueron and Alto Paraguay in Chaco region, both remote rural areas; on the one hand due to budget constraints and, on the other, because less than 3% of the population live within these two departments.

carried out in Paraguay, both by DGEEC, in a systematic and comparable way. The previous censuses were carried out in 1950, 1962, 1972 and 1982. All censuses are carried out during the month of August, and cover the entire population living within Paraguayan territory, including foreign residents.

## 1.3 Results

### 1.3.1 Regression Results

In order to estimate per capita income for every household in the census, a set of household head individual characteristics, such as age, years of schooling or economic activity, characteristics of the spouse and the family group, such as number of children, their schooling or economic activity, characteristics of the habitat, access to basic services and assets within the household are used. We also tried some local infrastructure data such as kilometers of asphalt road, number public schools, the existence of a post office, public transport or government organized market places in the district.<sup>8</sup> To reinforce empirical evidence on spatial effects we controlled the number of direct neighboring districts<sup>9</sup> and percentage of economic activity by different sectors in neighboring districts.

As the aim of the regression models is to predict as precise as possible per capita income for every household in the census, using coefficients from regressions based on household surveys, including only common variables from survey and census, household survey regressions results need not be understood as regressions on determinants of income, but as regressions on variables which are correlated with income. Variables that are correlated with income, and not only variables which determine income, are used in order to achieve good results.

---

8 Most of these happened to be insignificant and were excluded from the final models (see footnote 4). Only for the “Remaining Urban Area in 2002” variable did the Kilometer of asphalt roads in each district (ROAD) happen to be significant and was included in the final model (see Table 1.9).

9 The number of direct neighboring districts (NVD) for Paraguayan economy can be understood as a proxy for closeness to areas of higher economic dynamics. Four out of the five most economically important cities are border cities, three of these with twin cities on the other side of the border. For being border districts the number of direct neighbor districts is smaller than for other districts within the country. The NVD variable can be a proxy to measure local effects of these more dynamic border cities.

**Table 1.1 Variable definitions for 1992 estimates**

AECON	Years of schooling spouse	RADIO	Household has radio (dummy)
AEDUJ2	Years of schooling head squared	SANITA1	Household has latrine (dummy)
AEJEFE	Years of schooling household head	TECHO1	Roof of cement (dummy)
AEM12S2	Sum of years of schooling > 12 years sqrd	TIPO2	Habitat is a house (dummy)
AEM15S2	Sum of years of schooling > 15 years sqrd	VIVI1	Own habitat (dummy)
AEM18S	Sum of years of schooling > 18 years	VIVI2	Rented habitat (dummy)
AEM18S2	Sum of years of schooling > 18 years sqrd	AEDUJ_2_1	Cluster mean of years of schooling head sqrd
AGUA1	Household connected to public network of water supply (dummy)	AEDUJ3_1	Cluster mean years of schooling head cubed
D3	Department Cordillera (dummy)	AEM18S_1	Cluster mean of sum of years of schooling > 18 years
D6	Department Caazapa (dummy)	AEM18S_2	Cluster mean of sum of years of schooling > 18 years squared
D7	Department Itapua (dummy)	AGUA1_1	Cluster mean of household connected to public network of water supply (dummy)
D8	Department Misiones (dummy)	CONAGR_1	Cluster mean of spouse employed or self-employed in agricultural sector
D10	Department Alto Parana (dummy)	D3_1	Cluster mean of Department Cordillera
EDADJE	Age of household head	D4_1	Cluster mean of Department Guaira
EDJ2	Age of household head squared	FUENTE_1	Cluster mean of water supply inside the house
EDJ3	Age of household head cubed	HELA_1	Cluster mean of household has refrigerator
FUENTE1	Water supply inside the house (dummy)	JAGRO_1	Cluster mean of head working in agriculture
HELA	Household has refrigerator (dummy)	JCOM_1	Cluster mean of household head working in commercial sector
JAGRO	Head working in agriculture sector (dummy)	JTRAN_1	Cluster mean of household head working in transport sector
JEFEOC	Head employed or self-employed (dummy)	LUZ_1	Cluster mean of household has electricity
JSERV	Household head working in service sector (dummy)	MAT182_1	Cluster mean of number of household members > 18 enrolled in education squared
LNTOTP	Log total number of persons in household	MAT183_1	Cluster mean of number of household members > 18 enrolled in education cubed
LUZ	Household has electricity (dummy)	MATM18_1	Cluster mean of number of household member > 18 years enrolled in education
MATTOT	Total number of household member enrolled in any educational institution	MATTOT_1	Cluster mean of total number of household member enrolled in education
MCOSER	Household has sewing-machine (dummy)	NDOR2_1	Cluster mean of number of bedrooms squared
NDOR	Number of bedrooms	NDOR3_1	Cluster mean of number of bedrooms cubed
NVD	Number of direct neighbor districts	OCUM18_1	Cluster mean of number of household members > 18 employed or self-employed
OCUM122	Number of household member > 12 years employed or self-employed squared	PEPS_1	Cluster mean of percentage of secondary employment in neighbor districts
OCUM123	Number of household member > 12 years employed or self-employed cubed	PET_1	Cluster mean of percentage of tertiary sector employment in district
OCUM183	Number of household member > 18 years employed or self-employed cubed	TECHO1_1	Cluster mean of roof of cement
PEA	Number of household member in economic activity	TIPO1_1	Cluster mean of habitat is a house
PEPT	Percentage of tertiary sector employment in neighbor districts	TIPO2_1	Cluster mean of habitat is an apartment
PES	Percentage of secondary sector employment in district	TOT3_1	Cluster mean of total number of household members cubed

Source: CNPV and ECV 1992

**Table 1.2 Regression results Asuncion 1992**

Var.	Coef.	Std.E.	t	Prob> t
Intercept	13.444	0.551	24.41	<.0001
AEDUJ2	0.002	0.000	9.600	<.0001
PEA	0.238	0.017	13.61	<.0001
LNTOTP	-1.262	0.051	-24.85	<.0001
VIVI1	0.250	0.034	7.454	<.0001
NDOR	0.152	0.016	9.415	<.0001
AECON	0.024	0.003	7.488	<.0001
FUENTE1	0.197	0.048	4.103	<.0001
EDJ2	0.000	0.000	2.390	0.017
RADIO	0.314	0.065	4.790	<.0001
MATTOT	0.056	0.014	4.138	<.0001
TECHO1	-0.332	0.085	-3.918	<.0001
AEM18S	0.014	0.002	5.570	<.0001
JSERV	-0.088	0.032	-2.708	0.0069
MCOSER	0.100	0.030	3.340	0.0009
AEM18S2	0.000	0.000	-4.541	<.0001
TIPO2	0.211	0.074	2.833	0.0047
AEDUJ3_1	0.000	0.000	3.916	<.0001
JAGRO_1	2.185	0.466	4.687	<.0001
FUENTE_1	-0.455	0.108	-4.204	<.0001
AGUA1_1	0.315	0.106	2.962	0.0031
AEM18S_2	0.000	0.000	-3.088	0.0021
NDOR2_1	0.030	0.007	3.983	<.0001
LUZ_1	-1.692	0.559	-3.025	0.0025
MAT183_1	0.417	0.169	2.474	0.0135
JCOM_1	-0.365	0.117	-3.117	0.0019
MATM18_1	1.975	0.573	3.449	0.0006
MAT182_1	-1.989	0.652	-3.051	0.0023
Nr. of obs.		1,346		
Adj. R-sqrd.		0.636		
Cluster		89		
var( $\hat{\eta}_s$ )		0.084		
var(var( $\hat{\eta}_s$ ))		<0.001		
var( $\hat{\epsilon}_{sh}$ )		5.987		

Source: Author's calculations based on ECV 1992

**Table 1.3 Regression results Central Urban 1992**

Var.	Coef.	Std.E.	t	Prob> t
Intercept	11.351	0.371	30.599	<.0001
AEDUJ2	0.002	0.000	9.742	<.0001
PEA	0.381	0.027	14.230	<.0001
LNTOTP	-0.981	0.051	-19.268	<.0001
AGUA1	0.154	0.046	3.331	0.0009
HELA	0.208	0.054	3.835	0.0001
VIVI1	0.215	0.045	4.804	<.0001
NVD	-0.040	0.016	-2.517	0.0121
TECHO1	-0.309	0.086	-3.577	0.0004
NDOR	0.079	0.025	3.185	0.0015
OCUM123	-0.015	0.003	-4.924	<.0001
JAGRO	-0.453	0.126	-3.587	0.0004
OCUM183	0.013	0.003	3.773	0.0002
MCOSE	0.077	0.038	2.021	0.0437
AEM18S_1	0.080	0.019	4.192	<.0001
PET_1	1.014	0.308	3.290	0.0011
OCUM18_1	-0.372	0.096	-3.881	0.0001
TOT3_1	0.001	0.000	2.396	0.0168
HELA_1	-0.591	0.218	-2.717	0.0068
NDOR3_1	0.015	0.004	3.590	0.0004
MATTOT_1	-0.196	0.060	-3.255	0.0012
CONAGR_1	10.986	3.456	3.179	0.0015
AEM18S_2	-0.001	0.000	-2.103	0.0359
Nr. of obs.		679		
Adj. R-sqrd.		0.617		
Cluster		46		
$\text{var}(\hat{\eta}_s)$		0.056		
$\text{var}(\text{var}(\hat{\eta}_s))$		<0.001		
$\text{var}(\hat{e}_{sh})$		5.676		

Source: Author's calculations based on ECV 1992

**Table 1.4 Regression results Remaining Urban 1992**

Var	Coef.	Std.E.	T	Prob> t
Intercept	10.721	0.167	64.175	<.0001
AEJEFE	0.036	0.004	8.757	<.0001
PEA	0.370	0.023	16.047	<.0001
LNTOTP	-0.990	0.039	-25.44	<.0001
HELA	0.183	0.031	5.897	<.0001
AECON	0.026	0.004	6.850	<.0001
EDADJE	0.006	0.001	6.250	<.0001
PEPT	0.837	0.233	3.595	0.0003
JAGRO	-0.403	0.047	-8.527	<.0001
D3	-0.171	0.047	-3.650	0.0003
VIVI2	-0.196	0.039	-5.085	<.0001
AGUA1	0.056	0.028	1.984	0.0474
JSERV	-0.176	0.034	-5.235	<.0001
D7	0.340	0.041	8.352	<.0001
D10	0.294	0.051	5.798	<.0001
SANITA1	-0.130	0.030	-4.322	<.0001
OCUM122	-0.032	0.007	-4.430	<.0001
NDOR	0.070	0.014	4.947	<.0001
AEM15S2	0.000	0.000	-2.385	0.0172
MATTOT	0.034	0.011	3.056	0.0023
AEM18S	0.010	0.003	3.723	0.0002
PES	0.731	0.165	4.429	<.0001
D6	0.478	0.199	2.402	0.0164
AEDUJ2_1	0.002	0.000	5.173	<.0001
TIPO1_1	0.437	0.134	3.252	0.0012
TIPO2_1	0.937	0.346	2.710	0.0068
PEPS_1	-0.532	0.334	-1.594	0.111
Nr. of obs.		1997		
Adj. R-sqrd.		0.604		
Cluster		157		
var( $\hat{\eta}_s$ )		0.079		
var(var( $\hat{\eta}_s$ ))		<0.001		
var( $\hat{e}_{sh}$ )		5.870		

Source: Author's calculations based on ECV 1992

**Table 1.5 Regression results Rural 1992**

Var	Coef.	Std.E.	t	Prob> t
Intercept	10.795	0.126	85.638	<.0001
NDOR	0.100	0.025	4.010	<.0001
LUZ	0.452	0.059	7.626	<.0001
PEPT	2.975	0.352	8.462	<.0001
LNTOTP	-0.930	0.047	-19.815	<.0001
HELA	0.227	0.056	4.057	<.0001
D8	-0.480	0.106	-4.527	<.0001
PEA	0.183	0.031	5.874	<.0001
JAGRO	-0.403	0.057	-7.123	<.0001
OCUM122	-0.026	0.006	-4.585	<.0001
AEDUJ2	0.002	0.001	2.662	0.0079
JEFE0C	0.358	0.084	4.267	<.0001
EDJ3	0.000	0.000	3.557	0.0004
MCOSER	0.113	0.046	2.448	0.0146
AEM15S2	-0.001	0.000	-3.945	<.0001
AEM18S	0.022	0.005	4.123	<.0001
AEM12S2	0.000	0.000	3.853	0.0001
TECHO1_1	0.715	0.147	4.861	<.0001
D4_1	-0.330	0.060	-5.544	<.0001
NDOR3_1	0.009	0.003	3.219	0.0013
JTRAN_1	-3.436	0.971	-3.537	0.0004
D3_1	-0.188	0.089	-2.122	0.0341
Nr. of obs.	978			
Adj. R-sqrd.	0.574			
Cluster	31			
var( $\hat{\eta}_s$ )	0.075			
var(var( $\hat{\eta}_s$ ))	<0.001			
var( $\hat{e}_{sh}$ )	5.012			

Source: Author's calculations based on ECV 1992

**Table 1.6 Variable definitions 2002**


---

AEJ	Years of schooling of household head
AG_POCB	Water source is a well with pump (dummy)
AIRE	Household has air conditioning (dummy)
ALB	Household with at least one illiterate (dummy)
ANTENA	Household has an antenna for satellite TV (dummy)
AUTO	Household has a car (dummy)
BA_REC	Household with public removal of garbage (dummy)
CABLE	Household has cable TV dummy)
CEL	Household has at least one cell phone (dummy)
CO_GAS	Household cooks with gas (dummy)
CO_LENA	Household cooks with firewood (dummy)
DEP_AMA	Department Amambay (dummy)
DEP_CAAG	Department Caaguazu (dummy)
DEP_SANP	Department San Pedro (dummy)
EDADJ	Age of household head
ESPV	Percentage of primary sector employment in neighboring districts
HELA	Household has refrigerator (dummy)
JOCAGR	Household head working in agriculture sector (dummy)
JOCMAC	Household head working as machine operator (dummy)
LNTOTP	Log total number of persons in household
LUZ	Household has electricity (dummy)
MOTO	Household has a motorcycle (dummy)
NBIVIVI	Household with unmet basic needs in housing (dummy)
NHIJOS	Total number of children in household
NM15OC	Number of household members > 15 year occupied
NPER_PIE	Number of individuals per bedroom
NPIEZ	Number of rooms in habitat
OCJ	Household head employed or self employed
PA_LAD	Habitat walls of brick (dummy)
PC	Household has a PC (dummy)
PI_LAD	Habitat with floor of brick (dummy)
PI_TIER	Habitat with floor of earth
ROAD	Km of asphalt roads in district
TE_ZIN	Habitat with roof of zinc (dummy)
TEL	Household with telephone line (dummy)
VDVD	Household has a DVD player (dummy)
ANTENA_1	Cluster mean of household with an antenna for satellite TV
ASI18A_1	Cluster mean of number of household members 18 to 24 enrolled in education
BA_HOY_1	Cluster mean of households which trough their garbage in a hole
NBIVIV_1	Cluster mean of households with unmet basic needs in housing
TE_ZIN_1	Cluster mean of households with a roof of zinc

---

Source: CNPV and ECV 1992



**Table 1.7 Regression results Asuncion 2002**

Var.	Coef.	Std.Err.	t	Prob> t
Intercept	3.382	0.095	140.276	<.0001
TEL	0.370	0.061	6.050	<.0001
NPER_PIE	0.173	0.024	-7.132	<.0001
VDVD	0.266	0.061	4.391	<.0001
CABLE	0.247	0.061	4.064	<.0001
PC	0.259	0.067	3.872	0.0001
CO_LENA	-0.468	0.139	-3.370	0.0008
AEJ	0.012	0.006	2.161	0.0312
LNTOTP	-0.284	0.058	-4.934	<.0001
ALB	-0.140	0.058	-2.421	0.0159
Nr. of obs.		476		
Adj. R-sqrd.		0.612		
Cluster		51		
var( $\hat{\eta}_s$ )		0.145		
var(var( $\hat{\eta}_s$ ))		<0.001		
var( $\hat{\epsilon}_{sh}$ )		4.938		

Source: Author's calculations based on EPH 2002

**Table 1.8 Regression results Central Urban 2002**

Var.	Coef.	Std.Err.	t	Prob> t
Intercept	12.677	0.095	132.741	<.0001
NPIEZ	0.113	0.013	8.365	<.0001
LNTOTP	-0.781	0.067	-11.517	<.0001
TEL	0.270	0.060	4.498	<.0001
AIRE	0.255	0.066	3.843	0.0001
AEJ	0.027	0.005	5.119	<.0001
NM15OC	0.209	0.018	11.164	<.0001
OCJ	0.356	0.048	7.415	<.0001
NHIJOS	-0.043	0.016	-2.613	0.0093
BA_REC	0.142	0.0456	3.158	0.0017
Nr. of obs.		495		
Adj. R-sqrd.		0.632		
Cluster		68		
var( $\hat{\eta}_s$ )		0.087		
var(var( $\hat{\eta}_s$ ))		<0.001		
var( $\hat{\epsilon}_{sh}$ )		4.890		

Source: Author's calculations based on EPH 2002

**Table 1.9 Regression results Remaining Urban 2002**

Var.	Coef.	Std.Err.	t	Prob> t
Intercept	13.344	0.288	46.286	<.0001
PI_TIER	-0.226	0.061	-3.712	0.0002
PI_LAD	-0.197	0.051	-3.811	0.0001
NPIEZ	0.097	0.012	7.634	<.0001
LNTOTP	-0.702	0.034	-20.62	<.0001
HELA	0.240	0.042	5.705	<.0001
AUTO	0.221	0.043	4.947	<.0001
TEL	0.202	0.051	3.971	<.0001
CABLE	0.200	0.053	3.705	0.0002
PC	0.268	0.077	3.455	0.0006
AEJ	0.041	0.005	8.063	<.0001
OCJ	0.231	0.040	5.738	<.0001
JOCAGR	-0.281	0.059	-4.725	<.0001
ROAD	0.001	0.000	4.143	<.0001
EDADJ	0.005	0.001	3.799	0.0002
ESPV	-0.529	0.112	-4.73	<.0001
ANTENA_1	0.667	0.183	3.633	0.0003
ASI18A_1	-0.621	0.203	-3.049	0.0023
NBIVIV_1	-0.712	0.237	-3.005	0.0027
Nr. of obs.		1158		
Adj. R-sqrd.		0.576		
Cluster		136		
$\text{var}(\hat{\eta}_s)$		0.083		
$\text{var}(\text{var}(\hat{\eta}_s))$		<0.001		
$\text{var}(\hat{\epsilon}_{sh})$		5.416		

Source: Author's calculations based on EPH 2002

**Table 1.10 Regression results Rural 2002**

Var.	Coef.	Std.Err.	t	Prob> t
Intercept	12.306	0.112	109.256	<.0001
PA_LAD	0.164	0.036	4.527	<.0001
PI_TIER	-0.273	0.037	-7.236	<.0001
TE_ZIN	0.085	0.047	1.804	0.0715
LUŽ	0.157	0.047	3.603	0.0003
AG_POCB	0.188	0.044	4.218	<.0001
NPIEZ	0.094	0.015	6.241	<.0001
LNTOTP	-0.564	0.054	-10.347	<.0001
CEL	0.173	0.045	3.783	0.0002
AIRE	0.972	0.130	7.475	<.0001
AUTO	0.270	0.062	4.324	<.0001
MOTO	0.155	0.044	3.519	0.0004
ANTENA	0.404	0.085	4.728	<.0001
CO_GAS	0.0731	0.050	1.438	0.1507
NHIJOS	-0.039	0.010	-3.912	<.0001
OCJ	0.321	0.050	6.356	<.0001
JOCAGR	-0.306	0.038	-7.881	<.0001
JOCMAC	0.281	0.094	2.968	0.003
NBIVIVI	-0.084	0.041	-2.023	0.0432
DEP_SANP	-0.228	0.053	-4.271	<.0001
DEP_CAAG	-0.215	0.046	-4.669	<.0001
DEP_AMA	-0.404	0.107	-3.749	0.0002
ESPV	-0.329	0.109	-3.015	0.0026
BA_HOY_1	0.645	0.171	3.76	0.0002
TE_ZIN_1	0.537	0.111	4.817	<.0001
Nr. of obs.		1537		
Adj. R-sqrd.		0.632		
Cluster		207		
var( $\hat{\eta}_s$ )		0.125		
var(var( $\hat{\eta}_s$ ))		<0.001		
var( $\hat{e}_{sh}$ )		4.568		

Source: Author's calculations based on EPH 2002

All eight models produce acceptable results with most of the adjusted R sqrd. > 0.6; var( $\hat{\eta}_s$ ) between 0.15 and 0.06; var(var( $\hat{\eta}_s$ )) < 0.001 and var( $\hat{e}_{sh}$ ) between 4.6 and 5.9. For 1992 a higher number of significant variables (13 to 22) were identified. Additionally, 1992 models include between 4 and 11 cluster means variables, identifying local effects in the error term. In 2002, only 2 models include cluster means and the number of significant variables from census is much smaller. For 1992 and 2002, all coefficients from census variables have the expected signs, but not all the cluster means do. Interestingly, variables included to produce additional evidence on more extended spatial effects (passing district borders) produce the expected results. In 1992, NVD has a negative coefficient for Central Urban and percentage of tertiary sector employment in neighboring districts has

a positive sign for Remaining Urban and Rural models. For 2002, the percentage of primary sector employment in neighboring districts has a negative sign for the Remaining Urban and Rural areas. Infrastructure data such as availability of public transport, education and health care institutions, post offices and public market places did not produce any significant results for income estimation. It was only for the Remaining Urban area 2002 that we found a significant and positive effect of kilometers of asphalt roads.

### 1.3.2 Poverty Estimates

The poverty measures calculated are the poverty headcount index (P0), poverty gap index (P1), and poverty severity index (P2) from the FGT family of poverty measures.<sup>10</sup> Meanwhile, the inequality measure calculated is the Gini ratio and General Entropy measures (GE0; GE1; GE2), as deciles mean incomes, but only P0 and Gini results are reported in this paper. In addition to the estimates of poverty and inequality indicators as usually presented, the results of the simulated-welfare mapping exercise also provide the standard errors of these estimates as a measure of their precision.

Tables 1.11 and 1.12 compare the estimated headcount poverty rate as reported in the household surveys in 1992 and 2002 and those estimated from the Population Census data (standard error of simulations in brackets, standard error as percentage of estimated poverty in squared brackets).

**Table 1.11 Percentage of Poverty – 1992**

	Survey				
	ASU	CU	RU	RUR	TOTAL
Extreme Poverty	4.4	11.2	16.1	28.2	20.4
Moderate Poverty	19.0	29.6	25.8	18.3	22.2
Total Poverty	23.4	40.8	41.9	46.5	42.6
Mean income	318,869	175,631	156,957	69,243	136,116
	Simulations				
Extreme Poverty	6.8 (0.7) [10.3]	12.9 (2.7) [20.7]	21.8 (0.8) [3.8]	31.9 (1.3) [4.1]	23.5 (1.4) [5.8]
Moderate Poverty	16.6	27.2	25.1	17.6	20.8
Total Poverty	23.4 (1.2) [5.1]	40.1 (5.3) [13.2]	46.9 (0.9) [1.9]	49.5 (1.3) [2.6]	44.3 (1.9) [4.3]
Mean income	335,400	188,143	145,187	79,710	142,519

Notes: Extreme Poverty Line 54,286 (ASU); 53,333 (CU); 47,500 (RU); 39,840 (RUR)

Poverty Line: 110,738 (ASU); 108,000 (CU); 89,000 (RU); 42,258 (RUR)

Source: Author's calculations based on ECV 1992 and CNPV 1992.

For 1992, extreme poverty is higher in the Asuncion, Remaining Urban and Rural simulations, even when considering standard errors. Simulated mean income is higher than that observed in the Asuncion, Central Urban and Rural regions. This seems to be a consequence of differences in the income distribution between observed incomes in the survey and simulated incomes with census data. Overall poverty is almost the same except for the Remaining Urban area. We overestimate extreme poverty and underestimate moderate poverty but have an almost exact result for overall poverty except in Remaining Urban where simulated overall poverty exceeds observed poverty by 12%. General over-estimate is 4%. Since we overestimate mean income but nevertheless get higher poverty rates, there seem to be differences in the observed income distribution in survey and the simulated income distribution in census estimates.<sup>11</sup>

**Table 1.12 Percentage of Poverty – 2002**

	Survey				
	ASU	CU	RU	RUR	TOTAL
Extreme Poverty	7.7	15.9	16.2	31.1	21.7
Moderate Poverty	21.5	37.9	22.7	19.4	24.6
Total Poverty	29.1	53.8	38.9	50.5	46.3
Mean income	664,097	374,642	353,579	188,998	317,063
	Simulations				
	6.7	16.2	16.6	34.4	24.1
Extreme Poverty	(1.1)	(1.3)	(0.9)	(0.9)	(0.9)
	[16.4]	[8.0]	[5.4]	[2.6]	[3.7]
Moderate Poverty	23.2	39.6	23.9	19.9	24.5
	29.8	55.7	40.5	54.3	48.5
Total Poverty	(1.9)	(1.4)	(1.0)	(0.9)	(1.1)
	[6.4]	[2.5]	[2.5]	[1.7]	[2.3]
Mean income	725,053	375,413	355,013	186,634	322,467

Notes: Extreme Poverty Line 142,308 (ASU); 140,717 (CU); 106,802 (RU); 73,501 (RUR)

Poverty Line: 321,229 (ASU); 317,998 (CU); 197,895 (RU); 118,483 (RUR)

Source: Author's calculations based on EPH 2002 and CNPV 2002.

Considering standard errors, the estimates for extreme poverty for 2002 fit for all regions except the Rural area. In this model, simulated total poverty exceeds observed poverty in all areas by an average of about 5%. We are still overestimating mean income, although not as much as in 1992.

To understand what could be the possible reasons for the differences between observed results in the surveys and estimated results in census database, sev-

11 One possible source for these distribution differences is the fact that household observed household income in the survey “is not continuous” (this is that several of the higher centiles [ $>80$ ] are empty [no observations]). Nevertheless, in simulation exercises there will be estimates for a “continuous distribution”.

eral tests were carried out. There is almost no difference between observed (survey) and simulated (census) covariate levels. Only in one (Central Urban 2002) out of the eight models in use, simulated census covariate levels differed from the observed ones in the survey in more than 1%. Furthermore, there is no problem with the residual assumption in survey regressions. In all cases, residuals have a mean zero and a low variance. Consequently, if there are no structural problems in the models themselves, one possible source for biases could be a sampling problem in the actual surveys, for example classification of households in urban and rural areas. In fact, for 2002 the maximum difference for population share by region between survey and census is 0.2%. However, in 1992 there is a classification problem for the Asuncion and Rural areas (Asuncion has 11.6% of population in census and 14.9% in survey; Rural area has 48.6% in census and 45.6% in survey). The maximum difference for Central Urban and Remaining Urban is 0.5%. Recall that the 1992 survey was the first nationwide household survey ever carried out in Paraguay, so there may have been some lack of experience in paying attention to all the details. This hypothesis is consistent with differences between 1992 and 2002 simulations, where biases in 2002 are much smaller than in 1992. The way sampling problems can introduce biases in the results seems to be through the variance of household specific error ( $e_{sh}$ ). The variance of 1992 errors exceeds that of 2002 errors between 9 and 23 percent.

**Table 1.13 Inequality measures**

	Survey					Simulations				
	ASU	CU	RU	RUR	TOTAL	ASU	CU	RU	RUR	TOTAL
Gini income 1992	0.519	0.408	0.464	0.499	0.558	0.493 (0.021) [0.043]	0.439 (0.017) [0.039]	0.480 (0.007) [0.015]	0.547 (0.017) [0.031]	0.589 (0.015) [0.025]
Gini income 2002	0.437	0.419	0.446	0.534	0.561	0.444 (0.015) [0.034]	0.394 (0.011) [0.028]	0.460 (0.009) [0.019]	0.538 (0.011) [0.020]	0.510 (0.011) [0.022]

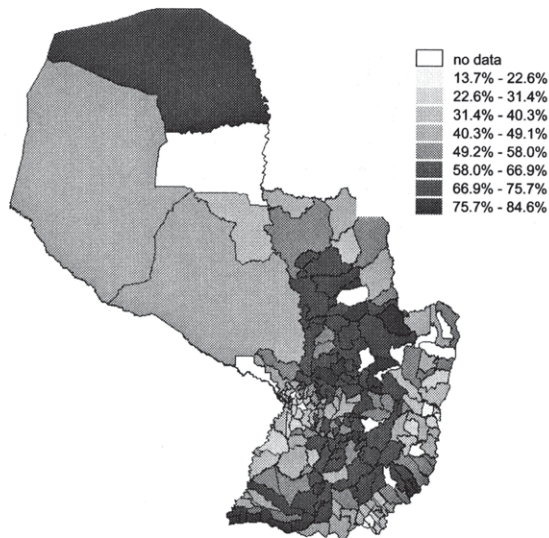
Source: Author's calculations based on ECV 1992, CNPV 1992, EPH 2002 and CNPV 2002.

For 1992 we underestimate inequality for Asuncion but overestimate inequality in all other regions. These differences also seem to be a consequence of allocation differences between rural and urban areas in 1992, between census and survey, for the Asuncion and Rural areas. For 2002, the estimated inequality in the Asuncion and Rural areas are slightly higher than those observed and we underestimate inequality in the Central Urban and Remaining Urban areas.

### 1.3.3 Poverty and Inequality Maps

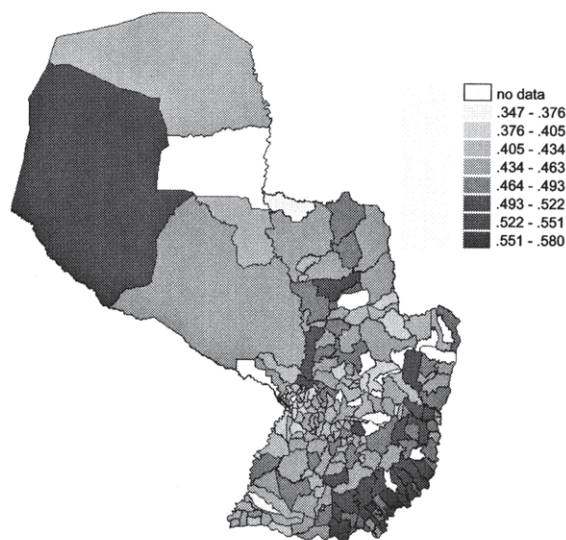
When examining these maps, it should be kept in mind that they have been created using the *expected* headcount. The *true* headcount for a location will differ from the expected headcount because of sampling and modeling errors. The maps do not take the errors into account.

**Figure 1.1 FGT0 Per capita income 1992 at district level**



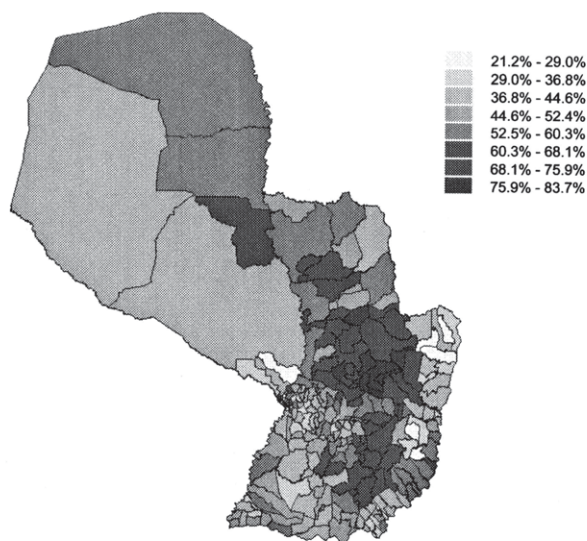
Source: Author's calculations based on ECV 1992 and CNPV 1992

**Figure 1.2 Gini Per capita income 1992 at district level**



Source: Author's calculations based on ECV 1992 and CNPV 1992

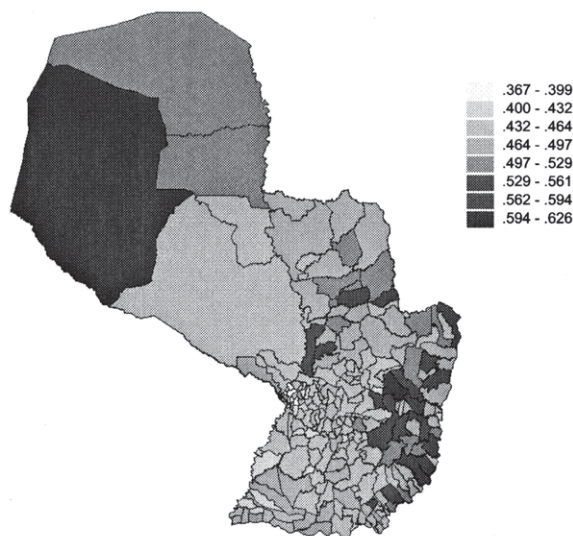
**Figure 1.3 FGT0 Per capita income 2002 at district level**



Source: Author's calvulations based on EPH 2002 and CNPV 2002

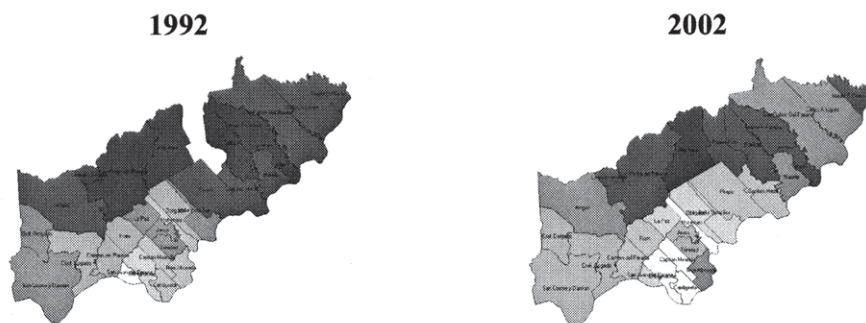


**Figure 1.4 Gini Per capita income 2002 at district level**



Source: Author's calculations based on EPH 2002 and CNPV 2002

**Figure 1.5 FGT0 Per capita income Itapua department 1992 and 2002 at district level**



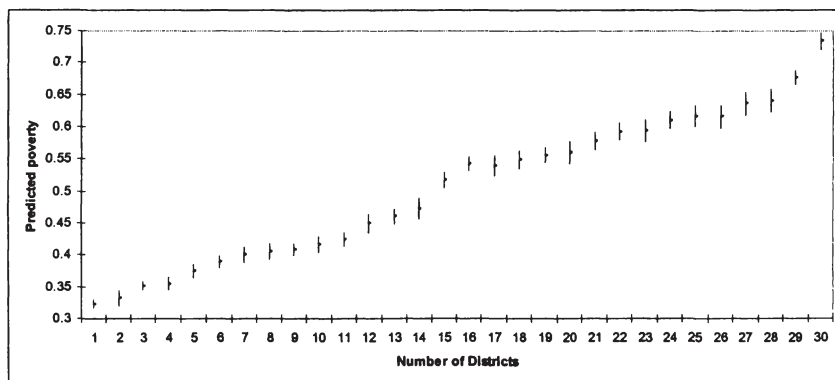
Source: Author's calculations based on ECV 1992, CNPV 1992, EPH 2002 and CNPV 2002.

The maps in Figure 1.5 show the heterogeneity of poverty levels in small areas, using the Itapua department example. Itapua is one of the most prosperous departments in Paraguay, considering its economic performance (important GDP growth driven by mechanized Soya agro-industry) between 1992 and 2002. Re-

garding departmental mean poverty levels in 2002, Itapua ranked 11 out of 18 departments, about 1 percentage point above the national poverty rate, despite its high GDP growth. What is behind this apparent contradiction can be seen by mapping Itapua's poverty levels by district. In 1992, there was a belt of poor, up to extremely poor districts in the north of the department (darker colors), with poverty rates of up to 68%, versus a more prosperous zone in the south (brighter colors) with poverty rates down to 14%. Until 2002, and as a consequence of GDP growth, poverty generally decreased in Itapua (more districts with bright colors), but the poor district belt in the north still remained. Some of these poor districts even increased their poverty levels, now up to 72% of the population. The lowest levels in the south were 29% in 2002. So, considering a pro poor policies intervention, maybe by conditional cash-transfer programs to extremely poor households, Itapua might not be selected for program participation, considering department mean levels of poverty. Nevertheless, disaggregating poverty estimates by districts, it happens to be that some of the poorest districts of the country are located in Itapua, being direct neighbors of some of the most prosperous districts in the country. Small welfare estimates help to improve the targeting of pro-poor policies.

To show an example of what precision can be achieved at the district level, Figure 1.6 shows the predicted poverty headcount in rural Itapua for 2002, along with a confidence interval from one standard error below to one standard error above the point estimate. The department of Itapua was selected because covers almost the complete range of standard errors for point estimates observed for the 1992 and 2002 exercise, varying from 0.015 to 0.075.

**Figure 1.6 Rural poverty estimates Itapua 2002**



Source: Author's calculations based on EPH 2002 and CNPV 2002

Apart from the standard errors for point estimates, regression models have structured and unstructured errors as seen above. To check for spatial patterns of these kinds of errors they are mapped in the annex, as well as relative changes in poverty and inequality.

As mentioned above, there is a set of former poverty map exercises in Paraguay. Robles (2000) combined the 1992 census with 1997/98 survey using the Hentschel et al (2000) method, which differs slightly from the method applied in this paper, so they can not easily be compared. The second poverty map exercise carried out by Otter (2003), using the same method applied in this paper, combines the 2002 survey with a 10% sub-sample of the census and combines the estimated poverty levels with weighted unmet basic needs percentages. Since the data bases are not the same, there may be some difficulties in comparing the results of this paper with the Otter (2003) exercise. Finally, Robles and Santanders (2004) poverty map exercise is most similar to this paper. Using the same method, they combine 2002 census with 2003 household survey data, mostly because the 2003 survey sample allows to run a separate regression model for every department (18 models) and not only 4 different models by region as in 2002. Since poverty rates changed considerably (dropping by 6 percentage points in the national mean) between 2002 and 2003, the best way to compare the results of these two exercises is to compare rankings of districts by poverty level, which should not change strongly, even if poverty percentages decrease considerably. When comparing the rankings, we observe that 64% of all districts are ranked within the same deciles, meanwhile the standard deviation of ranking differences is only 0.94 points. Consequently, the results of both poverty mapping exercises are consistent between each other, and differences should be a consequence of the more detailed estimates by Robles and Santander and poverty changes between these two years.

### 1.3.4 Pro-poor growth evidence

Although this paper is not about pro-poor growth, its results provide empirical evidence on such growth from poverty map exercises. Even if the existence of pro-poor growth evidence could easily be confirmed from the household survey data, doing this with simulated incomes based on census data will allow the identification of whether there are any spatial patterns in pro-poor growth, for example the concentration of a huge number of households benefiting from pro-poor growth which could be concentrated in a small and limited geographic area.

According to international organizations pro-poor growth is simply defined as economic growth that benefits the poor (e.g., UN 2000a; OECD 2001). This

definition, however, provides little information on how to measure or how to implement it. What remains to be specified are, first, whether economic growth benefits the poor and, second, if this is the case, to what extent. Klasen (2004) provides more explicit requirements that a definition of pro-poor growth needs to satisfy. The first requirement is that the measure differentiates between growth that benefits the poor and other forms of economic growth, and it must answer the question by how much the poor have been benefited. The second requirement is that the poor must have benefited disproportionately more than the non-poor. The third requirement is that the assessment must be sensitive to the distribution of incomes amongst the poor. The fourth requirement is that the measure must allow an overall judgement of economic growth and not only focus on the gains of the poor.

To identify the existence of pro-poor growth of per capita income according to point estimates from the poverty map exercise, specific inflation rates by region and income deciles were calculated. To obtain these measures as realistic as possible they are based on consumption profiles by deciles, built up as a mean of 1997/98 and 2000/01 consumption profiles<sup>12</sup> (only during these two periods did Paraguayan household surveys include a consumption module). Table 1.14 shows the deciles specific inflation rates. In general, inflation is lower for lower income deciles and lower for less urban or more rural areas. These results seem to drive the specific results of pro-poor growth.

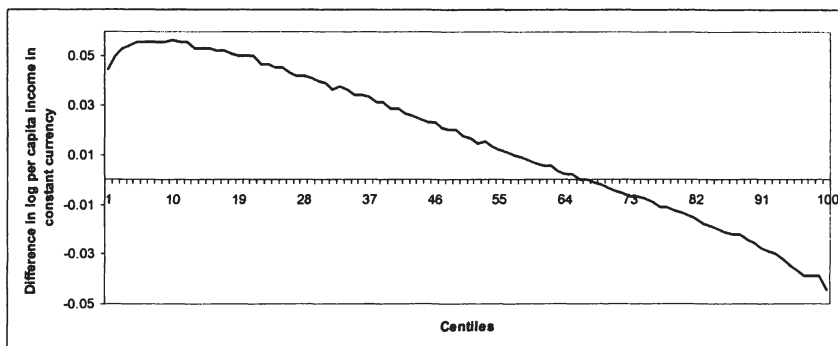
**Table 1.14 Mean inflation rates by deciles, Paraguay 1997/98 – 2000/01 (%)**

	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
ASU	112.5	115.7	115.2	115.3	116.4	117.5	121.7	122.1	123.5	123.5
CU	114.4	114.9	114.4	114.7	115.3	115.7	117.7	118.0	120.7	120.5
RU	111.1	111.1	112.8	111.9	112.6	115.5	114.5	116.9	116.4	119.5
RUR	107.5	108.0	108.2	108.5	109.4	110.1	110.3	111.0	111.8	113.6
TOTAL	111.4	112.4	112.7	112.6	113.4	114.7	116.1	117.0	118.1	119.3

Source: Author's calculations based on EIH 1997/98 and EIH 2000/01.

Figure 1.7 shows the growth incidence curve of log per capita income for constant currency in 1992 values. There is a clear pro-poor growth pattern for deciles 5 to 25. Growth incidence curves produced separately for the four different regions show that there is almost no pro-poor growth in the Asuncion and Central Urban areas, very few in Remaining Urban, but mostly in rural areas.

12 Carried out by groups of goods and services: food, clothing, housing, health, transport, education, various.

**Figure 1.7 Growth incidence curve of log per capita income**

Source: Author's calculations based on ECV 1992, CNPV 1992, EPH 2002 and CNPV 2002.

## 1.4 Discussion

In the regression results, all adjusted R squared are not very high; approximately 0.6. On the one hand, this is due to the fact that only variables whose coefficients are highly significant were included in order to make the models as robust as possible, as explained above. On the other hand, reduced levels of adjusted R squared may result from a considerable number of dummy variables in the models, which may have reduced the power of explanation for probably having lower variance than other kind of variables.

In regression results, it seems that the less homogeneous a population is, the higher the probability of identifying locational components in the error term. If there are locational components in the error term, it is easier to identify them in smaller geographic areas (this is why the Asuncion model produces more significant cluster effects than the other regions in 1992).

If this is the case, the question is now why the 2002 models produce much less significant cluster effects? Could it be true that the population became more homogeneous? Even if poverty and inequality changes between 1992 and 2002 are not big, the real story household survey data tell us that there was a poverty reduction between 1992 and 1997 and an increase between 1998 and 2002. Urban inequality tended to decrease while rural inequality tended to increase over the whole period. Additionally, we have a growing urban migration of poor and a growing urban poverty.

There are two main differences between the 1992 and the 2002 results. First, the 1992 results include much more significant cluster effects than those for 2002,

which capture part of a locational effect. In 2002, this fact seems to have disappeared. Noticeably, the locational effect is not only or directly related to the geographic location, but is in relation to the population group and their characteristics, living in the observed area, at any moment. Even if the 2002 regression models include less variables and less significant cluster effects, their prediction power for poverty is higher than 1992 while their prediction power for inequality is almost the same as in 1992. Consequently, geographic location seems to be less important for 2002 than for 1992.

Although the income regression models are not modeling determinants, but variables which are correlated with income, most of these have the expected signs. This is also true for the cluster means. The interpretation of these cluster means, which try to capture part of the locational error, is difficult. Nevertheless, for some cluster mean variables there can be a kind of intuitive understanding. For example, in some models the percentage of primary sector employment appears as a significant variable with a negative sign for the household (individual) level. However, its cluster mean also has a positive sign. This may be understood as a positive effect at community level *ceteris paribus* and for the given mean level of income in the region.

At least concerning empirical evidence from Paraguay, the methodology seems to work better for the prediction of higher incomes, since extreme poverty is overestimated (by underestimating lower incomes), as are mean incomes. Consequently, the associated distributions are not the same as those observed in household surveys. Several reasons can be attributed for this. For example, rural incomes (where most of the low incomes are located) depend strongly on climate and market price changes not captured by variables included in a census. Lowest incomes in urban areas may be difficult to simulate correctly due to sampling and measurement problems in household surveys.

Poverty maps show that there is a concentration of poverty in the center of eastern Paraguay (where 98% of the population is living). Changes of poverty during the observation period neither altered significantly the spatial distribution of poverty nor of inequality. In general, structured and unstructured errors are higher in more rural areas. All these results are consistent with Paraguayan rural economic history over the period, with a crisis of small scale cotton cash crop farming and an increase of large scale Soya bean mechanized farming, deepening poverty and inequality in rural areas.



## 1.5 Conclusions

As shown, the method of Elbers et al is a reliable method for small area welfare estimates, producing poverty point estimates for sub-national levels. Obviously, there are several sources of errors in the methodology and other errors from sampling and measurement problems in the household surveys. Nevertheless, the income estimates are consistent with economic history in Paraguay and, since most of the errors made during the estimation procedure can be quantified, it is possible to determine their reliability. In any case, the gain in additional information is crucial for politics and policies design and implementation.

Poverty analysis is often based on national level indicators that are compared over time or across countries. The broad trends that can be identified using aggregate information are useful for evaluating and monitoring the overall performance of a country. For many policy and research applications, however, the information that can be extracted from aggregate indicators is not sufficient, since these hide significant local variation in living conditions within countries.

The detailed poverty maps of small administrative areas, that are the ultimate output of the simulated-welfare mapping method, provide benefits that help address the shortcoming of aggregate poverty analysis in the following ways:

*(i) Poverty maps capture the heterogeneity of poverty within a country.*

Almost all countries in the world have regions that are better off and others that are left behind. Such differences are often lost in national level statistics. Poverty maps can reveal the variation in local poverty levels when small area information is available. As shown, seemingly homogeneous regions can actually have a large degree of local heterogeneity.

*(ii) Poverty maps improve targeting interventions.*

In designing poverty reduction programs, resources can be used more effectively if the most needed groups can be better targeted. This reduces the leakage of benefits from a poverty reduction program to non-poor households and, on the other hand, reduces the risk that poor households will be missed by a program. This requires an adequate targeting to poor areas, but also a correct beneficiary selection.

*(iii) Poverty maps can help governments – national and local – to articulate their policy objectives.*

Basing allocation decisions on observed geographic poverty data, rather than subjective rankings of regions, increases the transparency of government decision making. Such data can thus help limit the influence of special interests in

allocation decisions. There is a related role for well-defined poverty maps to lend credibility to government and donor decision-making. By increasing transparency, poverty maps can help prevent the regional autonomy policy from being hijacked by the local elite.

*(iv) Poverty maps have an important role in communicating information on welfare distribution to the civic population within a country.*

Poverty maps are not only useful to governments and decision makers, but also to local communities. Compiling disaggregated information on human welfare generates locally relevant information. This provides local stakeholders with the facts that are required for local decision making and for negotiation with government agencies. Poverty maps thus become an important tool for local empowerment and decentralization.

*(v) Poverty maps are useful for evaluating the impact of various programs.*

Poverty maps offer opportunities to undertake detailed empirical research on the causal relationships between local poverty, income inequality, and various other social outcomes, both at the individual and community levels. Until now, scarcity of welfare indicators for small areas has prevented researchers from studying the relationship between various programs, poverty, inequality, and various outcomes, such as health, education, crime, and the environment. Poverty maps open up more opportunities for researchers to examine these relationships.

*(vi) Estimation of small area indicators of poverty allows their incorporation into geographical information systems (GIS).*

This feature of poverty maps facilitates the combination of poverty information with other indicators from policy-relevant subject areas. Examples are geographic databases of transport infrastructure, public service centers, access to input and output markets, or information on natural resources quality and vulnerability. Using geographic overlay techniques and spatial analysis methods, the newly constructed databases on poverty can thus be used to address a range of multidisciplinary questions. The databases can also be used by the private sector to guide them in determining the locations for new investment opportunities.





## **Chapter 2**

### **Does Inequality Harm Income Mobility and Growth?**

#### **An Assessment of the Growth Impact of Income and Education Inequality in Paraguay 1992 – 2002**

##### **2.1 Introduction**

Latin America is the most unequal region of the world in terms of income or expenditure, as well as regarding other aspects of economic or social exclusion. The region suffered the lost decade of the nineteen eighties, and experienced a modest recovery in the nineteen nineties. In the nineteen nineties, most of the governments implemented stabilization politics, more or less close to the proposals of the Washington Consensus. Paraguay itself, however, neither suffered a debt crisis nor a mayor economic instability during the eighties, so the stabilization policies would not have been necessary or useful for the Paraguayan economy and business cycles in the nineties. Nevertheless, many of the macro-economic policies applied in Paraguay during the nineties were close to the Washington Consensus. The most striking macroeconomic result of the decade was a per capita income decrease beginning in late 1995, hand in hand with a poverty increase after 1996. Given the persistently high levels of poverty incidence in Paraguay to date, understanding the determinants of growth at the household level in Paraguayan economy remains an important but under-researched field in economics. This appears to be particularly true for the question whether inequality has a positive or negative effect on economic growth, a question that is both fundamental in (development) economics and highly relevant for poverty reduction policies. Although the effect of inequality on growth has important implications for poverty (Bourguignon, 2004; Ravallion, 1997), empirical evidence on this link is virtually inexistent for Paraguay.<sup>13</sup>

The effect of inequality on economic growth is the subject of a large literature. Aghion et al., 1999 and Thorbecke and Charumilind, 2002 review this literature

---

13 Different country analysis on aspects of financial liberalization and openness were run by a research team supported by CEPAL, UNDP and IADB. These studies include Paraguay, focussing on CGE simulation models and their counterfactual effects on households, but these analyses do not consider the effect of inequality on growth (Ganuza, Morley and Taylor 1998; Ganuza, Paes de Barros, Taylor and Vos 2001; Ganuza, Morley, Robinson and Vos 2004).

and show that theory does not provide firm predictions of the sign of the effect.<sup>14</sup> Empirical studies in the 1990s have been “.. impressively unambiguous ..” (Aghion et al., 1999, p.1617) in concluding that the growth effect of inequality is negative, but more recently some authors have obtained contrasting results (e.g. Forbes, 2000). The most common denominator in these studies is the nature of the data used: the empirical inequality-growth literature is largely based on cross-country data.

This paper contributes to the existing inequality-growth literature by providing empirical evidence that is new in a number of ways. First, the present study is based on micro data for Paraguay. This allows avoiding data comparability problems that affect cross-country studies (see Section 2). While there are a small number of inequality-growth studies using micro data (for example Joeng, 2001, Schipper and Hoogeveen, 2005), this is the first such study for Paraguay. Second, the data used consists partly of the so-called small area welfare estimates, which are obtained by combining information from a census and a survey. For this paper the small area welfare estimates were grouped in a pseudo panel. Third, theoretical and empirical studies have been criticized for their focus on income or expenditure inequality as the determinant of growth. Birdsall and Londono (1997) show that once land and human capital inequality are entered in a cross-country growth regression, income inequality no longer has a significant effect on growth. Elbers and Gunning (2004) address this issue theoretically using a Ramsey type household growth model and show that growth is affected by ‘underlying’ inequalities in assets, abilities and shocks. In particular, these authors show that higher ‘ability’ (human capital) inequality will positively affect growth if the production function is convex in ability. In that case, a mean-preserving spread in human capital results in a higher mean steady state level of output, and therefore in higher growth. In this paper, this issue is explicitly addressed by estimating the growth effect of inequality in human capital. The results indicate that it is income inequality rather than human capital inequality that affects growth and that this effect is negative. Nevertheless, there are also positive growth effects of human capital inequality, some less strong than income inequality results.

---

14 Positive inequality-growth effects can be attributed to a positive effect on savings, to the existence of investment indivisibilities or to positive incentive effects of inequality. A negative inequality-growth effect can be explained by political tension, instability and demands for redistribution due to inequality, by reduced investment opportunities for the poor, worsened borrowers’ incentives and by higher macro-economic volatility. A ‘unified’ model that aims to reconcile these conflicting effects is presented in Galor (2000); this paper predicts that the effect of inequality on growth is non-linear, with a positive effect at an ‘early stage of economic development’ and a negative effect at a ‘later stage’.

The paper is organized as follows. Section 2 discusses the existing empirical inequality-growth and some of the income mobility literature. Small area welfare estimates as an alternative source of data for this type of investigation are briefly described. In Section 3 the growth model and descriptive statistics are presented. Section 4 presents a discussion of econometric issues that need to be addressed given the model in use and, in particular, given that some of the variables have been imputed using small area welfare estimation. Results are presented in Section 5 and conclude in Section 6.

## **2.2 Data: macro, micro and small area welfare estimates**

Cross-country inequality-growth studies, while providing considerable empirical evidence, have been criticized for various reasons. A general problem with both macro (cross-country) and micro growth studies is the ‘open-endedness’ of the underlying theory: many variables potentially affect growth and theory may often not give clear guidance as to which specification is preferable. Data used in cross-country studies are national aggregates that are likely to lose valuable region or gender specific information (Deininger and Okidi, 2003). Brock and Durlauf (2001) reject causal interpretations in cross-country studies except under considerably exceptional conditions. Their main argument is that causal interpretation requires that estimated parameters can be assumed constant, which is not plausible given the importance of country-specific unobserved information (e.g. regarding policy).

Comparability of variables that are intended to measure the same concepts across countries is a further issue in cross-country studies. This is particularly problematic for cross-country inequality data (Atkinson and Brandolini, 2001). An issue that has not received much attention in the literature is that, even when variables are defined and measured in exactly the same way, national statistics (including GDP) are often estimates derived from, for instance, national household surveys – as is the case for inequality estimates. Even if these estimates are representative at the national level, they are still point estimates with a standard error, a fact that the analyst has to take into account when doing regressions: one should expect that properly accounting for the uncertainty with respect to these estimates, reflected by their standard error, translates into higher standard errors in the growth regression coefficient estimates. This problem is equivalent to the one encountered when using small area welfare estimates in regression analysis, as is discussed in detail in Section 4.

A problem with household data is that only surveys for very large countries provide sufficient data points to meaningfully include inequality indicators in a regression while census data typically do not provide the income or wealth

variables and covariates needed in a growth regression. As a result, only a small number of inequality-growth studies that use micro or regional data remain. Ravallion, 1998 estimates a linear household level growth model with local externalities and finds a significant negative effect of inequality for rural China. Balisacan and Fuwa, 2003, find a positive effect of inequality on provincial level growth for the Philippines, using a linear model. Schipper and Hoogeveen (2005), using downstream regressions for Uganda, found that it is human capital inequality rather than income inequality that affects growth and that the effect is positive.

An important advantage of regional or household data is that comparability problems are much less severe than in cross-country datasets: the definitions of variables or phrasing of survey questions are generally uniform across regions for a given dataset. Depending on the level of desegregations, regional analyses may also be able to use larger numbers of observations than cross-country analyses; household growth studies are especially advantaged in this sense.

Until recently, the unavailability of nationwide inequality data covering a larger period precluded the study of the inequality-growth relation for Paraguay<sup>15</sup>. However, the application of welfare estimation techniques for small area target populations has recently provided income estimates for all households in Paraguay for 1992 and 2002 (see Chapter 1). This now allows the study of the inequality-growth relation for Paraguay.

### 2.3 Small area welfare estimation

Part of the data used for this paper is derived using small area welfare estimating techniques first described in Hentschel et al., 2000 and refined in Elbers et al., 2003; the latter paper is referred for details of the technique and provide a brief review below.

Small area welfare estimation combines data from a census and a household survey in a three-stage process. First, a set of variables that are common to the survey and the census are identified. Second, household per capita expenditure is regressed on these common variables using the household survey data and census means obtained for the clusters from which the survey households originate; this yields coefficient estimates with the associated variance-covariance matrix and estimates of the distribution of household and cluster error terms. Third, out

---

15 The first nation wide inequality estimates in Paraguay are based on the household survey of 1992 (carried out by IADB and the National University). Only as from 1998 does the National Statistics Bureau (DGEEC) provide annual updates of nation wide household surveys.

of sample prediction on unit record census data is used. Predicted values are calculated typically 100 times, each time drawing variable coefficients and household specific and cluster level error terms from the relevant distributions. This yields, for each household in the census, predicted per capita income and its standard error. A close correspondence between census and survey household characteristics is needed to obtain reliable welfare estimates.<sup>16</sup> For this reason, small area welfare estimates have typically only been generated for the years close to a census year. Hoogeveen et al. (2003) show how, in the presence of panel survey data for which one of the waves was collected at the time of the census, the welfare estimates can be updated by associating household characteristics collected during the census year, with expenditures obtained for a more recent period. Since panel surveys do not exist for Paraguay, and since the analysis of the present paper is based on two different censuses, the inequality and growth analysis is based on a pseudo panel build up from income estimates for each household in each census.

More formally written small area welfare estimates can be estimated by using the following model:

$$\ln y_{cs, t+1} = E[\ln y_{cs, t+1} \mid X_{cs, t}] + \eta_{c, t+1} + \varepsilon_{cs, t+1} \quad (2.1)$$

where subscript  $t$ , survey households is represented with subscript  $s$ , census households is represented with subscript  $h$ , and the cluster from which census and survey households originate is represented with subscript  $c$ .

Predicted log per capita expenditure is now derived, for each household in the census, from:

$$\ln \tilde{y}_{ch, t+1} = X_{ch, t}^T \tilde{\beta} + \tilde{\eta}_c + \tilde{\varepsilon}_{c, t+1} \quad (2.2)$$

and welfare estimates are based on:

$$\tilde{\mu}_{t+1} = E[W_{t+1} \mid m_t, \tilde{y}_{h, t+1}] \quad (2.3)$$

---

16 Much attention is therefore devoted to identifying common variables by assuring that variable definitions are identical between the census and the survey, that questions are phrased the same way, that coding and enumerator instructions are identical and that the survey and census are fielded contemporaneously. When the latter condition is not met -and this is more of a problem in rapidly changing economic environments-, changes in the economic situation will be reflected in household characteristics. As a result, survey variables identified as common to the census, are actually not representative of the census and small area welfare estimates can not be derived.

Once predictions are made using (2.2) welfare estimates can be generated for any administrative unit, but their precision decreases with the degree of disaggregation. For Paraguay, accurate welfare estimates coming from household surveys are available for three levels; nation wide, by urban or rural area and by region.<sup>17</sup>

Our analysis makes use of two data sets: unit record data from Paraguay's 1992 population census, combined with 1992 household survey. Small area welfare estimates for all households are carried out. The same exercise is carried out with the 2002 population census and the 2002 household survey. The 1992 census was carried out in August 1992 and covers 526,050 urban households and 454,342 rural households. The 1992 household survey was carried out between October and December 1992. The 2002 census was carried out in August 2002 and covers 782,966 urban households and 505,567 rural households. The 2002 household survey was carried out during November and December 2002. Both censuses comprise, for all household members, information on household composition, ethnic background, marital status and educational attainment. Growth and inequality variables are calculated using the income values prepared for Chapter 1. The author shows that the income estimates for 1992 and 2002 are unbiased and closely correlated estimates of the 'true' welfare estimates derived from the national household surveys. Estimates of income and inequality were derived for all 224 districts<sup>18</sup> of Paraguay for both years. Based on comparable household income data, they represent the first data set for Paraguay with comparable inequality estimates for two points in time for a substantial number of observations.

**Table 2.1 Welfare estimates, Paraguay, Selected Years**

Region	Poverty ratio (per capita income)		Poverty change	Inequality change (Gini)	Income growth
	1992	2002			
Asuncion	0.234	0.298	0.274	-0.097	-0.077
Central Urban	0.401	0.577	0.389	-0.102	0.066
Rem. Urban	0.469	0.405	-0.136	0.042	0.065
Rural	0.495	0.543	0.097	-0.016	0.082
National	0.443	0.485	0.095	-0.059	0.058

Note: column entries are regional means of district estimates.

Source: Author's calculations based on results in Chapter 1

17 That is, the ratios of mean values to standard errors are about the same as those obtained in household surveys.

18 The Paraguayan "distrito" is a municipality, the smallest existing administrative unit.



A summary of the welfare estimates used in this paper is presented in Table 2.1. The Table confirms that on average poverty increased over the 1990s, except for the Remaining Urban region. Also, the increase in poverty was not distributed uniformly. Asunción and Central Urban were the most affected regions. At the same time, inequality decreased where poverty increased and vice versa. Mean income increased in all regions except Asunción. Even if this seems contradictory it is consistent with the macroeconomic history of Paraguay over the decade, with a growth period and poverty reduction until 1997. During this period, in general, income increased and inequality decreased. In the following period of recession (1998 to 2002) characterized by income decrease, not all of these mean income increases and inequality decreases were lost. Nevertheless poverty rose by means of the appearance of an important number of “new poor”.

At first, it seems to be contradictory that we observe a simultaneous income growth associated with a poverty increase and an inequality reduction. Poverty can increase despite income increase, so long as prices grow quicker than income (so poverty lines rise faster) or whether there are any other problems with the poverty lines, as such. In Paraguay, poverty is defined by four different poverty lines for the Asuncion, Central Urban, Remaining Urban and Rural areas. Official poverty lines are updated yearly by an official inflation measurement that is limited to the Asuncion and Central Urban areas. To apply this inflation data to the other two areas, an implicit Engel coefficient based on a consumption profiles measurement not updated since 1998, is applied. This methodology seems to create some bias in the poverty lines. The inequality decrease associated with poverty increase results from general income loss after 1998, where higher income groups suffer stronger losses than lower ones, resulting in decreasing inequality (more on this in chapter 3).

Since pseudo panels are used for the analysis, the results could also be interpreted as an indicator for income-mobility, since the growth rates of estimated mean household per capita income between 1992 and 2002 at a district level are used as dependent variables. However, since education inequality is used as one of the independent variables, we also have notions of human capital in the analysis. This brings the results close to the link between growth, inequality and social mobility.

One of the primary motivations for economic mobility studies is to gauge the extent to which longer-term incomes are distributed more or less equally than are single-year incomes. Krugman, for instance, stated: “If income mobility were very high, the degree of inequality in any given year would be unimportant, because the distribution of lifetime income would be very even (...). An increase in income mobility tends to make the distribution of lifetime income more



equal” (Krugman, 1992). Similar statements have been made by Shorrocks (1978), Atkinson, Bourguignon, and Morrisson (1992), Slemrod (1992), and Jarvis and Jenkins (1998).

Social mobility and income inequality together describe the “fairness” of an income distribution. If income is very unevenly distributed and social mobility is low, then there is a large gap between rich and poor and there is little chance of crossing that gap. However, since social mobility might be related to education, who has more mobility, better-educated individuals or less-educated people? The answers may depend on the mobility concept used. In the *intergenerational* context, the recipient unit is the family, specifically a parent and a child. In the *intragenerational* context, the recipient unit is the individual or family at two different dates. The pseudo panel used in this paper refers to an intergenerational model, but the observation period is not a whole generation, but only a ten year difference.

The literature distinguishes between six notions of mobility (Fields et al 2006, Scott and Lichtfield 1994). Briefly, they are: *time-dependence*, which measures the extent to which economic well-being in the past determines individuals' economic well-being at present; *positional movement*, which is what is measured when looking at individuals' changes in economic positions (ranks, centiles, deciles, or quintiles); *share movement*, which arises when individuals' shares of the total income change; *income flux*, which is what is gauged when looking at the size of the fluctuations in individuals' incomes but not their sign; *directional income movement*, which is what we measure when we determine how many people move up or down per amount of dollars; and *mobility as an equalizer of longer-term incomes*, which involves comparing the inequality of income at one point in time with the inequality of income over a longer period. If the results of this paper might be understood as an income mobility indicator, the study belongs in part to time dependence (because it considers initial levels of income and education inequality) and in part to positional movements (because it asks if there was some pro-poor growth).

Several papers show how the allocation of talent in an economy is important for the level of growth. Murphy, Shleifer, and Vishny (1991), for example, show that when talented people are attracted to the productive sector, they create high growth, but if they instead are attracted to rent seeking activities, they create stagnation. However, the use of talent needs the opportunity to be developed and exposed by a formal educational process.

Two papers have theoretically analyzed the relationship between social mobility and economic growth (Raut 1996; Hassler and Mora 1998). They both arrive to

the conclusion that high social mobility is associated with higher economic growth, but the direction of causality and the transmission mechanisms between mobility and growth differ slightly between the models. Raut (1996) develops a signaling model of endogenous growth in which innate talents and education levels of workers drive the basic scientific knowledge accumulation in the economy. The second study is by Hassler & Mora (1998). They analyze an economy with two types of individuals: workers and entrepreneurs. Entrepreneurs are those who generate new ideas and new technologies and make the economy grow. The more intelligent the entrepreneurs the higher the growth rate of the economy.

The implication of the above mentioned studies is that to achieve optimum growth it is important that people get the opportunity to work in the sectors where they are most productive. This requires that young people's educational and occupational choices be determined by talent and not limited by family background. Linking these ideas to the model used in this paper, initial income level could be a proxy for family background and initial education level as institutional opportunities to develop talent (which is supposed to be distributed randomly, in spite of the fact that educational levels are usually strongly determined by family background).

## 2.4 The model

For estimating yearly per capita income growth over the period 1992 – 2002 we build up a pseudo panel at the district level, to be able to compare 1992 and 2002 results. The pseudo panel takes into account the age of the household head (3 year steps), his years of schooling (3 year steps), his mother tongue (as a proxy for ethnicity), the district of residence and the condition of migration (only non-migrant households are included).<sup>19</sup> Groups with common characteristics in 1992 and 2002 with more than 29 observations were considered for the model. Only non-migrant households entered the model. This is, on the one hand, because migration is not an important phenomenon over the

---

19 The idea of excluding migrant households is based on the fact that, even if a pseudo panel is used for this exercise, it is still possible to identify locational or district effects. A "pure" district effect would only be found when considering non-migrant households, even if there are also arguments for including them, such as "pull and push" factors that make a certain district more or less attractive. Either way, migration levels in Paraguay over the nineties were not huge (only 8% of population older than 15 years moved from one department to another between 1992 and 2002, and only 75% of these are non-poor) (Otter, 2007).

whole period<sup>20</sup> and, on the other, to analyse growth determinants we can focus on the change of real conditions in each district, which are not biased by changes due to migration. Final estimates were carried out for five different models<sup>21</sup>; Asunción (471 panels), Central Urban (655 panels), Remaining Urban (762 panels), Rural (2388 panels) and pro-poor-growth households (1300 panels) sample which includes all groups from any region living below the poverty line but having experienced positive income growth. The purpose of this pro-poor-growth panel is to identify if there are any spatial patterns related to the geographic location of pro-poor growth. A separate panel for pro-poor-growth additionally allows us to identify if there are differences in household, family group or household heads characteristics between poor households with and without income growth. Nevertheless, this last step of the analysis was not carried out in this paper.

Estimate growth effects using a pseudo panel can eventually be problematic. All households in a panel, even if they are different between each other, have to observe the same panel mean income change; this can cause problems of heteroskedasticity. Even if all households grouped together ought to be similar, some differences still remain. Not all sources of heteroskedasticity can or should be captured via a relationship with an independent variable. For example, using grouped data leads to heteroskedasticity if the groups are not all the same size. In this case the error variances are proportional to the group sizes. Using weighting factors could be a solution for this problem. In our case, households are the elements composing panel groups. Every household enters the panel with “size one”, since characteristics of the household head are used as grouping criteria. Since this paper uses census data, no weighting factors are used. All size differences between groups reflect reality and should be taken as such since all households in the country are considered (only migrant households are left out of the analysis). For all five models to be run, panel groups contain between 30 and up to 1000 households. Nevertheless, in all cases, panels including between 30 and 250 households cover more than 90% of all observations. The distribution of these panel groups by size is almost the same. So if there is a heteroskedasticity problem caused by different panel sizes, it would be a systematic one.

In the model we estimate yearly per capita income growth of each panel group, over the period 1992 – 2002 as a function of, for 1992, per capita income, income inequality, human capital inequality, male and female human capital

- 
- 20 As a result, there are very few or no panels by district which fulfill the conditions of identical characteristics and more than 29 observations in the panel.
  - 21 Since income estimates in Otter (2006) were carried out for four different regions, each of these with its own poverty line, the growth analysis is based on the same regions as well as a growth analysis for poor households.

household demographics and employment sector. The model we estimate can be represented as:

$$g_{i,t} = (\tilde{y}_{i,02} - \tilde{y}_{i,92}) / 10 = \hat{y}_{i,92}\beta_1 + \tilde{I}^{exp}_{i,92}\beta_2 + \tilde{I}^{edu}_{i,92}\beta_3 + \mathbf{X}_{i,92}\gamma + \alpha_d + u_i \quad (2.4)$$

With the exception of the Gini coefficients, which are district averages, all other values are averages by panel  $i$ :  $g$  is the annual income growth rate between 1992 and 2002;  $y$  is the logarithm of per capita income;  $I^{exp}$  is the Gini coefficient for per capita household income;  $I^{edu}$  is the Gini coefficient for the number of years of formal education of the household head.  $\mathbf{X}$  is a matrix of other covariates consisting of human capital (number of years of formal education entered separately for household heads and for spouses), head age, gender of the household head, logarithm of the number of individuals in each household, number of children and dummy variables for employment sectors, changes of some of these variables (which are likely to be endogenous) and some departmental dummies.<sup>22</sup> Given this approach, we are limited in our choice of covariates in  $\mathbf{X}$  to what the census has to offer. District fixed effects, represented by  $\alpha_d$ , to control for unobserved spatial heterogeneity;  $u_i$  is an error term used.

A non-standard econometric issue lies in the fact that some of the variables are not observed but imputed as described in Section 3. The imputed variables, income growth and income inequality, are denoted using tildes. See Table 2.2 for definitions and summary statistics.

An important issue in regional growth studies is the measurement of the dependent variable. In our case, the smallest available geographical subdivision in the database is the district, and within the district, households are grouped in panels. Growth for a panel  $i$  is usually specified as:

$$gr_i = \frac{y_{i,t} - y_{i,0}}{t} \quad (2.5)$$

where  $y$  is a measure of panel income or expenditure. This measure is often specified as the logarithm of the mean of per capita income over households  $h$  for group  $i$  (e.g. in Balisacan and Fuwa, 2003), i.e.:

$$y_i = \ln \left( \frac{\sum_{h=1}^H y_{h,i}}{H} \right) \quad (2.6)$$

---

22 Potential changes in employment sectors could be considered proxies for structural changes in the productive sector.

However, as pointed out by Ravallion, 1998, the use of the logarithm of mean expenditure rather than the mean of log expenditure introduces a measure of the change in inequality in the error term of the regression equation. The argument is as follows: a general inequality measure is

$$I(y_i) = \ln M(y_i) - M(\ln y_i) \quad (2.7)$$

where  $I(.)$  is an inequality measure and  $M(.)$  denotes an average. Rearranging these terms we have:

$$\ln M(y_i) = M(\ln y_i) + I(y_i) \quad (2.8)$$

The LHS of (2.8) is the income of (2.6). However, if we think that the log of household income is the variable of interest we should use:

$$y_{i,t} = \frac{\sum_{n=1}^N \log(y_n)}{N} \quad (2.9)$$

which is the first term in the RHS of (2.8). It is clear from (2.8) that we introduce a measure of inequality if we use the log of mean incomes as our regional income variable. Consequently, we introduce as measure of the change in expenditure inequality in our growth variable if we calculate mean expenditure using (2.6).

In an inequality growth regression, this is likely to introduce a correlation between the error and the inequality variable which will affect estimates through omitted variable bias. For example, consider the case where increases in inequality have a negative effect on growth, while the level of (initial) inequality has a positive correlation with the change in inequality. Then omitting the change in inequality will cause a spurious (negative) effect of inequality on growth (Ravallion, 1998). Since we have access to household level per capita income estimates aggregated by pseudo panels, it could be useful comparing the estimates of a growth regression using both types of dependent variable (mean-log(exp) and those using log-mean(exp)). Nevertheless, this comparison is still pending and has not yet been carried out. In this paper, only the mean log income is used in the regression models.

**Table 2.2 Variables and descriptive statistics - Asunción**

Variable	Definition	Mean	Standard error	Minim.	Max.
gry	Annual growth of log per capita inc, 1992-2002: $\text{Ln}(\text{pcy}_{92}) - \text{Ln}(\text{pcy}_{02})/10$	-0.01	0.02	-0.06	0.05
lny92	Log income per capita 1992: $\text{Ln}(\text{pcy}_{92})$	12.39	0.55	11.37	13.62
hhedu92	Household head's education 1992, number of years	8.77	5.31	1.30	17.73
dlntot	Changes in log total individuals per household: $(\text{Intot}_{02} / \text{Intot}_{92}) - 1$	-0.02	0.08	-0.29	0.29
dsedu	Changes in spouse's education in number of years	0.76	0.84	-0.11	6.78
desp	Changes in household head's primary sector employment percentage: $(\text{esp}_{02} / \text{esp}_{92}) - 1$	0.40	0.29	-0.43	0.63
dhijo	Changes in number of children: $(\text{nhijo}_{02} / \text{nhijo}_{92}) - 1$	0.52	0.58	-0.17	6.19
dhedu	Changes in household head's education in number of years: $(\text{hedu}_{02} / \text{hedu}_{92}) - 1$	-0.07	0.06	-0.16	0.04
hage92	Age of household head 1992	46.83	16.14	21.32	82.06
Intot92	log total individuals per household 1992	1.46	0.19	0.65	1.79
nhijo92	Number of children 1992	1.66	0.65	0.18	3.20
hsedu92	Spouse's education 1992, number of years	5.75	3.29	0.72	13.69
hfem92	Percentage of female head 1992	0.32	0.13	0.06	0.82
dest	Changes in household head's tertiary sector employment percentage: $(\text{est}_{02} / \text{est}_{92}) - 1$	-0.03	0.05	-0.10	0.03
dginiy	Changes in income inequality: $(\text{giniy}_{02} / \text{giniy}_{92}) - 1$	-0.18	0.08	-0.28	-0.08
giniy92	Income inequality: Gini coefficient wrt pcy 1992	0.49	0.02	0.47	0.52

Note: All observations are panel (sub-district) means of the household values of the variables mentioned, with the exception of the Inequality measures, which are district means.  
No. of observations: 471.

Source: Author's calculations based on results of income estimates in Chapter 1.

**Table 2.3 Variables and descriptive statistics – Central Urban**

Variable	Definition	Mean	Standard error	Min.	Max.
gry	Annual growth of log per capita inc, 1992-2002: $\text{Ln}(\text{pcy}_{92}) - \text{Ln}(\text{pcy}_{02})/10$	-0.01	0.02	-0.07	0.05
lny92	Log income per capita 1992: $\text{Ln}(\text{pcy}_{92})$	11.89	0.38	11.05	13.24
hhedu92	Household head's education 1992, number of years	6.23	3.74	1.18	17.70
dlntot	Changes in log total individuals per household: $(\text{Intot}_{02} / \text{Intot}_{92}) - 1$	-0.02	0.07	-0.29	0.32
lntot92	Log total persons in household 1992	1.41	0.17	0.88	1.82
hage92	Age of household head 1992	45.49	15.04	21.48	82.37
hetran92	Percentage of household head employed in transport and communications 1992	0.06	0.04	0.00	0.26
heagro92	Percentage of household head employed in agriculture sector 1992	0.03	0.04	0.00	0.21
dhijo	Changes in number of children: $(\text{nhijo}_{02} / \text{nhijo}_{92}) - 1$	-0.01	0.16	-0.67	0.80
giniy92	Income inequality: Gini coefficient wrt pcy 1992	0.46	0.01	0.44	0.51
dginiy	Changes in income inequality: $(\text{giniy}_{02} / \text{giniy}_{92}) - 1$	-0.02	0.02	-0.08	0.03
ginie92	Education inequality: Gini coefficient wrt hhedu <sub>92</sub>	0.29	0.01	0.27	0.31
dginie	Changes in education inequality: $(\text{ginie}_{02} / \text{ginie}_{92}) - 1$	-0.05		-0.09	0.06

Note: All observations are panel (sub-district) means of the household values of the variables mentioned, with the exception of the Inequality measures, which are district means.

No. of observations: 655.

Source: Author's calculations based on results of income estimates in Chapter 1.



**Table 2.4 Variables and descriptive statistics – Remaining Urban**

Variable	Definition	Mean	Standard error	Min.	Max.
gry	Annual growth of log per capita inc, 1992-2002: $\text{Ln}(\text{pcy}_{92}) - \text{Ln}(\text{pcy}_{02})/10$	0.01	0.04	-0.11	0.13
lny92	Log income per capita 1992: $\text{Ln}(\text{pcy}_{92})$	11.72	0.54	10.57	13.31
D13	D13 equals one of department equals 13 (Amambay)	0.09	0.28	0.00	1.00
dlntot	Changes in log total individuals per household: $(\text{Intot}_{02} / \text{Intot}_{92}) - 1$	-0.01	0.09	-0.32	0.48
Intot92	log total individuals per household 1992	1.42	0.22	0.58	1.87
hsedu92	Spouse's education 1992, number of years	4.38	2.38	0.44	13.08
dsedu	Changes in spouse's education in number of years	0.74	0.95	-0.48	7.15
D15	D15 equals one of department equals 15 (Presidente Hayes)	0.02	0.13	0.00	1.00
D6	D6 equals one of department equals 6 (Caazapá)	0.01	0.08	0.00	1.00
hana92	Percentage of households with at least one analphabetic 1992	0.29	0.28	0.00	0.94
dhedu	Changes in household head's education in number of years: $(\text{hedu}_{02} / \text{hedu}_{92}) - 1$	0.38	0.85	-0.70	3.23
D10	D10 equals one of department equals 10 (Alto Paraná)	0.28	0.45	0.00	1.00
dginiy	Changes in income inequality: $(\text{giniy}_{02} / \text{giniy}_{92}) - 1$	0.02	0.04	-0.11	0.15
dhana	Changes in percentage of households with at least one analphabetic: $(\text{dhana}_{02} / \text{dhana}_{92}) - 1$	3.62	5.72	-0.71	60.92
D8	D8 equals one of department equals 8 (Misiones)	0.03	0.16	0.00	1.00
dhfem	Changes in percentage of female household head: $(\text{hfem}_{02} / \text{hfem}_{92}) - 1$	0.48	0.92	-0.67	7.57
hage92	Age of household head 1992	44.93	15.13	20.88	82.38
giniy92	Income inequality: Gini coefficient wrt pcy 1992	0.47	0.02	0.41	0.56
dginie	Changes in education inequality: $(\text{ginie}_{02} / \text{ginie}_{92}) - 1$	-0.02	0.02	-0.07	0.09
nhijo92	Number of children 1992	2.24	0.91	0.31	4.42
hfem92	Percentage of female household head 1992	0.25	0.12	0.03	0.64
ginie92	Education inequality: Gini coefficient wrt hhedu <sub>92</sub>	0.31	0.02	0.27	0.36

Note: All observations are panel (sub-district) means of the household values of the variables mentioned, with the exception of the Inequality measures, which are district means. No. of observations: 762.

Source: Author's calculations based on results of income estimates in Chapter 1.



**Table 2.5 Variables and descriptive statistics – Rural**

Variable	Definition	Mean	Standard error	Min.	Max.
gry	Annual growth of log per capita inc, 1992-2002: $\text{Ln}(\text{pcy}_{92}) - \text{Ln}(\text{pcy}_{02})/10$	0.02	0.04	-0.11	0.16
lny92	Log income per capita 1992: $\text{Ln}(\text{pcy}_{92})$	10.92	0.55	9.82	12.74
dginiy	Changes in income inequality: $(\text{giniy}_{02} / \text{giniy}_{92}) - 1$	-0.01	0.05	-0.17	0.17
D5	D5 equals one of department equals 5 (Caaguazú)	0.14	0.34	0.00	1.00
D13	D13 equals one of department equals 13 (Amambay)	0.02	0.13	0.00	1.00
dlntot	Changes in log total individuals per household: $(\text{Intot}_{02} / \text{Intot}_{92}) - 1$	-0.01	0.09	-0.32	0.46
ginie92	Education inequality: Gini coefficient wrt $\text{hhedu}_{92}$	0.28	0.01	0.25	0.33
hhedu92	Household head's education 1992, number of years	3.50	1.78	0.48	11.31
Intot92	log total individuals per household 1992	1.48	0.26	0.67	2.03
hage92	Age of household head 1992	48.17	15.78	20.43	82.62
dhecom	Changes of percentage of household head employed in commercial sector: $(\text{hecom}_{02} / \text{hecom}_{92}) - 1$	0.61	1.70	-1.00	25.10
hecom92	Percentage of household head employed in commercial sector 1992	0.03	0.04	0.00	0.26
dhecs	Changes of percentage of household head employed in community services sector: $(\text{hecs}_{02} / \text{hecs}_{92}) - 1$	0.36	1.42	-1.00	22.19
giniy92	Income inequality: Gini coefficient wrt $\text{pcy}$ 1992	0.54	0.03	0.41	0.63
hsedu92	Spouse's education 1992, number of years	2.71	1.27	0.29	6.60
hfem92	Percentage of female head 1992	0.19	0.11	0.00	0.58
dhedu	Changes in household head's education in number of years: $(\text{hedu}_{02} / \text{hedu}_{92}) - 1$	0.19	0.24	-0.13	3.61
dhetran	Change of percentage of household head employed in transport and communications sector: $(\text{hetran}_{02} / \text{hetran}_{92}) - 1$	0.13	0.95	-1.00	16.16
dsedu	Changes in spouse's education in number of years	0.94	0.89	-0.04	7.67
dginie	Changes in education inequality: $(\text{ginie}_{02} / \text{ginie}_{92}) - 1$	0.01	0.05	-0.13	0.14

Note: All observations are panel (sub-district) means of the household values of the variables mentioned, with the exception of the Inequality measures, which are district means. No. of observations: 2,388.

Source: Author's calculations based on results of income estimates in Chapter 1.

**Table 2.6 Variables and descriptive statistics – Pro-Poor-Growth-Panels**

Variable	Definition	Mean	Standard error	Min.	Max.
gry	Annual growth of log per capita inc, 1992-2002: $\text{Ln}(\text{pcy}_{92}) - \text{Ln}(\text{pcy}_{02})/10$	0.04	0.03	0.01	0.16
hage92	Age of household head 1992	43.40	11.60	21.39	81.80
hhedu92	Household head's education 1992, number of years	3.73	1.74	0.94	11.76
giniy92	Income inequality: Gini coefficient wrt $\text{pcy}_{92}$	0.54	0.02	0.44	0.63
ginie92	Education inequality: Gini coefficient wrt $\text{hhedu}_{92}$	0.28	0.02	0.25	0.36
lny92	Log income per capita 1992: $\text{Ln}(\text{pcy}_{92})$	10.61	0.35	9.82	12.38
nhijo92	Number of children 1992	3.31	1.09	0.55	5.69
dginiy	Changes in income inequality: $(\text{giniy}_{02} / \text{giniy}_{92}) - 1$	-0.01	0.05	-0.28	0.17
dginie	Changes in education inequality: $(\text{ginie}_{02} / \text{ginie}_{92}) - 1$	0.02	0.05	-0.16	0.14
dlnatot	Changes in log total persons in household: $(\text{lnatot}_{02} / \text{lnatot}_{92}) - 1$	-0.02	0.08	-0.32	0.35
dhijo	Changes in number of children: $(\text{nhijo}_{02} / \text{nhijo}_{92}) - 1$	-0.02	0.19	-0.42	1.94

Note: All observations are panel (sub-district) means of the household values of the variables mentioned, with the exception of the Inequality measures, which are district means. No. of observations: 1,300.

Source: Author's calculations based on results of income estimates in Chapter 1.

## 2.5 Estimation

Before discussing the results obtained in regressions it is necessary to make sure that these results can be taken as true. There could be some important bias in the results, given that the independent variable was estimated and not observed. The properties of estimators obtained from downstream<sup>23</sup> regressions using imputed values for welfare indicators are investigated in Elbers et al., 2005. Their main proposition is that coefficients from regressions involving imputed welfare indicators which have been derived from small area estimation techniques, either in the LHS or in the RHS, do not differ systematically from regressions with true indicators ('real data'). The intuition for this consistency result is that imputed variables can be regarded a special kind of instrumental variables and may

23 It is convenient to refer to our inequality-growth regression as a 'downstream' model so as to distinguish it from the 'upstream' expenditure model which has been used to generate the imputed values.

therefore be safely used in estimation. We briefly explore the issues involved in estimation for the general case with imputed values in both the LHS and the RHS of a regression equation.

We consider a simple version of our downstream regression model (omitting inequality measures):

$$g_i = y_i\beta + \mathbf{x}_i\gamma + u_i \quad (2.10)$$

The dependent  $g$  and the independent  $y$  are obtained from upstream imputation; in what follows, imputed variables have tildes in order to distinguish them from ‘true’ values or observations. Writing imputed values as the difference between ‘true’ values and an error term,  $\tilde{g} \equiv g - \omega$  and  $\tilde{y} \equiv y - \xi$ , we obtain:

$$\tilde{g}_i = \tilde{y}_i\beta + (\xi_i\beta - \omega_i) + \mathbf{x}_i\gamma + u_i \quad (2.11)$$

The  $\beta$  coefficient can be consistently estimated provided that (a) the imputed values  $\tilde{g}$  and  $\tilde{y}$  are consistent estimators of the conditional expectation of the true welfare measures and (b) the error terms  $\xi$  and  $\omega$  are uncorrelated with the regressors  $\tilde{y}$  and  $\mathbf{x}$ .

Elbers et al. (2005) show that when small area welfare estimates are used (a) is satisfied and (b) is likely to be satisfied. To see the latter, first note  $\tilde{y}$  is imputed Per Capita Income (PCI) or a non-linear measure calculated from PCI, e.g. inequality.<sup>24</sup> Both  $\xi$  and  $\omega$  are prediction errors and are thus orthogonal to the predicted values  $\tilde{y}$  and  $\tilde{g}$ , respectively. Moreover, since  $\tilde{y}$  and  $\tilde{g}$  are based on the same prediction model, the prediction errors should be orthogonal with respect to both  $\tilde{y}$  and  $\tilde{g}$ .<sup>25</sup>

The prediction errors should also be uncorrelated with regressors in  $\mathbf{x}$ : since the upstream modeling process makes use of as many available instruments as possible, these regressors will have been considered as instruments in the upstream PCI prediction model, ruling out the presence of any remaining correlation.

However, a correction of the estimated standard errors of the coefficients is necessary because the (upstream) imputation process creates correlation between the welfare estimates. Following Elbers et al. (2005), the prediction error of imputed variables, e.g. expenditure, can be decomposed as:

24 Other variables could in principle be imputed or predicted as well; however, we consider PCI imputations.

25 This holds *a fortiori* when either  $y$  or  $z$  is a non-linear transformation of PCI or its distribution, such as a poverty or inequality measure.

$$\xi \equiv y - \tilde{y} = [y - E(y)] + [E(y) - \tilde{y}] \quad (2.12)$$

where  $E(y)$  is the conditional expectation of expenditure. The first term in the RHS of (2.12) is termed the idiosyncratic error, which is due to unobserved factors that determine expenditure, and the second part is the model error, which reflects uncertainty about the upstream model's parameters. Applying this error decomposition to both  $g$  and  $y$  (2.11) can be written as

$$\begin{aligned} \tilde{g}_i = & [\tilde{y}_i\beta + \mathbf{x}_i\boldsymbol{\gamma}] + [(E(y_i) - \tilde{y}_i)\beta - (E(g_i) - \tilde{g}_i)] \\ & + [(y_i - E(y_i))\beta - (g_i - E(g_i)) + u_i] \end{aligned} \quad (2.13)$$

The RHS of the equation consists of three parts, each in square brackets. First we have a structural part consisting of imputed and non-imputed regressors and their respective coefficients. The second part represents the model error, the third part the sum of upstream idiosyncratic error and downstream error.

We simplify notation by rewriting these three parts as  $\tilde{g}_i = \mathbf{z}_i^*\boldsymbol{\lambda} + \varphi_i + e_i$  where  $\mathbf{z}^* = (\tilde{y}, \mathbf{x})$  represents all regressors, both observed and imputed, and  $\boldsymbol{\lambda} = (\beta, \boldsymbol{\gamma})$ ;  $\varphi$  represents the 'model part' of the error and  $e$  the idiosyncratic part. Assuming that the idiosyncratic part of the error is i.i.d., the variance matrix of the OLS coefficient estimates of (2.13) is:

$$V(\boldsymbol{\lambda}) = \sigma_e^2 (\mathbf{Z}'\mathbf{Z})^{-1} + (\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{Z}' V(\varphi) \mathbf{Z} (\mathbf{Z}'\mathbf{Z})^{-1} \quad (2.14)$$

where the model part variance is:

$$V(\varphi) = \beta^2 V(E(y) - \tilde{y}) + V(E(g) - \tilde{g}) - 2\beta \text{Cov}[E(y) - \tilde{y}, (E(g) - \tilde{g})] \quad (2.15)$$

Equation (2.14) shows that, compared to OLS variance estimates, variance has to be adjusted upwards. As (2.15) shows, this adjustment depends on the variance in the model error. The more imputed variables are used, the more terms will have to be added: with  $n$  imputed variables, the number of terms in the RHS of (2.15) equals  $n$  variance terms plus  $n(n-1)/2$  covariance terms. For example, if one imputed variable is used in the RHS only, the adjustment is limited to the first term. In our regression model (equation (2.4)), two imputed variables are used in the RHS, one in the LHS.

In sum, using imputed values of expenditure or other welfare (inequality) measures will lead to unbiased regression estimates. The coefficients of a model like equation (2.4), may be estimated using OLS under the assumption that the idiosyncratic prediction errors and the error term  $u_i$  are i.i.d.

Two additional econometric problems affect our growth model. First, Caselli et al., 1996 show that estimating a cross-section growth model using a fixed effects estimator will lead to substantial bias when the number of periods is small, especially on the coefficient for initial income ( $y_{92}$ ). The empirical growth literature suggests a number of solutions to this problem, most notably the Arrelano-Bond estimator. Such estimators, however, need at least three periods to estimate the model, using the first period to instrument for the initial conditions of the second period which explain growth between periods two and three. Since we only have two periods, we cannot follow this approach. However, although the bias on the ‘convergence coefficient’ may be significant, Monte Carlo experiments indicate that the bias in the other RHS coefficients tends to be small (Forbes, 2000).

The second problem is endogeneity. Even though our model does not contain ‘flow’ variables but only beginning-of-period ‘stocks’, initial expenditure  $y_{92}$  has been used to construct the growth variable and is thus correlated with the error term. The same could be true for the changed variables on the RHS. Initial inequality may also be an endogenous variable, as the literature suggests that growth affects inequality (e.g. Aghion et al., 1999; Lundberg and Squire, 2003). One would expect this to be more problematic for changes in inequality rather than for initial inequality. Put to scrutiny, a Hausman test rejects exogeneity of expenditure inequality, but cannot reject exogeneity of education. Consequently, we deal with the endogeneity of initial expenditure and expenditure inequality.

Since we do not have lagged values, e.g.  $y_{t-1}$ , to use as instruments, we have to find instruments amongst the (few) available sub-county census means. We have chosen the following instruments. In the Asunción regression, the instrument for income is a variable that measures the ‘education deficit’ (the number of years of schooling missed) of children below the age of 13. The (initial) education deficit for children in this age group is strongly negatively correlated with initial income, but arguably, does not affect growth in the period analysed. The instrument for income inequality is the ‘ethnic fractionalization’, which is the probability that any two citizens randomly chosen from panel population are from different ethnic groups. For the Central Urban regression the instruments are the same as for Asunción. For the Remaining Urban regression the instruments are “education deficit”, as before, for income and for inequality a dummy indicating whether the household head is working in agricultural sector is used. For the rural model, once again the “education deficit” is instrument for income and the number of children is used as an instrument for inequality. The Pro-Poor-Growth Panel regression instruments are the same as for the Asunción regression.

We tested the validity of the instruments by including them in the different models. They do not alter the other coefficient estimates in any significant way. Fi-

nally, we note that the instrumentation also affects the calculation of the model's variance: imputed endogenous variables have to be instrumented first and then instrumented values are used in the calculation of the variance-covariance matrix  $V(\varphi)$ .

## 2.6 Results

The estimated standard errors in all our regressions are adjusted to account for prediction errors following the approach outlined in Section 5. The adjustments – illustrated for the baseline equation are found in Tables 2.7 to 2.11.

**Table 2.7 Variance adjustments – Asunción**

<i>Dependent: growth</i>	<i>coef</i>	<i>t-val(adj)</i>	<i>t(2SLS)</i>	<i>incr(se)</i>
lny92	-0.0508	-16.4051	-6.7010	2.4482
hhedu92	0.0030	7.8283	4.1620	1.8809
Dlntot	-0.0534	-8.5381	-5.5010	1.5521
Dsedu	0.0031	4.0523	3.6090	1.1228
Desp	-0.0314	-5.9729	-2.6030	2.2946
Dhijo	-0.0048	-4.1943	-2.1540	1.9472
Dhedu	-0.0349	-1.9666	-2.4500	0.8027
hage92	0.0004	6.1914	3.0510	2.0293
lntot92	-0.0404	-5.6680	-4.5660	1.2414
nhijo92	0.0059	3.2711	2.8930	1.1307
hsedu92	0.0015	2.8324	1.9170	1.4775
hfem92	0.0220	3.1821	2.2030	1.4444
Dest	0.1508	3.7251	2.3730	1.5698
Dginiy	0.0279	1.1205	2.1420	0.5231
giniy92	-0.8042	-5.2783	-2.4380	2.1650
Observations	471			
R squared adjusted	0.601			
F	48.194			
Standard error	0.011			

Source: Author's calculations based on results in Table 2.2

**Table 2.8 Variance adjustments – Central Urban**

<i>Dependent: growth</i>	<i>Coef</i>	<i>t-val(adj)</i>	<i>t(2SLS)</i>	<i>incr(se)</i>
lny92	-0.0701	-38.8157	-4.6400	8.3655
hhedu92	0.0038	23.6922	3.3430	7.0871
Dlntot	-0.0519	-9.1051	-6.2010	1.4683
lntot92	-0.0289	-14.7102	-5.4780	2.6853
hage92	0.0003	10.6186	1.7520	6.0608
hetran92	-0.0443	-5.1810	-3.0360	1.7065
heagro92	0.0487	5.0584	2.6940	1.8776
Dhijo	-0.2044	-4.9576	-1.5610	3.1759
giniy92	-0.0092	-3.4300	0.0930	-36.882*
Dginiy	-0.0690	-2.8767	-1.2340	2.3312*
ginie92	0.1866	4.2781	2.0620	2.0747*
Dginie	0.1425	7.0851	2.5620	2.7655
lny92	-0.0701	-38.8157	-4.6400	8.3655
Observations	655			
R squared adjusted	0.699			
F	127.408			
Standard error	0.011			

\* Variable lost significance in 2SLS estimation at 10% level.

Source: Author's calculations based on results in Table 2.3

**Table 2.9 Variance adjustments – Remaining Urban**

<i>Dependent: growth</i>	<i>Coef</i>	<i>t-val(adj)</i>	<i>t(2SLS)</i>	<i>incr(se)</i>
lny92	-0.0666	-30.9315	-3.4870	8.8705
D13	0.0252	12.8894	6.3370	2.0340
dlntot	-0.0792	-13.9280	-6.8930	2.0206
lntot92	-0.0484	-6.5188	-4.4400	1.4682
hsedu92	0.0075	13.6337	2.0810	6.5515
dsedu	0.0096	9.2672	3.7070	2.4999
D15	0.0104	2.7217	1.1640	2.3382*
D6	-0.0366	-6.6286	-3.5290	1.8783
hana92	-0.0477	-10.2479	-4.7720	2.1475
dhedu	0.0124	8.8931	5.6710	1.5682
D10	-0.0180	-9.0377	-1.3820	6.5396*
dginiy	-0.1630	-9.0003	0.5020	-17.929*
dhana	-0.0004	-4.4230	-3.3880	1.3055
D8	0.0099	3.5585	3.2520	1.0942
dhfem	-0.0013	-2.0242	-2.2550	0.8977
hage92	0.0004	5.7010	1.1960	4.7667*
giniy92	-0.0222	-0.7406	1.8190	-0.4072
dginie	-0.1410	-5.1656	-2.1640	2.3871
nhijo92	0.0055	3.2441	2.3510	1.3799
hfem92	0.0188	2.3932	-0.1220	-19.617*
Observations	762			
R squared adjusted	0.836			
F	186.461			
Standard error	0.016			

\* Variable lost significance in 2SLS estimation at 10% level.

Source: Author's calculations based on results in Table 2.4

**Table 2.10 Variance adjustments – Rural**

<i>Dependent: growth</i>	<i>Coef</i>	<i>t-val(adj)</i>	<i>t(2SLS)</i>	<i>incr(se)</i>
lny92	-0.0774	-68.2896	-3.0580	22.3314
dginiy	-0.1161	-15.4914	-1.5740	9.8420*
D5	-0.0216	-21.4285	-5.8820	3.6431
D13	-0.0538	-18.2676	-0.4700	38.867*
dlntot	-0.0891	-21.9526	-3.0640	7.1647
ginie92	0.4167	15.1772	2.8610	5.3049
hhedu92	0.0036	8.0220	2.7240	2.9449
lntot92	-0.0364	-17.8537	-2.3880	7.4764
hage92	0.0005	13.0275	2.4280	5.3655
dhocom	0.0015	7.7851	2.7760	2.8044
hecom92	0.1196	10.2076	2.0490	4.9818
dhcscs	0.0011	4.7100	2.4600	1.9146
giniy92	-0.0405	-2.5215	-1.7450	1.4450
hsedu92	0.0024	2.7427	-0.4250	-6.4533*
hfem92	0.0170	3.5152	-0.9800	-3.5870*
dhedu	-0.0090	-3.7366	-1.3310	2.8074*
dhetran	0.0010	2.8341	1.6040	1.7669*
dsedu	0.0011	1.4881	-0.2880	-5.1670*
dginie	-0.0243	-3.2285	-1.8790	1.7182
Observations	2388			
R squared adjusted	0.467			
F	111.125			
Standard error	0.034			

\* Variable lost significance in 2SLS estimation at 10% level.

Source: Author's calculations based on results in Table 2.5

**Table 2.11 Variance adjustments – Pro-Poor-Growth Panels**

<i>Dependent: growth</i>	<i>Coef</i>	<i>t-val(adj)</i>	<i>t(2SLS)</i>	<i>incr(se)</i>
hage92	0.0004	6.6849	1.834	10.295*
hhedu92	0.0038	10.7338	1.598	4.1833
giniy92	-0.0613	-2.7456	1.919	5.5934
ginie92	0.6057	18.1587	-1.704	1.6113
lny92	-0.0625	-29.3277	3.848	4.7190
nhijo92	-0.0052	-8.7441	-2.006	14.620
dginiy	-0.1540	-15.3952	-1.699	5.1466
dginie	0.0140	1.2152	-4.021	3.8287*
dlntot	-0.0747	-8.6839	-0.907	-1.3398*
dhijo	-0.0083	-2.2094	0.912	-9.5218
Observations	1300			
R squared adjusted	0.087			
F	13.336			
Standard error	0.049			

\* Variable lost significance in 2SLS estimation at 10% level.

Source: Authors's calculations based on results in Table 2.6



In all four regions and in the poor household sample, the result is an increase in estimated standard errors for all coefficients. The last column of Tables 2.7 to 2.11 gives the ratio of the adjusted standard error estimates to the standard 2SLS estimates. The increase varies over coefficients between a factor 0.5 and up to 22.3, considering all variables that did not lose significance in the 2SLS estimation. The results in Tables 2.7 to 2.11, illustrate the general decrease in significance when taking into account the fact that estimates or predictions, and not data, are used. In many cases the adjustment even ‘destroys’ a significant result, that is, causes the significance level to increase to over ten percent. This is the typical trade-off when analysing small area welfare estimates: the gain in the number of ‘observations’ obtained by using imputed variables is partly offset by the loss in precision due to (downstream) model prediction errors.

The main findings are presented in Tables 2.12 to 2.16 in a series between seven and ten regressions. They were separated in different regression models, because the estimation of income is based on different models as well.

In the 2SLS regression the complete Asuncion model (regression 1) loses quality. Adjusted R squared decreases from 0.748 (Table 2.12) to 0.601 (Table 2.7), and the models standard error increases from 0.008 (Table 2.12) to 0.011 (Table 2.7). All variables except one have the expected signs. Even if total number of individuals per household decreased during the observation period, for Asuncion we get a negative sign for this change, significant at 1% level in all specifications.

Conditional convergence is pronounced in all specifications: the coefficient on initial income is negative, highly significant and has a value of around  $-0.05$  in all specifications. Apparently, sub-district panels with lower mean per capita income in 1992 have grown faster over the 1990s, *ceteris paribus*. However, note that the coefficient estimate is biased, so we should not attach significance to its exact value.

We have interesting and consistent results for growing primary sector employment of the household head, which ends up harming growth and a growing tertiary household head employment that benefits from growth. In three out of four specifications we find that decreasing household head education harms growth and surprisingly that female-headed households are better off, regarding their growth capacity in all specifications. Household heads education, age, spouses’ education and changes in the number of children have very small effects.

The main variable of interest, inequality, has been entered using income inequality (gini). For Asuncion, education inequality is correlated with income inequality.

ity and was left out. The results show that income inequality (gini) has a significant negative effect on growth in all specifications. The change in income inequality (income inequality decreased in Asuncion) has a significant but negative effect only in model 6. The positive effects of a decrease in education inequality are up to three times stronger than the positive effect of an initial education inequality (considering standardized coefficients). In the 2SLS regression the complete Central Urban model (regression 1) loses quality. Adjusted R squared decreases from 0.889 (Table 2.13) to 0.699 (Table 2.8), and the models standard error increases from 0.007 (Table 2.13) to 0.011 (Table 2.8).

Conditional convergence is pronounced in all specifications: the coefficient on initial income is negative, highly significant and has a value of approximately  $-0.07$  in all specifications.

All variables except one have the expected signs. Even if the total number of individuals per household decreased during the observation period for Central Urban area, we get a negative sign for this change, significant at 1% level in all specifications.

The main variable of interest, inequality, has been entered using income inequality and education inequality; these variables have been entered in linear and quadratic form in alternative specifications. Income inequality has a negative effect on growth, significant in three specifications at the 5% level and once at the 1% level. In contrast, education inequality has a changing effect on growth. In three times out of four significant specifications, the effect is positive. When only education inequality is entered, – without income inequality, (column 2) – there is no significant effect. Including education and income inequality squared, produces mixed results (positive and negative coefficients), so there is no strong evidence for a relation of u-shape or inverted u-shape, but the small decrease of income inequality observed in Central Urban has a negative effect on growth. At the same time, the observed decrease in education inequality has a strong and significant positive effect on growth. The observed effects of changes in household heads employment sector, composition of household or family group or household age and initial education are very small.

**Table 2.12 Regression results – Asuncion**

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent	gry	gry	gry	gry	gry	gry
(Constant)	1.02 (11.92)***	0.99 (11.73)***	0.99 (12.13)***	0.60 (17.62)***	1.00 (12.30)***	1.00 (17.72)***
lny92	-0.05 (-16.41)***	-0.05 (-16.63)***	-0.05 (-16.98)***	-0.05 (-15.87)***	-0.05 (-16.83)***	-0.05 (-16.04)***
hhedu92	0.003 (7.83)***	0.003 (8.15)***	0.003 (8.16)***	0.003 (7.40)***	0.003 (8.10)***	0.003 (7.60)***
dIntot	-0.05 (-8.54)***	-0.05 (-8.67)***	-0.05 (-8.52)***	-0.06 (-8.59)***	-0.05 (-8.85)***	-0.05 (-8.65)***
dsedu	0.003 (4.05)***	0.003 (3.98)***	0.003 (3.98)***	0.003 (3.48)***	0.003 (4.06)***	0.003 (3.47)***
desp	-0.03 (-5.97)***	-0.03 (-6.22)***	-0.03 (-6.01)***	-0.01 (-2.73)***	-0.03 (-6.36)***	-0.03 (-4.23)***
dhijo	0.005 (-4.19)***	0.005 (-4.26)***	0.005 (-4.22)***	-0.01 (-4.44)***	0.005 (-4.28)***	0.005 (-4.46)***
dhedu	-0.03 (-1.97)*		-0.02 (-1.86)*	-0.01 (-0.67)		
hage92	0.0004 (6.19)***	0.0004 (6.31)***	0.0004 (6.41)***	0.0004 (6.10)***	0.0004 (6.25)***	0.0004 (6.16)***
Intot92	-0.04 (-5.67)***	-0.04 (-5.79)***	-0.04 (-5.79)***	-0.04 (-5.45)***	-0.04 (-5.74)***	-0.04 (-5.52)***
nhijo92	0.01 (3.27)***	0.01 (3.36)***	0.01 (3.31)***	0.01 (3.04)***	0.01 (3.36)***	0.01 (3.09)***
hsedu92	0.002 (2.83)***	0.001 (2.68)**	0.002 (2.81)**	0.002 (2.75)**	0.001 (2.61)**	0.001 (2.71)**
hfem92	0.02 (3.18)***	0.02 (3.26)***	0.02 (3.35)***	0.02 (3.04)***	0.02 (3.15)***	0.02 (3.08)***
dest	0.15 (3.73)***	0.11 (3.17)***	0.12 (4.41)***	-0.02 (-0.88)	0.14 (6.04)***	0.14 (-1.33)
dginiy	0.03 (1.12)	-0.01 (-0.93)		-0.02 (-0.94)		-0.78 (-2.63)***
giniy92	-0.80 (-5.28)***	-0.73 (-4.93)***	-0.74 (-5.25)***		-0.78 (-5.54)***	
Observations	471	471	471	471	471	471
R squared	0.756	0.754	0.755	0.741	0.754	0.741
R squared adjusted	0.748	0.746	0.748	0.733	0.747	0.734
F	94.043	99.856	100.614	93.275	107.505	100.536
Stand. error	0.008					

Notes: Absolute value of t statistics in parentheses.

Significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%.

Source: Author's calculations based on results of income estimates in Chapter 1.

In the 2SLS regression the complete Remaining Urban model (regression 1) loses quality. Adjusted R squared decreases from 0.909 (Table 2.14) to 0.836 (Table 2.9), and the model's standard error increases from 0.012 (Table 2.14) to 0.016 (Table 2.9).

All variables except one have the expected signs. Even if the total number of individuals per household decreases during the observation period for the Remaining Urban area, we get a negative sign for this change, significant at 1% level in all specifications.

Conditional convergence is pronounced in all specifications: the coefficient on initial income is negative, highly significant and has a value of approximately  $-0.07$  in all specifications. Income inequality has a significant effect only in three out of eight specifications; two of these three are negative. Education inequality has a significant negative effect in all specifications. The observed increase in income inequality has a negative effect in all specifications and the observed smaller increase in education inequality has a negative and significant effect in all specifications.

Positive effects of household heads' education and age are still small but a little more important than in the Asuncion and Central Urban areas. Again, we have some evidence that female-headed households are better off regarding growth. For five out of 16 possible departments we find dummies with the expected signs regarding their overall economic performance. So sub-regional differences in growth performance exist, but their effect is considerably small.

In the 2SLS regression the complete Rural model (regression 1) loses quality. Adjusted R squared decreases from 0.867 (Table 2.15) to 0.467 (Table 2.10), and the model's standard error increases from 0.015 (Table 2.15) to 0.034 (Table 2.10). All variables except one have the expected signs. Even if the total number of individuals per household decreased during the observation period for rural area, we get a negative sign for this change, significant at 1% level in all specifications. Conditional convergence is pronounced in all specifications: the coefficient on initial income is negative, highly significant and has a value of approximately  $-0.08$  in all specifications.

Income inequality has a changing significant effect (two times positive, two times negative). Education inequality has a significant positive effect in all specifications. The observed increase in income inequality has an important negative effect on growth in all specifications, as well as the increase in education inequality. Household heads' age and education do not have important effects on growth. Female-headed households are better off regarding their growth

capacities, as are households whose head is working in the commercial sector. Nevertheless, the positive effect of an increase in commercial employment, even if highly significant, ends up being very small.

For two out of 17 possible departments we find dummies with the expected signs regarding their overall economic performance. Consequently, sub-regional differences in growth performance exist, but their effect is considerably small.

Before running a separate fifth regression model on a sub-sample of panels for which pro-poor-growth has been determined, we checked on the veracity of this data (see Annex). About 97% of the sub-sample for pro-poor-growth is from rural areas. There are no spatial patterns, the Pro-Poor-Growth (PPG) panels are distributed all over the country, so PPG seems to be not the result of specific geographic area or any special districts, with better economic performance. It is a consequence of activities carried out by certain groups of people, permitting them to overcome part of their poverty. This phenomenon is observed in almost any part of the country (in 15 out of 18 departments and in 154 of the 224 districts).

If PPG is a consequence of group dynamics and not of spatial structures we should know more about these group characteristics. In all PPG panels the mother tongue is Guaraní (indicator for low ethnical fragmentation), and 98.4% of the household heads have less than 5 years of education. The maximum geographic concentration is of 29 panel groups in the same district (2.4% of the sample). The 1300 identified PPG panel groups represent approximately 5% of all households and some 10% of poor households. The age distribution of PPG panel household heads follows the age distribution of all household heads.

In the 2SLS regression the complete PPG model (regression 1) loses almost all its quality. Adjusted R squared decreases from 0.601 (Table 2.16) to 0.087 (Table 2.11), and the model's standard error increases from 0.017 (Table 2.16) to 0.049 (Table 2.11).

All variables except one have the expected signs. Even if the total number of individuals per household decreased during the observation period, for PPG sub-sample we get a negative sign for this change, significant at 1% level in all specifications.

Conditional convergence is pronounced in all specifications: the coefficient on initial income is negative, highly significant and has a value of approximately – 0.06 in all specifications.

Income inequality has an important negative effect. By construction, this is to be expected at least if a household is poor. Education inequality has a positive effect in four out of nine specifications. The small decrease observed in income inequality has a negative and highly significant effect in all specifications. No significant effect is caused by the increase in education inequality. Income inequality squared produces significant positive effects in five out of six specifications, so there seems to be a u-shape relation. Only in one specification, letting out initial education inequality, education inequality squared produces a significant positive effect.

Household heads age and education, the number of children and the change in their number (small decrease observed) do have significant but very small effects on PPG in our case.

**Table 2.13 Regression results – Central Urban**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent	gry	gry	gry	gry	gry	gry	gry	gry	gry
(Constant)	1.03 (3.20)***	0.80 (2.54)**	0.88 (2.76)**	1.45 (5.43)***	1.10 (3.29)***	1.44 (5.38)***	0.73 (2.30)**	0.80 (45.17)***	0.83 (47.75)***
lny92	-0.07 (-39.15)***	-0.07 (-38.73)***	-0.07 (-46.45)***	-0.07 (-38.95)***	-0.07 (-37.98)***	-0.07 (-38.96)***	-0.07 (-46.44)***	-0.07 (-46.51)***	-0.07 (-55.22)***
hhedu92	0.004 (23.96)***	0.004 (23.56)***	0.004 (25.18)***	0.004 (23.83)***	0.004 (23.04)***	0.004 (23.84)***	0.004 (25.07)***	0.004 (25.18)***	0.004 (27.66)***
dlntot	-0.05 (-9.25)***	-0.05 (-9.04)***	-0.05 (-9.10)***	-0.05 (-9.27)***	-0.05 (-8.61)***	-0.05 (-9.27)***	-0.05 (-8.96)***	-0.05 (-9.01)***	-0.05 (-8.57)***
lntot92	-0.03 (-14.89)***	-0.03 (-14.63)***	-0.03 (-14.37)***	-0.03 (-14.86)***	-0.03 (-14.28)***	-0.03 (-14.87)***	-0.03 (-14.29)***	-0.03 (-14.34)***	-0.03 (-14.58)***
hage92	0.001 (10.84)***	0.001 (10.58)***	0.001 (10.19)***	0.001 (10.69)***	0.001 (10.84)***	0.001 (10.70)***	0.001 (10.20)***	0.001 (10.20)***	0.001 (12.27)***
hetran92	-0.04 (-5.29)***	-0.04 (-5.17)***	-0.04 (-5.06)***	-0.04 (-5.22)***	-0.04 (-4.80)***	-0.04 (-5.22)***	-0.04 (-5.01)***	-0.04 (-5.01)***	-0.04 (-4.79)***
heagro92	0.05 (5.20)***	0.05 (5.01)***	0.04 (4.55)***	0.05 (5.20)***	0.05 (4.53)***	0.05 (5.21)***	0.04 (4.53)***	0.04 (4.55)***	0.04 (4.50)***
dhijo	-0.01 (-3.35)***	-0.01 (-3.45)***	-0.01 (-3.78)***	-0.01 (-3.34)***	-0.01 (-3.60)***	-0.01 (-3.35)***	-0.01 (-3.79)***	-0.01 (-3.79)***	-0.01 (-3.69)***
giniy92	-5.06 (-3.27)***		-3.84 (-2.53)**	-2.53 (-2.25)**	2.00 (1.57)	-2.59 (-2.30)**			
dginiy	-0.08 (-3.48)***	-0.07 (-2.79)**		-0.08 (-3.17)***	-0.01 (-0.41)	-0.08 (-3.18)***			
ginie92	7.07 (2.37)**	0.36 (0.16)	5.62 (1.89)*		-5.07 (-1.94)*	0.17 (3.91)***	0.50 (0.23)		
dginie	0.19 (7.48)***	0.14 (6.87)***	0.15 (6.57)***	0.16 (7.35)***		0.16 (7.39)***	0.12 (6.25)***	0.12 (6.39)***	
ginie92 sqrd	-12.05 (-2.31)**	-0.30 (-0.08)	-9.52 (-1.83)*	0.30 (3.87)***	9.14 (2.01)**		-0.55 (-0.14)	0.31 (4.12)***	0.28 (3.61)***



giniy92 sqrd	5.22 (3.14)***	-0.22 (-4.83)***	4.02 (2.45)**	2.50 (2.07)**	-2.25 (-1.62)	2.56 (2.12)**	-0.12 (-4.11)***	-0.12 (-4.18)***	-0.06 (-2.02)**
Observations	655	655	655	655	655	655	655	655	655
R sqrd.	0.892	0.890	0.890	0.891	0.882	0.891	0.889	0.889	0.882
R sqrd. adj.	0.889	0.888	0.887	0.889	0.880	0.889	0.887	0.887	0.880
F	376.596	398.712	397.732	402.232	369.609	402.420	426.771	466.252	479.167
Stand. error	0.007								

Notes: Absolute value of t statistics in parentheses.

Significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%.

Source: Author's calculations based on results of income estimates in Chapter 1.



**Table 2.14 Regression results – Remaining Urban**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent (Constant)	gry 0.91 (27.39)***	gry 0.88 (25.34)***	gry 0.90 (29.73)***	gry 0.82 (28.41)***	gry 0.93 (29.09)***	gry 0.81 (31.89)***	gry 0.80 (24.80)***
lny92	-0.07 (-31.05)***	-0.07 (-30.96)***	-0.07 (-31.48)***	-0.06 (-31.00)***	-0.07 (-32.92)***	-0.06 (-31.51)***	-0.07 (-29.82)***
D13	0.03 (12.87)***	0.03 (17.70)***	0.03 (13.25)***	0.03 (14.03)***	0.03 (17.55)***	0.03 (14.44)***	0.03 (16.55)***
dIntot	-0.08 (-13.99)***	-0.08 (-14.27)***	-0.08 (-14.30)***	-0.08 (-14.26)***	-0.08 (-13.84)***	-0.08 (-14.57)***	-0.09 (-14.27)***
Intot92	-0.05 (-6.50)***	-0.06 (-7.71)***	-0.05 (-6.52)***	-0.04 (-5.88)***	-0.06 (-7.83)***	-0.04 (-5.90)***	-0.06 (-7.55)***
hsedu92	0.01 (13.75)***	0.01 (13.46)***	0.01 (13.82)***	0.01 (12.75)***	0.01 (14.30)***	0.01 (12.83)***	0.01 (12.58)***
dsedu	0.01 (9.14)***	0.01 (8.46)***	0.01 (9.12)***	0.01 (9.40)***	0.01 (8.75)***	0.01 (9.39)***	0.01 (8.65)***
D15	0.01 (2.73)**	0.02 (6.28)***	0.01 (2.70)**	0.01 (3.23)***	0.03 (7.38)***	0.01 (3.20)***	0.03 (6.95)***
D6	-0.04 (-6.65)***	-0.03 (-5.73)***	-0.04 (-6.64)***	-0.04 (-7.13)***	-0.03 (-5.63)***	-0.04 (-7.12)***	-0.04 (-6.31)***
hana92	-0.05 (-10.34)***	-0.04 (-9.22)***	-0.05 (-10.33)***	-0.05 (-10.67)***	-0.05 (-9.43)***	-0.05 (-10.67)***	-0.05 (-10.04)***
dhedu	0.01 (9.03)***	0.01 (8.52)***	0.01 (9.04)***	0.01 (9.37)***	0.01 (8.39)***	0.01 (9.39)***	0.01 (9.00)***
D10	-0.02 (-8.96)***	-0.01 (-4.97)***	-0.02 (-9.11)***	-0.02 (-9.22)***	-0.01 (-3.86)***	-0.02 (-9.36)***	-0.01 (-3.10)***
dginiy	-0.16 (-8.93)***		-0.16 (-9.61)***	-0.15 (-8.21)***		-0.15 (-8.84)***	
dhana	-0.0001 (-4.49)***	-0.0001 (-4.18)***	-0.0001 (-4.50)***	-0.0001 (-4.49)***	-0.0001 (-4.10)***	-0.0001 (-4.50)***	-0.0001 (-3.70)***

D8	0.01 (3.54)***	0.01 (4.67)***	0.01 (3.57)***	0.01 (3.22)***	0.01 (4.73)***	0.01 (3.25)***	0.01 (3.87)***
dhfem	0.00 (-1.92)**	0.00 (-2.37)**	0.00 (-1.93)**	0.00 (-1.62)*	0.00 (-2.37)**	0.00 (-1.63)*	0.00 (-2.53)**
hage92	0.001 (5.73)***	0.001 (6.30)***	0.001 (5.69)***	0.001 (4.85)***	0.001 (6.97)***	0.001 (4.81)***	0.001 (6.21)***
giniy92	-0.02 (-0.66)	0.09 (3.41)***		-0.02 (-0.58)	-0.11 (-3.72)***		0.04 (1.54)
dginie	-0.14 (-5.19)***	-0.11 (-3.90)***	-0.14 (-5.19)***				-0.03 (-1.16)
nhijo92	0.01 (3.15)***	0.01 (3.79)***	0.01 (3.16)***	0.01 (2.94)***	0.01 (3.89)***	0.01 (2.95)***	0.01 (3.89)***
hfem92	0.02 (2.62)**	0.01 (0.55)	0.02 (2.55)**	0.02 (2.29)**	0.00 (0.75)	0.02 (2.23)*	-0.01 (-0.45)
ginie92	-0.33 (-6.74)***	-0.29 (-5.69)***	-0.33 (-7.17)***	-0.20 (-4.67)***	-0.23 (-4.78)***	-0.21 (-5.12)***	
Observations	762	762	762	762	762	762	762
R sqrd.	0.911	0.902	0.911	0.908	0.900	0.908	0.897
R sqrd. adj.	0.909	0.899	0.909	0.906	0.898	0.906	0.895
F	361.718	339.691	380.072	365.645	351.923	385.213	341.438
Stand. error	0.012						

Notes: Absolute value of t statistics in parentheses.

Significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%;

Source: Author's calculations based on results of income estimates in Chapter 1.

Table 2.15 Regression results – Rural

Dependent (Constante)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	gry 0.78 (40.47)***	gry 0.79 (39.33)***	gry 0.84 (42.82)***	gry 0.75 (50.12)***	gry 0.76 (40.78)***	gry 0.86 (41.74)***	gry 0.81 (54.41)***	gry 0.77 (39.02)***
lny92	-0.08 (-68.29)***	-0.08 (-71.20)***	-0.08 (-64.54)***	-0.08 (-73.87)***	-0.08 (-68.62)***	-0.08 (-67.58)***	-0.08 (-84.28)***	-0.08 (-70.84)***
dginiy	-0.12 (-15.49)***		-0.12 (-14.93)***	-0.11 (-15.36)***	-0.12 (-16.03)***			
D5	-0.02 (-21.43)***	-0.03 (-24.62)***	-0.03 (-25.67)***	-0.02 (-21.69)***	-0.02 (-21.25)***	-0.03 (-29.03)***	-0.03 (-24.58)***	-0.03 (-24.40)***
D13	-0.05 (-18.27)***	-0.05 (-17.48)***	-0.04 (-14.65)***	-0.05 (-18.37)***	-0.05 (-18.24)***	-0.04 (-14.05)***	-0.05 (-17.42)***	-0.05 (-17.41)***
dlntot	-0.09 (-21.95)***	-0.10 (-22.78)***	-0.09 (-21.58)***	-0.09 (-22.47)***	-0.09 (-21.97)***	-0.10 (-22.44)***	-0.10 (-22.73)***	-0.10 (-22.84)***
ginie92	0.42 (15.18)***	0.42 (14.60)***		0.39 (15.15)***	0.44 (16.71)***		0.44 (16.05)***	0.46 (16.63)***
hhedu92	0.004 (8.02)***	0.004 (7.71)***	0.004 (8.33)***	0.004 (7.79)***	0.004 (7.75)***	0.004 (8.03)***	0.004 (7.94)***	0.004 (7.24)***
lntot92	-0.04 (-17.85)***	-0.04 (-21.23)***	-0.03 (-15.75)***	-0.04 (-17.69)***	-0.04 (-17.71)***	-0.04 (-19.10)***	-0.05 (-22.60)***	-0.04 (-21.10)***
hage92	0.001 (13.03)***	0.001 (15.65)***	0.001 (10.78)***	0.001 (12.91)***	0.001 (13.14)***	0.001 (13.41)***	0.001 (15.90)***	0.001 (15.96)***
dhecom	0.002 (7.79)***	0.002 (8.46)***	0.002 (8.52)***	0.002 (7.88)***	0.002 (7.59)***	0.002 (9.16)***	0.002 (8.42)***	0.002 (8.17)***
hecom92	0.12 (10.21)***	0.15 (12.02)***	0.13 (10.50)***	0.12 (10.51)***	0.12 (10.29)***	0.16 (12.27)***	0.14 (11.93)***	0.15 (12.23)***
dhecs	0.001 (4.71)***	0.002 (6.29)***	0.001 (4.46)***	0.001 (4.76)***	0.001 (4.87)***	0.002 (6.01)***	0.002 (6.30)***	0.002 (6.62)***
giniy92	-0.04 (-2.52)**	0.03 (1.59)	0.04 (2.34)**		-0.04 (-2.63)**	0.11 (6.57)***		0.03 (1.59)

hsedu92	0.003 (2.74)**	0.003 (3.68)***	0.002 (1.87)*	0.003 (3.03)***	0.003 (3.20)***	0.003 (2.81)***	0.003 (3.55)***	0.004 (4.46)***
hfem92	0.02 (3.52)***	0.03 (4.99)***	0.03 (5.39)***	0.02 (4.22)***	0.02 (3.68)***	0.04 (6.78)***	0.02 (4.75)***	0.03 (5.32)***
dhedu	-0.01 (-3.73)***	-0.01 (-4.48)***	-0.01 (-2.11)**	-0.01 (-3.87)***	-0.01 (-3.92)***	-0.01 (-2.90)***	-0.01 (-4.42)***	-0.01 (-4.80)***
dhetran	0.001 (2.83)***	0.001 (2.70)**	0.001 (2.68)**	0.001 (2.85)***	0.001 (2.82)***	0.001 (2.56)**	0.001 (2.68)**	0.001 (2.67)**
dsedu	0.001 (1.48)	0.001 (1.15)	0.001 (1.34)	0.001 (1.57)	0.001 (1.74)*	0.001 (1.03)	0.001 (1.09)	0.001 (1.53)
dginie	-0.02 (-3.22)***	-0.04 (-5.09)***	-0.06 (-7.41)***	-0.02 (-3.31)***		-0.07 (-9.20)***	-0.04 (-5.09)***	
Observations	2388	2388	2388	2388	2388	2388	2388	2388
R sqrd	0.868	0.855	0.855	0.868	0.867	0.842	0.855	0.853
R sqrd adj	0.867	0.854	0.854	0.867	0.866	0.840	0.853	0.852
F	819.956	774.088	777.446	863.204	861.503	740.757	818.945	809.588
Stand. error	0.015							

Notes: Absolute value of t statistics in parentheses.

Significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%.

Source: Author's calculations based on results of income estimates in Chapter 1.

**Table 2.16 Regression results – Pro-Poor-Growth Panels**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(Constant)	0.86 (5.61)***	0.48 (4.88)***	1.01 (7.65)***	0.84 (5.04)***	0.90 (5.93)***	0.49 (5.02)***	0.94 (7.07)***	0.55 (18.88)***	0.82 (4.99)***
hage92	0.00 (7.05)***	0.00 (6.60)***	0.00 (7.12)***	0.00 (7.73)***	0.00 (7.17)***	0.00 (6.63)***	0.00 (7.09)***	0.00 (6.68)***	0.00 (7.70)***
hhedu92	0.00 (10.81)***	0.00 (10.71)***	0.00 (10.79)***	0.00 (10.55)***	0.00 (10.81)***	0.00 (10.72)***	0.00 (10.81)***	0.00 (10.73)***	0.00 (10.55)***
giniy92	-1.49 (-3.23)***		-1.37 (-2.99)***	-1.22 (-2.42)**	-1.47 (-3.18)***	-0.06 (-2.75)**	-1.42 (-3.11)***	-0.06 (-2.74)**	-1.23 (-2.45)**
ginie92	1.26 (1.96)*	0.99 (1.54)		0.91 (1.29)	1.04 (1.66)	1.00 (1.56)	0.59 (17.53)***	0.61 (18.15)***	1.03 (1.51)
lny92	-0.06 (-29.15)***	-0.06 (-29.17)***	-0.06 (-29.25)***	-0.07 (-27.36)***	-0.06 (-30.26)***	-0.06 (-29.15)***	-0.06 (-29.25)***	-0.06 (-29.32)***	-0.06 (-27.91)***
nhijo92	-0.01 (-8.94)***	-0.01 (-8.60)***	-0.01 (-9.12)***	-0.01 (-10.60)***	-0.01 (-9.11)***	-0.01 (-8.63)***	-0.01 (-9.06)***	-0.01 (-8.74)***	-0.01 (-10.58)***
dginiy	-0.16 (-15.54)***	-0.15 (-15.37)***	-0.15 (-15.46)***		-0.15 (-15.49)***	-0.15 (-15.40)***	-0.15 (-15.51)***	-0.15 (-15.39)***	
dginie	0.02 (1.43)	0.02 (1.32)	0.01 (0.98)	-0.01 (-0.72)		0.02 (1.32)	0.01 (1.22)	0.01 (1.21)	
dIntot	-0.07 (-8.64)***	-0.08 (-8.73)***	-0.07 (-8.46)***	-0.08 (-8.98)***	-0.07 (-8.55)***	-0.08 (-8.69)***	-0.07 (-8.58)***	-0.07 (-8.68)***	-0.09 (-9.09)***
dhijo	-0.01 (-2.16)**	-0.01 (-1.95)*	-0.01 (-2.70)**	-0.01 (-1.82)*	-0.01 (-2.24)**	-0.01 (-1.97)*	-0.01 (-2.49)**	-0.01 (-2.20)**	-0.01 (-1.78)*
giniy92 sqrd	1.33 (3.10)***	-0.05 (-2.59)***	1.22 (2.86)***	1.12 (2.40)***	1.31 (3.05)***		1.26 (2.98)***		1.13 (2.42)***
ginie92 sqrd	-1.16 (-1.04)	-0.66 (-0.59)	1.02 (17.43)***	-0.54 (-0.44)	-0.80 (-0.74)	-0.68 (-0.61)			-0.74 (-0.62)
Observations	1300	1300	1300	1300	1300	1300	1300	1300	1300
R sqrd adj	0.603	0.600	0.602	0.529	0.603	0.600	0.603	0.601	0.529
F	165.456	178.236	179.747	133.575	180.164	178.424	180.384	196.322	146.935
Stand. error	0.017								

Absolute value of t statistics in parentheses. ;Significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%;Source: Author's calculations based on results in Chapter I

## 2.7 Discussion

The two most important findings of this study are that (1) income inequality does not necessarily have a negative effect on growth, but the observed decrease in income inequality in all models carried out harms growth; and (2) education (human capital) inequality has mixed effects on growth, depending on the initial level of education inequality. An increase in education inequality harms growth and a decrease in education inequality benefits growth. Furthermore, (3) in the Paraguayan case, the effects of changes in inequality are larger than the effects of inequality itself and (4) inequality effects and the effects of their change are bigger than family-group or employment sector effects. (5) There is almost no PPG in urban areas and in rural areas it is related to groups of individuals but not to geographical location. (6) A lower population growth (decrease in the total number of individuals per household) is negative for growth and (7) female-headed households are better off, regarding income growth. The first of these findings is mainly in line with cross-country evidence in Birdsall and Londono (1997), while the second result contrasts with findings in that paper, but supports findings by Schipper and Hoogeveen (2004).

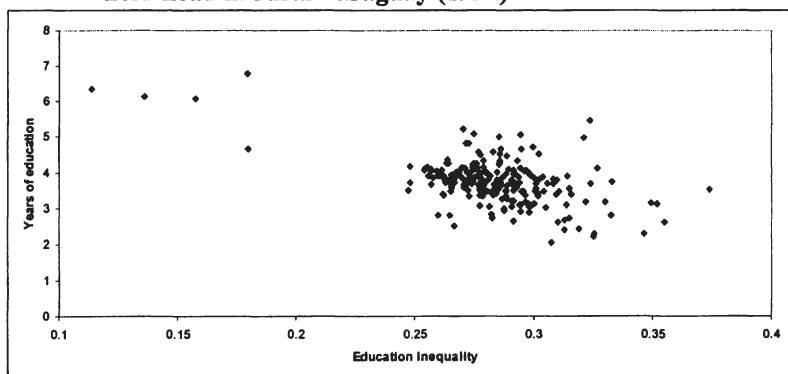
This second point may appear somewhat counter intuitive at first sight: growth is enhanced when human capital (or access to it) of the household head is more unequally distributed. The key to understanding what is going on is the fact that we control for district mean level of education: this means that our conclusion is that *at a given mean level* of human capital, a more unequal distribution of this capital is good for growth. Nevertheless, there is some weak evidence in Paraguayan data that this is not true at any level, because for higher levels of education inequality in Paraguay its effect on growth is negative (Remaining Urban region) or tends to be negative (Central Urban region) but has positive effects on growth in Rural area (and by this on PPG). Elbers and Gunning (2004) show that our result is to be expected in a Ramsey growth model: under the condition that the production function is convex in human capital, a mean-preserving spread in human capital results in higher output growth. For instance, suppose we were to redistribute one year of education from someone with low educational attainment to someone who is reasonably well educated. This would make the distribution of human capital more unequal while keeping the mean constant. However, if the increase in output by the well-educated person exceeds the decline for the less well-educated person, then the increased spread in education has a positive effect on growth – as long as the mean level of education is kept constant.

Mean preserving spreads in human capital are not possible within a given population; they only exist in theoretical experiments or in the long run, that is,



over generations. In reality, the mean level of education and inequality change simultaneously. In rural Paraguay, where a positive effect of education inequality on growth was found, education inequality – as measured by the Gini coefficient – has a negative correlation with the average level of education (see Figure 2.1). In theory, the implication of such a correlation is that, while raising the general level of education through policies like universal primary education will be good for growth; its positive effects will be partly offset by an expected associated decline in the education inequality. Nevertheless, for rural Paraguay the empirical evidence is that even if mean household heads education increased, education inequality also increased. This increase in education inequality harmed growth, even if the initial level of education inequality seems to have been an advantage. This evidence combined with results from different urban areas in Paraguay (where an education inequality higher than in rural area was harmful for growth) confirms the hypothesis that for a given level of inequality in relation to a given number of years of schooling, a higher level of education inequality can be a benefit, however, this is not that any higher level of education inequality has this same effect.

**Figure 2.1 District means of education and education inequality of household head in rural Paraguay (1992)**



Source: Author's calculations based on results in Chapter 1.

The larger effects of changes in inequality compared with the effects of inequality itself on growth are consistent with Paraguayan macro-economic and business cycles history, as well as with its education politics during this business cycle. A decreasing growth and beginning recession reduces growth. For all three different urban areas, annual growth rates of per capita income are negative, while the rural rate is positive but small. At the same time, an increase in education was driven by an education reform that started in Paraguay in 1994, producing a decrease in education inequality only in the Asuncion and Central Urban

areas. Within a context of economic recession, finally these effects happen to be stronger than most of the observed changes inside families or regarding employment opportunities.

The PPG related to groups of individuals and not to geographic location indicates that the sort of PPG we observe in Paraguay is related more to opportunities and less to structural changes or other effects. As a matter of fact, there were few structural changes in the rural economy in Paraguay over the period observed. Most of the rural PPG opportunities could be related to new cash crop farming and their export, such as sesame and some varieties of organic cotton. Unfortunately our census database cannot link the empirical evidence with the production sector data, since we are working with a pseudo panel and not pure household micro-data.

Finally, a decrease in the total number of individuals per household (consistent with fertility decrease during this period) is negatively related to growth, even if a decrease in the number of children per household is not. In a way, we could consider that less people per household in general equals a lower working force and a lower capacity of generating income. On the other hand, if annual income growth rates are negative, less people in a household should impulse an increase in per capita income. This is possibly a spurious relation, because per capita income and number of people per household decreased simultaneously.

Rethinking all these results from an income-mobility point of view; remember that the initial level of income can be understood as a proxy for family background and initial education level as a proxy for institutional opportunities to develop talent. Also remember that initial income inequality was considerably high and slightly decreasing during the observation period, while education inequality was lower and moved in different directions for the different regions. If initially higher levels of income facilitate upward income mobility and a higher income inequality benefits that process, we should expect that this effect benefits a more middle class kind of household. If inequality supports growth, it is easier to grow, but at the same time more difficult to reduce poverty, which in the end would be a strong income growth for low-income groups. Remember that in our results, initial levels of education have almost no effect on growth and education (human capital) inequality has mixed effects on growth, depending on the initial level of education inequality, with an increase in education inequality that harms growth and a decrease of education inequality that benefits growth. If so, the best combination for upward income mobility would be a high level of initial income in an area with high-income inequality and an institutional capacity to widen education opportunities in a way that education inequality decreases. This combination can be found in the PPG sub panel (97% rural area) and for Central Urban area. Nevertheless, in both areas, poverty increased during the observa-



tion period. Consequently, in Central Urban area, it might have been a middle class phenomenon and for PPG panels and rural area, even if there were positive growth and income mobility effects, they may not have been strong enough to get households out of poverty and there may have been some downward mobility as well.

## 2.8 Conclusions

We estimated the effect of income and education inequality on growth, using imputed data on income inequality and growth for small administrative units in Paraguay (districts), along with census data for education inequality; all this based on a pseudo panel data set. Carrying out this kind of analysis for a specific country has important benefits. First, it avoids data comparability problems that typically affect cross-country growth regressions. Moreover, by identifying the effects of inequality on growth for a given country, country specificity is taken into account. This enhances the relevance of our results for local policy makers.

In the empirical section we adjusted the standard errors of variable coefficients for the fact that some regressors are imputed; in our case initial income levels and income inequality, and therefore associated with a standard error. The adjustments are considerable; they typically increase standard errors from a factor 0.5 up to 22, using five different models for different areas or groups of households. Our models are not alone in using imputed variables. Most growth regressions do so by relying on GDP or survey based inequality estimates, for instance. This puts into question the significance of some of the inequality and growth results reported elsewhere.

Our results show for rural Paraguay that higher levels of education inequality enhance growth. Controlling for the level of educational attainment, larger variation in education is here good for growth. The latter finding is plausible if the production function is convex in ability, something that can be illustrated with a Ramsey type household growth model. Nevertheless, we find opposed results for urban areas, where education inequality is higher. Our results also show that higher income inequality does not have a uniform effect on growth (it tends to be more harmful in larger urban areas) and that effects of changes in inequality on growth are larger than the effect of inequality itself, this is for both, education and income.

What does this mean for policy in Paraguay? If policymakers are mostly interested in growth, they should be more concerned on income inequality in urban areas and on education inequalities in rural Paraguay. Income inequality is an important issue for income growth in urban areas (and more important in the

Asuncion and Central Urban areas), in a consistent way with the rapidly increasing urban poverty. Fighting urban poverty must consider income inequality. At the same time, the impact of income inequality in rural areas is much less of a problem. Also, education inequality is a greater problem in urban areas, but politics seem to be on track with a certain success of targeting urban education services, since urban education inequality tends to go down, which benefits income growth. For rural areas, the problem is more sophisticated. Even if initial education inequality benefited rural income growth, a badly targeted or non-universal policy implementation of education reform in rural area, increased education inequality, which in theory harms growth. If, for intrinsic reasons or otherwise, policy makers are interested in reducing education inequality, our results suggest that this would damage growth, but only if the policy was pursued by keeping the mean level of education constant. In practice a policy aimed at reducing inequality in education will almost always be mean increasing.

Finally, even if the poverty map exercise which preceded this paper suggested that there are important spatial effects on poverty levels, we did not find spatial effects for a PPG evidence, which seems to be more of a result of individuals and group dynamics and access to (labour and employment) opportunities. For politics this should mean that there is no need for a special growth strategy for special areas in the country as long as there will be new opportunities for almost all of the working force and not only opportunities for a few (which would increase income inequality).



## **Chapter 3**

### **Characterization of inequality changes through microeconometric decomposition**

#### **Paraguay 1992 – 2005**

##### **3.1 Introduction**

The main economic variables have oscillated widely during the 1992 – 2005 period in Paraguay, in association with some macroeconomic and structural transformations, but also following general growth trends and business cycles in the South American region. This can be separated into three sub-periods; 1992 to 1998, 1999 to 2002 and 2003 to 2005.

During the early eighties, the Paraguayan economy benefited from high public investment rates resulting from the construction of the Itaipu and Yacyreta hydro-electric power plants. The country made its own way of stability and growth during a period of hyperinflations and external debt crisis in many South American countries. Nevertheless, its economy fell into a growth crisis (still avoiding debt crisis and hyperinflation) during the second half of the eighties, once the construction period of the hydroelectric power plants came to an end. During the first half of the nineties, Paraguayan economy recovered from recession, now driven by agricultural production and a re-export business boom, based on special arrangements for duty rates for electric and electronic equipment imported to the MERCOSUR (Mercado Comun del Sur - regional free trade agreement established in 1991 by Argentina, Brazil, Paraguay and Uruguay) region via Paraguay. The agricultural success was based on a recovery of international cotton prices, combined with a successful cotton extension program for small farmers within the country and a quick and widespread expansion of mechanized soybean farming. The commercial success with electric and electronic components was based on re-export. Paraguayan import duty rates from outside MERCOSUR were so low, that Brazilian and Argentine enterprises would prefer to buy these products re-exported from Paraguay, rather than importing themselves from outside MERCOSUR, which would have meant higher duty rates. Before this background, Paraguayan GDP per capita grew until 1995 and then remained relatively stable until 1998. The per capita income<sup>26</sup> Gini coefficient fell from 55.8 to 54.0. Mean per capita growth was 0.63% and poverty dropped from 38.2% to 32.1%.

---

26 Including all kinds of income, labour income, non-labour income and imputed values for own housing.

The period between 1999 and 2002 saw great political instability. Weak and un-concluded structural reform processes in the economy, which had begun in the nineties, were terminated. Small-scale cotton farming entered a deep crisis due to falling international prices and considerable parasite problems, combined with adverse climatic conditions (El Niño phenomenon), affecting agriculture in general. External shocks such as the Brazilian devaluation during global finance crisis and the Argentine default strongly hit the country. Per capita growth was -2.6% and poverty leaped to more than 46%. Gini coefficient for income inequality ascended to 56.1. As from 2003, political changes brought the country back to a more stable course. A tax reform and institutional improvements in Government provided more and new revenues to the treasury. Public expenditure, including social expenditure went up. Economy was benefited by a regional recovery. In the production sector, this period is marked by an important growth of livestock and meat exports. Per capita GDP grew 2.0% on average, while poverty went down to 38.2% and Gini coefficient to 52.8.

However, the reasons behind these changes in inequality are more varied and complex than just a macroeconomic history could tell. The main purpose of this paper is to assess the relevance of some forces that are believed to have affected income inequality in Paraguay between 1992 and 2005. More specifically, the microeconometric decomposition methodology proposed by Bourguignon, Ferreira and Lustig (1998) has been used to measure the relevance of various factors, which appear to have driven changes in inequality. In particular, this methodology has been used to identify to what extent changes in the returns to education and experience, in the endowments of unobservable factors (such as individual's innate ability) and their returns, in the wage gap between men and women, in labour market participation and hours of work, and in the educational structure of the population contribute to explain the observed changes in income distribution.

The results of this paper suggest that the smaller change in inequality between 1992 and 1997/98 is mainly as a result of employment (including hours of work) and education effects, characterized by a primary schooling expansion. The larger inequality reduction effect after 1997 is due to returns to education, hours of work (since unemployment increased) and unobservable factors.

The rest of the paper is organized as follows. Section 2 presents the decomposition methodology implemented to assess the relevance of those factors. Section 3 shows the basic facts and discusses some factors that may have affected inequality during the last two decades, while section 4 explains the estimation strategy. The main results of the analysis are presented in section 5. The paper concludes with some brief final comments in section 6.

### 3.2 Methodology

Many different forces exist behind the long-run changes in income distributions or, more generally, distributions of economic welfare, within a population. Some of these forces have to do with changes in the distribution of factor endowments and socio-demographic characteristics, while others have to do with the returns these endowments produce and others with changes in populations' behaviour such as labour supply, consumption patterns or the decision on whether or not to have children. These forces are not independent from each other. This is what makes it difficult to precisely identify fundamental causes and mechanisms behind the dynamics of income distribution. Decomposition techniques are used to identify causes of distributional changes. Traditional techniques explain differences in scalar summary measures of distributions rather than in full distributions. The best known of these techniques is the Oaxaca-Blinder decomposition of differences in mean incomes across population groups with different characteristics (Blinder 1973; Oaxaca 1973) and the variance-like decomposition property of the so-called decomposable summary inequality measures (Bourguignon 1979; Cowell 1980; Shorrocks 1980). In both cases, the underlying logic is that the aggregate mean income (or inequality measure) in a population is the result of the aggregation of various socio-demographic groups of income sources. Thus, changes of overall mean or inequality measure can be explained by identifying changes in the means and inequality measures within those groups or income sources, and in their weights in the population or in total income.

The new focus on poverty and inequality reduction, which increasingly drives development policy, currently requires new analysis techniques on the shape of distribution, for example, in the vicinity of and below the poverty line. In terms of the Oaxaca-Blinder approach, the issue is to know not so much whether mean earnings are lower for women than for men, since the former may have less average education, as whether the differences are greater or smaller for the bottom part of the earnings distribution. Answering this kind of questions requires handling the whole distribution rather than summary measures. To assess the relevance of the various factors on income inequality changes, handling whole distributions, a microeconomic decomposition methodology first proposed by Bourguignon, Ferreira and Lustig (1998) was tailored to the Paraguayan case.<sup>27</sup>

---

27 Variants of the basic methodology have been applied in Altimir, Beccaria and González Rozada (2000), Bourguignon, Gurgand and Fournier (1999), Bouillon, Gasparini, Marchionni and Sosa (2000), Legovini and Lustig (1998) and Ferreira and Paes de Barros (1999), amongst others.

*The basic model*

The decomposition of a distributional change essentially consists of contrasting representations of the income-generation process (evaluating differences in estimated parameters) for two different distributions (two points in time) on the one hand, and accounting for changes in the joint distribution of endowments, on the other. Bourguignon, Ferreira and Lustig (1998) use parametric representation of inequality changes, because the parameters lead themselves directly to relevant economic interpretations.

More formally, a parametric representation of an income generation process can be defined by a set of variables  $X = (V, W)$  where specific combinations of individual characteristics  $V$  and the values of these individual characteristics  $W$  are defining groups. A general parametric representation of the conditional functions  $g^*(y|V, W)$  and  $h^*(V|W)$  relates  $y$  and  $(V, W)$  on the one hand, and  $V$  and  $W$  on the other, according to some predetermined functional form. These relationships can be denoted as follows:

$$y = G[V, W, \varepsilon; \Omega_\tau] \quad (3.1)$$

$$V = H[W, \eta; \Phi_\tau] \quad (3.2)$$

Where  $\Omega_\tau$  and  $\Phi_\tau$  are sets of parameters and  $\varepsilon$  and  $\eta$  are random variables -  $\eta$  is a vector if  $V$  is a vector. These random variables play a similar role to the residual term in standard regressions. They are meant to represent the dispersion of income  $y$  or individual characteristics  $V$  for given values of individual characteristics  $(V, W)$ , and  $W$ , respectively. They are also assumed to be distributed independently of these characteristics, according to density functions  $\pi^\tau$  and  $\mu^\tau$ . The functions  $G$  and  $H$  have pre-imposed functional forms.

If this model were to be applied to the distribution of individual earnings, the methodology would be rather simple. Ignoring the partition of  $X$  into exogenous characteristics ( $W$ ) and non-exogenous individual characteristics ( $V$ ), a simple parametric representation of individual earnings as a function of individual characteristics is given by:

$$\text{Log } y_i^\tau = X \cdot \Omega + \varepsilon \quad (3.3)$$

In this particular case, the function of  $G(\cdot)$  is thus as follows:

$$G(X, \varepsilon; \Omega) = e^{X \cdot \Omega + \varepsilon} \quad (3.4)$$

To obtain estimates for the set of parameters  $\Omega$  and for the distribution of the random term  $\varepsilon$ , one may rely on standard econometric techniques. Running a



regression on samples of the observations  $i$  available at time  $\tau$ ,

$$\text{Log } y_i^\tau = X_i^\tau \cdot \Omega_\tau + \varepsilon_i^\tau \quad (3.5)$$

yields an estimate of the set of parameters  $\Omega_\tau$ , as well as of the distribution  $\pi_\tau$  of the random term. Then, counterfactuals  $D$  can be computed easily. Without the  $(V, W)$  distinction, a counterfactual is defined as  $D(\chi, \pi; \Omega)$ , where  $\chi(W, \eta)$  is the joint distribution of the exogenous components of  $(V, W)$ . In discrete representation  $\{y_i\}^\tau = (y_1, y_2, \dots, y_{N_\tau})$  of the distribution at time  $\tau$ , where  $N_\tau$  is the number of observations in the sample available at time  $\tau = t, t'$ , it is identically the case that

$$D(\chi_t, \pi_t, \Omega_t) = \{y_i\}_t'. \quad (3.6)$$

The counterfactual,  $D(\chi_t, \pi_t, \Omega_{t'}) = \{y_i\}_\Omega^{t \rightarrow t'}$ , is obtained by computing:

$$\text{Log } (y_i)_\Omega^{t \rightarrow t'} = X_i^t \cdot \hat{\Omega}_{t'} + \hat{\varepsilon}_i^t \quad \text{for } i = 1, 2, \dots, N_t \quad (3.7)$$

where the notation  $\wedge$  stands for OLS estimates. The counterfactual is thus obtained by simulating the preceding model on the sample of observations available at time  $t$ . This simulation shows what would have been the earnings of each individual of the sample if the returns to each of the observed characteristics had been those observed at time  $t'$  rather than the actual returns at time  $t$ . The returns to the unobservable characteristics that may be behind the residual term  $\hat{\varepsilon}_i^t$  are supposed to be unchanged, nonetheless. This is equivalent to the evaluation of the price effect for observed characteristics of the Oaxaca-Blinder calculation. The difference is that the evaluation is carried out for every individual in the sample. The counterfactual of the distribution of the random term  $D(\chi_t, \pi_{t'}, \Omega_t) = \{y_i\}_\Omega^{t \rightarrow t'}$  is a little more difficult to construct. Importing the distribution of residuals from time  $t'$  to time  $t$  requires an operation known as *rank-preserving-transformation*, whereby the residual in the  $n^{\text{th}}$  percentile (of residuals) at time  $t$  is replaced by the residual in the  $n^{\text{th}}$  percentile at time  $t'$ , for all  $n$ . As the operation is not immediate when the number of observations is not the same in the two samples, an approximate solution is used. It consists of assuming that both distributions of residual terms are the same up to a proportional transformation. An example would be if residuals were normally distributed, with mean zero. The rank-preserving-transformation is then equivalent to multiplying the residual observed at time  $t$  by the ratio of standard deviation at time  $t'$  and  $t$ .  $D(\chi_t, \pi_{t'}, \Omega_t) = \{y_i\}_\Omega^{t \rightarrow t'}$  is thus defined by:

$$\text{Log } (y_i)_\pi^{t \rightarrow t'} = X_i^t \cdot \hat{\Omega}_t + \hat{\varepsilon}_i^t \cdot (\hat{\sigma}_\varepsilon^{t'} / \hat{\sigma}_\varepsilon^t) \quad \text{for } i = 1, 2, \dots, N_t \quad (3.8)$$



With those counterfactuals at hand, estimates for the contribution to the observed overall distributional change between  $t$  and  $t'$  of the change in the  $\Omega$  parameters, in the distribution of residuals ( $\pi$ ), and possibly even of these two changes taken together, may easily be found. The effect of changing the distribution of individual endowments,  $X$ , is obtained as the complement of the two previous changes:

$$\{y_i\}^{t'} - D(\chi_t, \pi_{t'}, \Omega_{t'}). \quad (3.9)$$

#### *Adaptation to Paraguayan data*

Let  $Y_{it}$  be individual's  $i$  labour income at time  $t$ , which can be written as a function  $F$  of the vector  $X_{it}$  of individual observable characteristics affecting wages and employment, the vector  $e_{it}$  of unobservable characteristics, the vector  $b_t$  of parameters that determine market hourly wages and the vector  $l_t$  of parameters that affect employment outcomes (participation and hours of work).

$$Y_{it} = F(X_{it}, e_{it}, \beta_t, \lambda_t) \quad i=1, \dots, N \quad (3.10)$$

The distribution of individual labour income can be represented as:

$$D_t = \{Y_{1t}, \dots, Y_{Nt}\} \quad (3.11)$$

We can simulate individual labour incomes by changing one or more arguments in equation (3.10). For instance, the following expression represents labour income that individual's  $i$  would have earned in time  $t$  if the parameters determining wages had been those of time  $t'$ , keeping all other things constant.

$$Y_{it}(\beta_{t'}, \cdot) = F(X_{it}, e_{it}, \beta_{t'}, \lambda_t) \quad i=1, \dots, N \quad (3.12)$$

More generally, we can define  $Y_{it}(k_t')$  where  $k$  is any set of arguments in (3.10). Hence, the simulated distribution will be:

$$D_t(k_t') = \{Y_{1t}(k_t'), \dots, Y_{Nt}(k_t')\} \quad (3.13)$$

The contribution to the overall change in the distribution of a change in  $k$  between  $t$  and  $t'$ , holding all else constant, can be obtained by comparing (3.11) and (3.13). Although we can make the comparisons in terms of the whole distributions, in this paper we only compare inequality indices  $I(D)$ . Therefore, the effect of a change in argument  $k$  is defined by:

$$I[(D_t(k_t'))] - I(D_t) \quad (3.14)$$

---

28 In the empirical implementation labour income distribution only is computed

This paper is devoted to discuss the following effects:

- *Returns to education* ( $k = \beta_{ed}$ ) measures the effect of changes in the parameters that relate education to hourly wages ( $\beta_{ed}$ ) on inequality.
- *Gender wage gap* ( $k = \beta_g$ ) measures the effect of changes in the parameters that relate gender to hourly wages ( $\beta_g$ ) on inequality.
- *Returns to experience* ( $k = \beta_x$ ) measures the effect of changes in the parameters that relate experience (or age) to hourly wages ( $\beta_x$ ) on inequality.
- *Endowment and returns to unobservable factors* ( $k = e^w$ ) measures the effect of changes in the unobservable factors and their remunerations affecting hourly wages ( $e^w$ ) on inequality.
- *Hours of work and employment* ( $k = \lambda$ ) measures the effect of changes in the parameters that determine hours of work and labour market participation ( $\lambda$ ) on inequality.
- *Education* ( $k = X$ ) measures the effect of changes in the educational levels of the population ( $X$ ) on inequality.

The previous discussion refers to the distribution of individual earnings. However, it is more relevant from a social point of view to study the distribution of household income since a person's utility usually depends not on their own earnings, but on their household income and demographic composition. Following Buhmann et al. (1988), equivalent household income is given by:

$$Y_{iht}^q = \sum_{jeh} (Y_{jy} + Y_{jy}^0) / \left( \sum_{jeh} a_j \right)^{\theta} \quad i = 1, \dots, N \quad (3.15)$$

where  $Y^q$  stands for equivalent household income,  $h$  is the household,  $Y^0$  is the income from other sources,  $a$  is the equivalent adult and  $q$  captures household economies of scale. The distribution of equivalent household income can be expressed as:

$$D_t = \{ Y_{1t}^q, \dots, Y_{Nt}^q \} \quad (3.16)$$

Changing argument  $k$  to its value in  $t'$  yields the following simulated equivalent household income in year  $t$ :

$$Y_{iht}(k_{t'}) = \sum_{jeh} (Y_{jy}(k_{t'}) + Y_{jy}^0) / \left( \sum_{jeh} a_j \right)^{\theta} \quad i = 1, \dots, N \quad (3.17)$$

Hence, the simulated distribution is:

$$D_t^q(k_{t'}) = \{ Y_{1t}^q(k_{t'}), \dots, Y_{Nt}^q(k_{t'}) \} \quad (3.18)$$

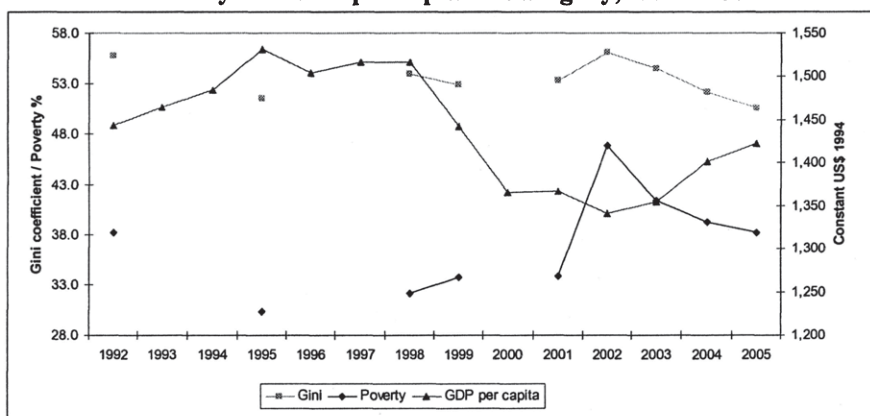
The effect of a change in argument  $k$ , holding all else constant, on equivalent household inequality is given by<sup>29</sup>:

$$I[(D_t^q(k_t'))] - I(D_t^q). \quad (3.19)$$

### 3.3 Income inequality in Paraguay: basic facts and sources of change

Per capita income inequality in Paraguay during the nineties has a generally negative tendency, rising during economic and political crisis between 1999 and 2002 and then recovering its path towards reduction. Figure 3.1 shows the Gini coefficient of per capita household income between 1992 and 2005 in Paraguay, combined with poverty headcount measures and GDP per capita.<sup>30</sup> Only since 2001 is there a yearly update of poverty and inequality measures in Paraguay.

**Figure 3.1 Gini coefficient of Per Capita Household Income Distribution, Poverty and GDP per capita in Paraguay, 1992 - 2005**



Source: Author's calculations based on EPH and EIH of National University and DGEEC for poverty and inequality. GDP data from the Banco Central del Paraguay.

For simplicity, this study focuses on three years of relative macroeconomic sta-

29 In the empirical implementation we ignore income from other than labour sources  $Y_{it}^q$  and we consider all individuals such that  $Y_{it}^q \geq 0$  and  $Y_{it}^q(k_t') \geq 0$ .

30 The 1992 survey was carried out by the Universidad Nacional de Asuncion, while the 1995 to 2005 surveys are Encuestas Permanentes de Hogares (EPH) or Encuestas Integrales de Hogares (EIH, only 1997/98 and 2000/01) carried out by the National Statistical Office (Direccion General de Estadistica, Encuestas y Censo - DGEEC).

bility separated by almost equal intervals: 1992, 1997/98 and 2005<sup>31</sup>. The analysis was restricted to labour income mainly for two reasons. (i) Permanent Household Surveys (EPH) and Integrated Household Surveys (EIH) have various deficiencies in capturing capital income, and (ii) modelling capital income and retirement payments is not an easy task, especially considering the scarce information included in the surveys. Households whose heads or spouses are older than 64, or receive retirement payments, were ignored. The following analysis concentrates on the distribution of individual labour income<sup>32</sup> and on the distribution of equivalent<sup>33</sup> household labour income.

**Table 3.1 Income Distribution in Paraguay, Selected Years (Gini coefficient)**

Type of distribution	1992	1997/98	2005
Labour income	0.589	0.553	0.490
Equivalent household labour income	0.606	0.584	0.528

Source: Author's calculations based on EPH and EIH of National University and DGEEC.

Table 3.1 shows the basic facts to be characterized in the paper: inequality in individual labour income and in equivalent household labour income, as measured by the Gini, dropped almost ten percentage points between 1992 and 2005. Interestingly, labour income inequality reduction is stronger than equivalent household labour income reduction. Reduction between 1997/98 and 2005 is much stronger than between 1992 and 1997/98, even if the first period contains a sub-period of economic, political and social crisis where inequality grew. One possible reason for the stronger decrease of inequality in labour income, compared with inequality in household income, may lie in changes in non-labour income sources. Main non-labour income sources in Paraguayan income survey measures are the imputed value on own housing and transfers from family members inside and outside the country. Mainly the poor are benefited from imputed values for own housing, no matter how precarious their housing might be. There is an underdeveloped market for renting houses or apartments. It is nearly exclusively an urban phenomenon, restricted to rich households that can afford to pay

31 1992 and 1995 surveys report income for September, EIH 1997/98 for February of 1998, 1999 for September, 2000/01 for March 2001 and starting from 2002 all incomes for November.

32 Labour income comprises wage earnings and self-employed earnings.

33 Following Buhmann et al. (1988) the equivalent household income is obtained by dividing household income by the number of equivalent adults raised to 0.8, a parameter which implies mild household economies of scale. Since there is no official measurement of equivalent adult scales from DGEEC in Paraguay, general scales with 0.4 for children < 5 years of age, 0.5 for children > 5 years of age, and < 16 years of age and 1.0 for all individuals > 14 years of age were applied.

rent. Cash transfer from private sector to Paraguayan households came mainly from Argentina during the nineties and since 2003 with an increasing degree from Spain and the US. Private sector cash transfers are distributed all over society, from very poor to very rich. In 1992 private cash transfers were 1.8% of total household income, 4.3% in 1997/98 and 4.7% in 2005. Other possible sources for differences in changes between labour and household income are changes in marriage markets or changes in household composition (total number of individuals per household). Regression results in the next chapter will give some hints on these points. There are almost no public cash transfers, apart from very small pension payments, but pension recipients were excluded from the analysis by definition. An innumerable number of factors may have caused the changes in inequality documented in Table 3.1. We will concentrate on seven of these: (i) returns to education, (ii) the gender wage gap,<sup>34</sup> (iii) returns to experience, (iv) the dispersion in the endowment of unobservable factors and their returns, (v) hours of work, (vi) labour market participation, and (vii) the education of the working-able population.

### 3.3.1 Returns to education

An increase in the returns to education implies a widening of the wage gap between high and low educated workers, which in turn would imply a more unequal distribution of individual earnings and probably a more unequal distribution of household income. Table 3.2 shows hourly earnings in constant Guaranies (Gs.) in 2005 for workers between 12 and 64 with valid and complete answers. The average wage increased 5.3% between 1992 and 1997/98 and dropped 7.8% during the next seven years. Changes were not consistent among educational groups. In the first period of the analysis we had winners and losers. While incomes for workers who had not finished primary education increased slightly, the wages for the next two groups, complete primary education and incomplete secondary education, dropped considerably. Dramatic increases were observed for complete secondary and complete or incomplete college education. In the 1997/98 – 2005 period the losses of income were generalized for all income groups except for primary incomplete education. Losses for higher education are stronger than for lower education. Table 3.2 is a first piece of evidence that changes in relative wages among schooling groups implied an increase in earnings inequality between 1992 and 1997/98 and a decrease thereafter.

---

34 Throughout this paper “wage” refers to hourly labour income earned by wage-workers and self-employed workers.

**Table 3.2 Hourly Earnings by Educational Level in Paraguay, Selected Years**

Educational Level	Means (Gs. 2005)			Changes (percent)		
	1992	1997/98	2005	1992-1998	1998-2005	1992-2005
Primary incomplete	3,000.6	3,131.7	3,207.5	4.4	2.4	6.9
Primary complete	4,635.7	4,276.3	3,547.7	-7.8	-17.0	-23.5
Secondary incompl.	6,035.0	5,477.5	3,998.4	-9.2	-27.0	-33.7
Secondary complete	8,145.8	10,226.9	6,525.8	25.5	-36.2	-19.9
College incomplete	11,763.0	16,081.3	10,040.8	36.7	-37.6	-14.6
College complete	20,957.6	28,811.8	18,594.8	37.5	-35.5	-11.3
Total	5,746.9	6,051.1	5,580.1	5.3	-7.8	-2.9

Source: Author's calculations based on EPH and EIH of National University and DGEEC.

Table 3.3 shows the results of Mincerian log hourly earning functions estimated using the Heckman procedure to correct for sample selection. The first three columns refer to household heads (mostly men) and the rest to spouses (nearly all women) and other members of the family (roughly half men and half women), respectively. A gender dummy, age and age squared and a dummy for youngsters less than 18 years old (only relevant for other members) are included in the regression. In addition to these variables, the selection equation includes marital status, number of children and a dummy that takes the value "1" when the individual attends school. Following Bourguignon et al. (1999) it is assumed that labour market participation choices are made within the household in a sequential fashion. Spouses take the head's labour market status into consideration to decide whether to enter the labour market or not. Other members of the family consider both the head and the spouse labour market status.

The coefficients for years of education are positive and returns to education are always positive. For family heads in 1992, one additional year of schooling increases in the mean hourly wages in 11.3%, keeping all other factors constant. The same figure for 1997/98 and 2005 is 15.4% and 10.9%, respectively. It is interesting to observe that spouses hourly wage determination follows the same path of the heads hourly wage. It also increases (from 7.5% to 13.2%) between 1992 and 1997/98 and then drops again (to 10.7%) in 2005, but there is no such path for other family members who lose income in each period. Figure 3.2 shows the predicted hourly earnings for all different years of education. The first panel refers to male heads and the second to other male members, both with age kept constant at 40.



**Table 3.3 Log-Hourly Earnings Equation Applied to Paraguay, Selected Years**

	Head of household			Spouse			Other family members		
	1992	1997/98	2005	1992	1997/98	2005	1992	1997/98	2005
Dependent variable: log (ln) hourly earnings equation									
Years of education	0.1136 (32.26) ***	0.1543 (34.51)***	0.1085 (24.8)***	0.0752 (8.99) ***	0.1321 (11.15)***	0.1073 (9.15)***	0.1284 (27.04) ***	0.0772 (12.9)***	0.0768 (14.77)***
Male	0.0256 (0.37)	0.1405 (1.93)*	0.1199 (2.21)**	-0.1386 (-0.94)	-0.0924 (-0.51)	0.0241 (0.16)	0.3447 (7.15) ***	-0.3199 (-7.64)***	-0.1713 (-4.75)***
Age	0.0505 (4.73) ***	0.0649 (5.11)***	0.0522 (4.17)***	-0.0203 (-0.94)	0.0167 (0.59)	-0.0041 (-0.14)	0.0720 (5.99) ***	-0.0654 (-4.27)***	-0.0055 (-0.4)
Age squared	-0.0005 (-3.82) ***	-0.0007 (-4.65)***	-0.0005 (-3.34)***	0.0003 (1.24)	0.0000 (-0.08)	0.0001 (0.36)	-0.0008 (-4.81) ***	0.0010 (4.44)***	0.0002 (1.04)
Married	0.0499 (0.94)	0.0614 (0.99)	0.1747 (3.21)***						
Number of children	-0.0249 (-2.85) **	-0.0480 (-4.85)***	-0.0131 (-1.26)	0.0329 (1.72) *	0.0207 (1.02)	0.0301 (1.61)			
Younger than 18							-0.0421 (-0.65)	-0.1258 (-1.75)*	0.0225 (0.35)
Attending school	0.4617 (3.45) **	-0.3056 (-1.87)*	0.2641 (2.69)***	0.3160 (1.51)	0.0185 (0.09)	0.1834 (1.35)	-0.6561 (-4.69) ***	0.4409 (7.66)***	0.2816 (5.8)***
Head employed					-0.0024 (-0.02)	-0.0131 (-0.15)			
Spouse employed		0.1101 (2.69)***	-0.1526 (-3.74)***						
Constant	6.5217 (27.84) ***	5.7516 (21.8)***	6.0895 (23.09)***	8.5570 (16.95) ***	6.9314 (9.87)***	7.3270 (9.64)***	5.4475 (25.07) ***	9.3568 (35.93)***	7.8517 (32.46)***
Selection equation: dependent variable = 1 if hourly earnings > 0									
Years of education	0.0397 (5.66) ***	0.0322 (3.6)***	0.0416 (6.63)***	0.0632 (10.6) ***	0.0627 (7.89)***	0.0625 (9.21)***	0.0856 (13.36) ***	0.0313 (5.32)***	0.0351 (6.46)***
Male	0.8162 (8.99) ***	1.1582 (13.01)***	0.6800 (10.3)***	0.8226 (5.31) ***	1.5301 (9.29)***	1.0532 (9.95)***	1.0269 (22.64) ***	0.4591 (12.46)***	0.3073 (8.5)***
Age	0.0827 (4.6) ***	0.1137 (5.54)***	0.0920 (5.78)***	0.1263 (7.8) ***	0.1470 (-8.45)***	0.1446 (-8.15)***	0.1208 (8.83) ***	0.1703 (12.52)***	0.1588 (12.03)***

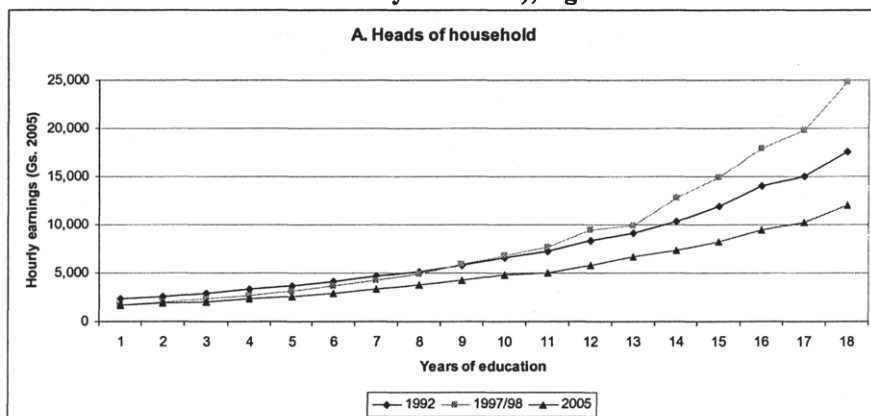
Age squared	-0.0012 (-5.82) ***	-0.0015 (-6.29) ***	-0.0013 (-7.05) ***	-0.0016 (-8.02) ***	-0.0017 (-8.45) ***	-0.0017 (-8.15) ***	-0.0017 (-9.29) ***	-0.0023 (-11.79) ***	-0.0021 (-11.44) ***
Married	0.2360 (2.65) *	0.1988 (1.83) *	-0.0662 (-0.89)						
Number of children	0.0023 (0.14)	-0.0427 (-2.52) **	-0.0265 (-2.03) **	-0.0786 (-4.97) ***	-0.0845 (-6.1) ***	-0.0442 (-2.9) ***			
Younger than 18							-0.1873 (-2.44) **	0.0021 (0.03)	-0.0453 (-0.7)
Attending school	-0.6388 (-3.35) **	-0.4046 (-1.66) *	-0.0641 (-0.4)	-0.3678 (-1.91)	0.2566 (1.14)	0.4632 (2.99) ***	-2.3547 (-32.67) ***	-0.6566 (-13.59) ***	-0.5385 (-11.4) ***
Head employed					-0.2269 (-2.3) **	-0.0389 (-0.42)			
Spouse employed		-0.2405 (-2.8) ***	-0.0763 (-1.33)						
Urban area	0.2152 (2.62) *	-0.2547 (-3.19) ***	0.7286 (14.37) ***	0.5563 (8.06) ***	0.2263 (3.55) ***	0.4028 (7.4) ***	-0.2762 (-4.51) ***	0.4805 (13.85) ***	0.5439 (15.37) ***
Constant	-1.2408 (-3.43) **	-1.4395 (-3.49) ***	-1.8780 (-5.72) ***	-3.5590 (-11.88) ***	-3.1358 (-10) ***	-3.6446 (-10.75) ***	-2.2127 (-9.88) ***	-3.2267 (-14.84) ***	-3.1761 (-15) ***
Number of observations	4,158	3,690	3,792	3,260	2,925	2,787	6,143	6,116	6,356
Censored observations	544	391	1,107	2,257	1,609	1,646	3,657	3,839	4,081
Uncensored observations	3,614	3,299	2,685	1,003	1,316	1,141	2,486	2,277	2,275
Chi squared	1,319.1	1,482.7	714.1	136.1	277.1	217.9	1,070.1	470.8	425.8
Log likelihood	-5,886.8	-5,651.1	-5,352.3	-2,955.3	-3,645.9	-3,077.5	-5,077.8	-5,789.3	-5,773.4
Rho	-0.2797	0.3074	-0.3532	-0.5951	-0.3101	-0.4089	0.5197	-0.8431	-0.7718
Sigma	0.8643	1.0072	0.8864	0.8852	1.0301	0.9115	0.8478	1.0756	0.9290
Lambda	-0.2417	0.3096	-0.3130	-0.5267	-0.3194	-0.3728	0.4406	-0.9068	-0.7171

Note: Data represent Heckman maximum likelihood estimation; z values are in parentheses. Data cover all individuals between 12 and 64 with valid answers.

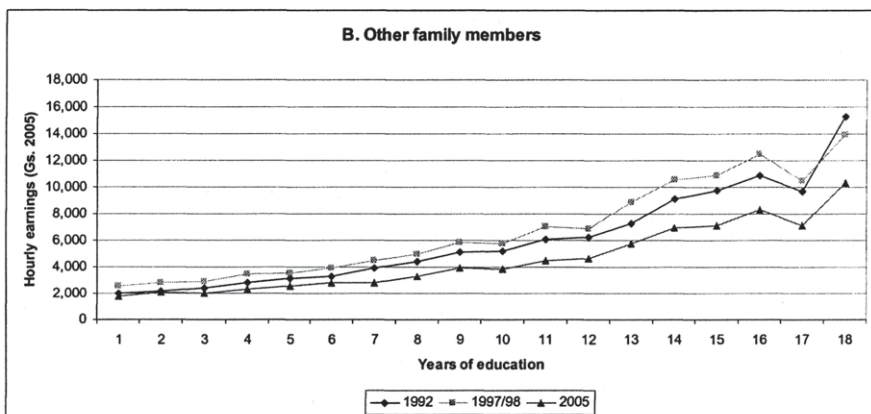
Source: Author's calculations based on EPH and EIH of National University and DGEEC.



**Figure 3.2 Hourly Earnings-Education Profiles for Men (Heads of Household and Other Family Members), Age 40**



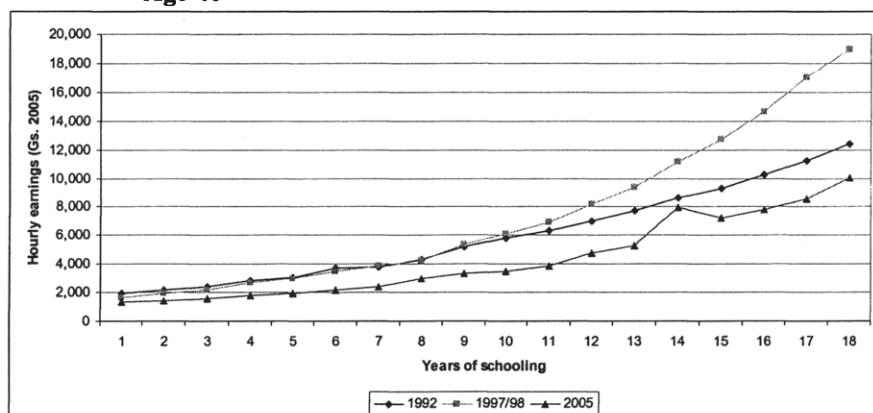
Source: Predicted hourly earnings from models on table 3.3.



Source: Predicted hourly earnings from models on table 3.3.

The wage-education profiles for family heads have a marked positive slope and are almost parallel everywhere, except for the substantial increase in the slope of 1997/98 in the highest levels, as from 13 years of education. This certainly contributes to increase earnings inequality among household heads with different educational levels. For male other-members the wage-education profile we have almost parallel slopes for all periods with some differences only for 17 and 18 years of education. So the changes in earnings of other family members could contribute only for high levels to widen inequality.

**Figure 3.3 Hourly Earnings-Education Profiles for Women (Spouses), Age 40**



Source: Predicted hourly earnings from models on table 3.3.

Figure 3.3 shows the profiles for 40 year old females. As in the case for men, the wage-education profiles show an increasing slope between 1992 and 1997/98, and an opposite movement between 1997/98 and 2005. It is interesting to see that for all three groups (household heads, spouses and other family members) there is a strong increase in returns for higher education in the 1992 – 1997/98 period. This is a real observation and unbiased by model specification, since years of education did not enter the model in its squared form. The reason for high returns for higher education might lie in pure market effects. Even if the Paraguayan economy and its industry are not very sophisticated, there is still a need for highly qualified human resources in any managerial post. As Table 3.10 will show further ahead, the percentage of the working force which completed college education did not exceed 2.4% over the whole observation period. Incomplete college education increased from 4.6% in 1992 to 9.4% in 2005, but nevertheless, these levels still remain low, and in a way, can explain why 2005 return profiles are much “smoother” than in previous years.

Summarizing, there is evidence of a positive relationship between hourly earnings and education which induces differences in incomes among individuals with different education. According to the evidence presented these differences, along with inequality, have increased between 1992 and 1997/98, and decreased in the next seven years. During this last period the wage-education profile has become smoother and less convex, which implies inequality reduction. Although this phenomenon seems widespread across groups, it appears to be more relevant for the groups of household heads and spouses.

### 3.3.2 Gender wage gap

Table 3.4 presents mean hourly wages by gender. Wages were higher for males in every year. Nevertheless, there are interesting dynamics within the gender wage gap which decreased from more than 16% in 1992 to less than 2% in 1997/98, and then increased again to some 6% in 2005. Over the whole period, female mean wage gain was about 6.9%, while male wages increased only in 0.6%. This implies inequality reduction over the whole observation period.

**Table 3.4 Hourly Earnings by Gender in Paraguay, Selected Years**

Gender	Means (Gs. 2005)			Changes (percent)		
	1992	1997/98	2005	1992-1997/98	1997/98-2005	1992-2005
Female	5,679	7,390	6,072	30.1	-17.8	6.9
Male	6,443	7,472	6,484	16.0	-13.2	0.6
Total	6,046	7,411	6,318	22.6	-14.8	4.5

Source: Author's calculations based on EPH and EIH of National University and DGEEC.

A conditional analysis also shows a shrinking gap for household heads. From Table 3.3 the coefficient for the male dummy is not always positive and significant, but clearly decreasing over time. Surprisingly, for other members we observe an important male income loss in 1997/98. However, since the number of working individuals in this group is considerably less than in the household heads group, the global conclusion of a narrowing gender wage gap holds. This shrinking gap has undoubtedly been an equalizing factor on the individual earnings distribution. The effect of this phenomenon on the equivalent household labour income distribution will basically depend on the position of working women in that distribution. Section 5 will expand on this further ahead.

### 3.3.3 Returns to experience

Age is used in this paper as a proxy for experience in the labour market. Table 3.5 shows average hourly earnings for different age groups. In general the wage-age profile has an inverted U shape. Between 1992 and 2005, hourly wages only increased for labour force younger than 30 years of age, which is the worst paid group of workers. In principle, this would imply an equalizing effect on the earnings distribution. However, the main benefit is for men and women less than 20 years of age. In 1992 they represented less than 13% of the total working population. During the 1997/98 to 2005 period these gains were lost again. All age groups lost income considerably in the 1997/98 to 2005 period, with a stronger loss for young workers. Once more, since this group is small, its effect is not big on the overall distribution. Since all other age groups lost more or less similar percentages of their wage there seems to be only a very small, but in the end positive equalizing effect.

**Table 3.5 Hourly Earnings by Age Groups, Paraguay, Selected Years**

Age	Means (Gs. 2005)			Changes (percent)		
	1992	1997/98	2005	1992-1997/98	1997/98-2005	1992-2005
12 to 19	2,726	3,431	2,019	25.9	-41.1	-25.9
20 to 29	5,026	5,128	4,198	2.0	-18.1	-16.5
30 to 39	6,671	6,096	4,942	-8.6	-18.9	-25.9
40 to 49	6,769	5,943	4,818	-12.2	-18.9	-28.8
50 to 59	6,940	6,034	4,652	-13.1	-22.9	-33.0
60 to 64	6,155	5,351	4,381	-13.1	-18.1	-28.8

Source: Author's calculations based on EPH and EIH of National University and DGEEC.

Throughout the whole observation period, the age group between 20 and 29 years of age is the less affected one. This implies certain equalizing effects, since in 2005 this group represented almost 40% of the working force. More negatively affected groups are those for working force older than 40 years of age. Nevertheless, in 2005 all three of these groups together represented less than half of the working force (48%). Since their mean wages are lower than the mean wages of the largest age group (20 to 29), their inequality increasing effects should be lower than the equalizing effects of the 20 to 29 year old group. Summing up, there are some reasons to believe that changes in the returns to experience have led to higher inequality and some reasons to believe the opposite. The analysis of Section 5 will help us to assess the quantitative relevance of each argument.

### 3.3.4 Unobservable factors

Earnings equations allow the estimation of returns to observable factors like education and experience. The error term is usually interpreted as capturing the joint effect of the endowment of non-observable factors (like individual ability) and its market value on earnings. In general terms, the variance of this error term captures the contribution of dispersion in unobservable factors to general inequality. Table 3.3 reports the standard deviation of the error terms of each log hourly earnings equation (labelled as "sigma"). For instance, for household heads the standard deviation took a value of 0.86 in 1992, 1.01 in 1997/98, and 0.89 in 2005. The substantial increase between 1992 and 1997/98 is also present in the spouses and other members' equations, as well as the reduction towards 2005. According to these results, the effect of changes in unobservable factors would have been strongly unequalizing between 1992 and 1997/98, reducing some of this additional inequality in the 1997/98 to 2005 period.

### 3.3.5 Hours of work

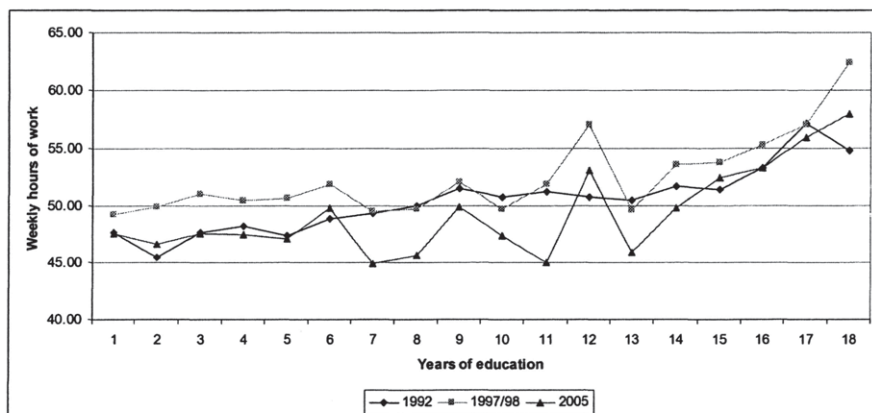
During the period under analysis there has been an increase in weekly hours of work between 1992 and 1997/98 and a decrease in the next seven years, almost to the same overall level observed in 1992. Table 3.6 classifies workers by educational level and records the average hours of work of each group. While there are clear gains for completed cycles in the 1992 to 1997/98 period, which deepens inequality, losses in 1997/98 to 2005 are more equally distributed. So, over the whole period we still observe important gains for completed cycles of secondary and college education, but at the same time important losses for groups who did not complete an education cycle. Since gains are larger for higher educational groups, this change would have a non-negligible unequalizing effect in the individual earnings distribution. A conditional analysis yields similar results. Figure 3.4 shows predicted weekly hours of work for male heads from the Tobit censored data model presented in Table 3.7. While hours clearly increased between 1992 and 1997/98 for less-educated (1 to 6 years of education) and for well educated (more than 13 years of education) male heads workers, changes in hours for the rest of the educational groups were only marginal between 1992 and 1997/98. In the

**Table 3.6 Weekly Hours of Work by Educational Levels in Paraguay, Selected Years**

Educational Level	Means			Changes (percent)		
	1992	1997/98	2005	1992-1997/98	1997/98-2005	1992-2005
Primary incomplete	47.9	51.0	45.4	6.0	-12.3	-5.5
Primary complete	50.3	51.3	49.3	2.0	-4.0	-1.9
Secondary incomplete	50.1	51.9	49.6	3.4	-4.7	-1.1
Secondary complete	50.0	48.7	50.6	-2.6	3.8	1.2
College incomplete	45.5	46.5	44.8	2.2	-3.8	-1.5
College complete	45.6	48.1	48.0	5.2	-0.3	4.9
Total	49.0	50.4	47.9	2.8	-5.2	-2.2

Source: Author's calculations based on EPH and EIH of National University and DGEEC.

**Figure 3.4 Weekly Hours of Work by Educational Level for Men (Heads of Household), Age 40**



Source: Predicted hourly earnings from models on table 3.7.

following seven years, the reduction of weekly hours of work is generalized for levels above six years of education. Consequently, we have evidence over the whole observation period for an equalizing effect for workers between 1 and 6 years of education, an unequalizing effect for workers between 7 and 15 years of education and a nearly neutral effect for the highest education levels.

### 3.3.6 Labour market participation

Household income inequality can change, not only after changes in hours of work but also as a result of changes in labour market participation. In Table 3.7 individuals are grouped according to whether they are employed, unemployed or inactive. The percentage of unemployed individuals dropped from 4.4% in 1992 to 3.4% in 1997/98 and rose again to 3.8% in 2005. However, notice that the increase in unemployment between 1997/98 and 2005 in 0.4 percentage points was accompanied by a decrease in inactivity in 3.2 percentage points. Notice that for inequality measures it is irrelevant whether an individual has zero income as a result of unemployment or due to not looking for a job. Hence the important indicator for possible inequality changes is the overall employment rate which increased from 57% to 61% and 63% during the observation period. These changes might have played a role in inequality changes depending on the distribution of wage levels and hours of work obtained by the additional working force.

**Table 3.7 Labour Status by Household Role, Paraguay, Selected Years**

	Proportions by group (percent)		
	1992	1997/98	2005
<b><i>All</i></b>			
Employed	56.96	60.55	63.36
Unemployed	4.40	3.35	3.75
Inactive	38.64	36.10	32.89
<b><i>Head</i></b>			
Employed	89.75	88.73	87.11
Unemployed	2.02	2.33	2.61
Inactive	8.23	8.94	10.28
<b><i>Spouse</i></b>			
Employed	23.80	47.38	57.10
Unemployed	6.76	2.15	2.44
Inactive	69.44	50.46	40.46
<b><i>Other</i></b>			
Employed	41.92	49.84	51.93
Unemployed	5.60	4.54	5.00
Inactive	52.48	45.61	43.06

Source: Author's calculations based on EPH and EIH of National University and DGEEC.



**Table 3.8 Hours of Work Equation for Paraguay, Selected Years**

	Head of household			Spouse			Other family members		
	1992	1997/98	2005	1992	1997/98	2005	1992	1997/98	2005
Years of schooling	0.0270 (0.34)	0.0575 (0.55)	0.2827 (2.71)***	3.8209 (12.59)***	2.2727 (9.67)***	1.7095 (8.15)***	1.3522 (9.44)***	1.1323 (7.1)***	0.7832 (5.01)***
Male	16.6394 (12.43)***	25.3464 (18.32)***	21.1308 (16.46)***	51.5352 (7)***	43.5240 (10.77)***	39.9850 (12.71)***	27.9596 (24.45)***	23.0687 (22.24)***	21.6306 (20.34)***
Age	1.4114 (5.91)***	1.2101 (4.25)***	2.2679 (7.4)***	6.9762 (8.06)***	5.4225 (8.84)***	5.3229 (9.05)***	3.2107 (9.05)***	4.6314 (11.92)***	4.8489 (12.46)***
Age squared	-0.0215 (-7.69)***	-0.0194 (-5.82)***	-0.0297 (-8.34)***	-0.0883 (-8.23)***	-0.0651 (-8.62)***	-0.0604 (-8.45)***	-0.0478 (-9.6)***	-0.0695 (-12.54)***	-0.0673 (-12.28)***
Married	4.7894 (3.9)***	0.5888 (0.41)	2.0695 (1.47)						
Children	-0.4056 (-1.94)*	-0.2636 (-1.15)	-0.7328 (-2.85)***	-5.0412 (-6.16)***	-3.7189 (-7.31)***	-1.1558 (-2.29)**			
Younger than 18							-9.1676 (-4.56)***	-1.5339 (-0.81)	-1.3841 (-0.72)
Attending school	-10.9443 (-3.69)***	-13.5369 (-3.65)***	-7.2588 (-2.64)***	-15.8972 (-1.55)	1.4317 (0.19)	4.9718 (1)	-76.3365 (-36.2)***	-32.0246 (-23.68)***	-27.2135 (-19.72)***
Head of household employed					-6.4897 (-1.89)*	4.5380 (1.45)			
Spouse employed		1.0494 (1.09)	0.9154 (0.86)						
Constant	7.4928 (1.59)	12.8926 (2.27)**	-13.0795 (-2.1)**	-173.7523 (-10.72)***	-97.1573 (-8.38)***	-113.2848 (-9.87)***	-46.9400 (-8.22)***	-53.7711 (-8.75)***	-63.1164 (-10.17)***
Number of observations	4,158	3,690	3,792	3,260	2,925	2,786	6,143	6,116	6,356
Censored	574	279	490	1,027	1,396	1,196	2,551	2,716	3,056
Chi squared	804.59	791.28	658.02	367.25	396.83	366.26	4,469.86	2,572.47	2,364.31
Log likelihood	-17,603.9	-16,081.6	-16,078.0	-6,770.5	-8,921.5	-9,078.8	-13,919.3	-18,632.4	-18,414.9
Pseudo R squared	0.0223	0.024	0.0201	0.0264	0.0218	0.0198	0.1383	0.0646	0.0603
Sigma	22.40	24.60	26.61	59.36	44.08	41.85	33.79	35.26	36.42

*Note: Data represent Tobit maximum likelihood estimation; t ratios are in parentheses. Data cover all individuals between 12 and 64 with valid answers.*

*Source: Author's calculations based on EPH and EIH of National University and DGEEC.*



Table 3.7 suggests three different stories in the labour market for heads, spouses and other members. Some household heads lost or quit their jobs, especially during the last seven years, becoming either unemployed or leaving the labour force. In contrast, many (30 percentage points) of the spouses left their homes in search of a job: most of them found one between 1992 and 1997/98, however, some did not during the 1997/98 to 2005 period. The other members of the family were less fortunate: even if the participation rate also increased dramatically (10 percentage points), their unemployment rate remained nearly unchanged, doubling spouses unemployment during the last period.

**Table 3.9 Labour Status and Education, Paraguay, Selected Years**

	Proportions by group (percent)		
	1992	1997/98	2005
<i><b>Primary incomplete</b></i>			
Employed	47.09	55.55	61.54
Unemployed	5.57	2.26	2.25
Inactive	47.35	42.19	36.20
<i><b>Primary complete</b></i>			
Employed	52.26	62.87	64.65
Unemployed	5.59	2.82	3.22
Inactive	42.16	34.31	32.13
<i><b>Secondary incomplete</b></i>			
Employed	44.99	54.30	51.97
Unemployed	4.21	4.72	4.34
Inactive	50.80	40.98	43.70
<i><b>Secondary complete</b></i>			
Employed	57.67	73.40	73.65
Unemployed	4.28	5.45	6.61
Inactive	38.04	21.15	19.75
<i><b>College incomplete</b></i>			
Employed	65.76	82.53	79.31
Unemployed	1.73	4.80	4.60
Inactive	32.52	12.66	16.09
<i><b>College complete</b></i>			
Employed	88.57	89.50	92.42
Unemployed	1.71	1.66	2.89
Inactive	9.71	8.84	4.69
<i><b>All</b></i>			
Employed	56.96	60.55	63.36
Unemployed	4.40	3.35	3.75
Inactive	38.64	36.10	32.89

Source: Author's calculations based on EPH and EIH of National University and DGEEC.

Table 3.9 presents the proportion of employed, unemployed and inactive individuals by educational group. In the 1992 to 1997/98 period, data show a strong increase in employment rates, jointly with a decrease in inactivity rate, for all

educational levels but college complete group. Employment increases and inactivity decreases are stronger for higher educational levels. This should imply an increase in inequality. In the 1997/98 to 2005 period, employment only keeps growing (and inactivity shrinking) for primary education and for complete university education. Overall, for this period we should expect an equalizing effect on income inequality. Over the whole period the unequalizing effect of the first period is expected to be stronger than its compensation in the next period.

### 3.3.7 Education

In Paraguay, as in many developing countries, substantial changes in the educational composition of the population have been taking place during the nineties. Table 3.10 presents the proportion of individuals between 12 and 64 years of age by educational level. Between 1992 and 1997/98 there was a contraction in the proportion of youngsters and adults with incomplete primary education and an expansion for incomplete secondary education. Both are groups with low or up to medium wages. In the 1997/98 to 2005 period the participation of incomplete primary education kept falling, primary complete remained almost unchanged and larger changes were observed in the higher income groups with secondary and college education.

**Table 3.10 Composition of Sample by Educational Level in Paraguay, Selected Years**

	1992	1997/98	2005
Primary incomplete	44.43	36.86	27.46
Primary complete	23.18	25.32	23.05
Secondary incomplete	12.27	22.64	26.93
Secondary complete	13.18	8.36	11.00
College incomplete	4.60	5.40	9.42
College complete	2.35	1.42	2.14

Note: Data cover individuals between 12 and 64 with valid answers.

Source: Author's calculations based on EPH and EIH of National University and DGEEC.

Education is usually viewed as an equalizing force. The traditional argument points out that income disparities in one generation can be reduced in the next, if poor children have access to more and better education, so that the educational gap with rich-families' children decreases. However, following Kuznets (1955), one can tell a different story if the high-educated rich are a minority and only some poor children manage to make it all the way up to the highest educational (and income) levels. In that case, it is likely that inequality grows as the average education of the population increases; at least until the high education group is relatively large. With multiple educational levels, a similar unequalizing outcome emerges if there is a net outflow from the lowest educational levels and a similar net inflow to the highest levels, with minor changes in the intermediate

levels. Changes in the educational structure from 1997/98 to 2005 have more or less taken this form, which feeds the assumption of an unequalizing education effect for this period and a more equalizing effect for the 1992 to 1997/98 period. Between 1992 y 1997/98, 5% of working age population left the primary education group. Almost all of these entered the secondary education group. In the 1997/98 to 2005 period, 12% left primary education. Seven percent entered in secondary education but the other 5% passed on to college level.

So far we have analyzed several factors that might have affected inequality. Although we have offered some evidence to argue about each effect, we still do not have a consistent framework to confirm the sign of each effect and where to assess its quantitative relevance. Were changes in the returns to education really an unequalizing force? Were they really a significant force? What about gender, employment or education effects? The next section presents a framework to tackle these issues.

### 3.4 Estimation strategy

To compute expressions (3.14) and (3.19) in section two, we need to have estimates of parameters  $\beta$  and  $\lambda$  and the residual terms  $e$ . Also, since we do not have panels, we need a mechanism to assign observable and unobservable individual characteristics in period  $t'$  to individuals in  $t$ . This section is dedicated to explain the strategies to deal with these problems.

#### *Estimation of $\beta$ and $\lambda$*

Let  $L_i$  denote the number of hours worked by person  $i$ , and with  $w_i$  the hourly wage perceived. Total labour income is given by  $Y_i = L_i w_i$ . The number of hours of work  $L_i$  comes from a utility maximization process which determines optimal participation in the labour market, whereas wages are determined by market forces. The estimation stage specifies models for wages and hours of work which are used in the simulation stage described above.

The econometric specification of the model is similar to the one used by Bourguignon Fournier and Gurgand (2001), which corresponds to the reduced form of the labour decisions model originally proposed by Heckman (1974). Heckman shows how it is possible to derive an estimable reduced form starting from a structural system obtained from a utility maximization problem of labour - consumption decisions. Leaving technical details aside, the scheme proposed by Heckman has the following structure. Individuals allocate hours to work and domestic activities (or leisure) so as to maximize their utility subject to time, wealth, wages and other constraints. As usual, the solution to this optimization problem can be characterized as demand relations for goods and leisure as func-

tions of the relevant prices. Under general conditions it is possible to invert these functions to obtain prices and wages as functions of quantities of goods and leisure consumed (or its counterpart, hours of work). In particular, the wages obtained in this manner (denoted as  $w^*$ ) are to be interpreted as marginal valuations of labour, which will be a function of hours of work and other personal characteristics, and represent the minimum wage for which the individual would accept to work a determined number of hours. In equilibrium, if the individual decides to work, the number of hours devoted to labour should equate their marginal value  $w^*$  with the wage effectively perceived. On the contrary, if the individual decides not to work, it is because this marginal value is greater than the wage offered, given the individual's personal characteristics.

This discussion suggests how to determine wages as of which individuals are willing to work. On the same note, it is possible to model market determinants of wages offered ( $w$ ) as a function of characteristics such as years of education, experience and age as a standard Mincer equation (Mincer, 1974). In equilibrium it is assumed that the number of hours of work adjusts to make  $w=w^*$ .

The demand-supply relations discussed so far are structural forms in the sense that they reflect relevant economic behaviour in which wages offered and asked depend on the number of hours of work, which equate in equilibrium. Under general conditions it is possible to derive a reduced form for the equilibrium relations, in which wages and hours of work are expressed as functions of the variables taken as exogenous. In this way, the model has two equations, one for wages ( $w^*$ ) and one for the number of hours of work ( $L^*$ ), both as function of factors taken as given which affect wages ( $X1$ ) and hours ( $X2$ ) which may or may not have elements in common. The error terms  $e1$  and  $e2$  will represent non-observable factors affecting the determination of endogenous variables. According to the characteristics of the problem, for a particular individual we observe positive values of  $w^*$  and  $L^*$  if and only if the individual actually works. If the person does not work, we only know that the wage offered is less than the salary asked. Consequently, the reduced form model for wages and hours of work is specified as:

$$w_i^* = X1_i b + e1_i \quad i = 1, \dots, N \quad (3.20)$$

$$L_i^* = X2_i l + e2_i \quad (3.21)$$

with

$$w_i = w_i^* \text{ if } L_i^* > 0$$

$$w_i = 0 \text{ if } L_i^* \leq 0$$

$$L_i = L_i^* \text{ if } L_i^* > 0$$

$$L_i = 0 \text{ if } L_i^* \leq 0$$

where  $w_i$  and  $L_i$  correspond to observed wages and hours of work, respectively. This notation emphasizes that, consistently with the data used for the estimation, observed wages for a non-working individual are zero.

Following Heckman (1979), for estimation purposes we will assume that  $ei_1$  and  $ei_2$  have a bivariate normal distribution with  $E(e_{1i}) = E(e_{2i}) = 0$ , variances  $s_{12}$  and  $s_{22}$  and correlation coefficient  $r$ . This particular specification corresponds to the "Tobit type III" model in Amemiya's (1985) classification.

Although it is possible to estimate all the parameters using a full information maximum likelihood method, the implemented methodology adopted a limited information approach, which has notorious computational advantages. If instead of hours of work we only had information about whether the individual works or not, the model would correspond to the "Type II" model in Amemiya's classification, whose parameters can be estimated based on a simple selectivity model. More specifically, the regression equation would be the wage equation and the selection equation would be a censored version of the labour supply equation, simply indicating whether the individual works or not. Table 3.3 shows the estimation results of these equations.

On the other hand, the hours of work equation corresponds to the "Tobit type I" model in Amemiya's classification where the variable is observed only if it is positive. In this case, the parameters of interest could be estimated using a standard censored regression Tobit model. This strategy is consistent though not fully efficient. In any case, the efficiency loss is not necessarily significant for a small sample. The results of the estimation are shown in Table 3.7.

### *Unobservable Factors*

Unobservable characteristics affecting wages are modelled as regression error terms of the wage equation (3.20). Their mean is trivially normalized to zero and their variance is estimated as an extra parameter in the Heckman procedure. In order to simulate the effect of changes in those unobservable factors between  $t$  to  $t'$  on inequality, the estimated residuals of the wage equation of year  $t$  are rescaled by  $st'/st$ , where  $s$  is the estimated standard deviation of the wage equation. This captures the effect of differences between years in dispersion in the unobservable factor affecting wages, which include non-observable factors and their market value.<sup>35</sup>

35 It is important to mention that under bivariate normal assumption implicit in the Heckman model, once the correlation between unobservables affecting wages and hours worked is kept constant, all remaining effects on unobservables on wages come through the variance. Machado and Mata (1998) allow for heterogeneous behavior of the error term using quantile regression methods.

To study employment effects the decomposition methodology requires simulating earnings for people who do not work. Since we do not observe wages we cannot apply equations (3.20) and (3.21) to estimate the unobservable factors. For each individual in that situation, we assign as “error term” a random draw from the bivariate normal distribution implicit in the wage - labour supply model (3.20) and (3.21), whose parameters are consistently estimated by the Heckman procedure. Residuals are sampled from the distribution of unobservable factors but conditional to the fact that the behaviour of the individual is observed. That is, error terms are drawn from the bivariate normal distribution and a prediction (based on observable characteristics, estimated parameters and sampled errors) is computed for wages and hours worked. If the resulting prediction yields positive hours worked (so the prediction is inconsistent with the observed behaviour in this group), the error term is sampled again until non-positive hours of work are predicted.

#### *Individual characteristics*

For the estimation of the education effect it is necessary to simulate the educational structure of year  $t'$  on year  $t$  population since we do not have the same individuals in both years. Instead of following Bourguignon, Fournier and Gurgand (2001) and estimating a parametric equation that relates individual educational level to other individual characteristics (age and gender), a rough non-parametric mechanism was applied. Adult population was divided in ten homogeneous groups by gender and age and then the educational structure of a given cell in year  $t'$  was replicated into the corresponding cell in year  $t$ .

#### *Poverty*

Poverty, measured officially by income, summing all kind of income, decreased in Paraguay between 1992 and 1997/98, increased until 2002 and turned back to a slight reduction as from 2003. Poverty and inequality are not the same, but they are closely related. A higher level of inequality reduces poverty reduction driven by economic growth, since the additional income and benefits from growth are not equally distributed amongst the population. Inequality reduction is not a poverty reduction tool in itself, but it can improve performance and impact of poverty reduction processes. So, it would be interesting to figure out which could be the effects the simulated inequality changes on poverty levels. Since the simulations of changes of inequality are based on labour income, we need a labour income poverty line. There is no such line officially fixed for Paraguay. As a proxy, the mean equivalent household labour income of all households classified as poor by the official per capita income poverty line, was taken as an income poverty line to check on poverty reduction effects of simulated inequality changes. Since this is only a very rough proxy, measurement results should not be taken as real changes in poverty, related to inequality



changes, because there are other kinds of income and many more factors related to poverty change as a whole. Anyway, the simulated poverty changes could be understood as a proxy for the sign of poverty reduction impact of an observed change in inequality, and to identify which kind of inequality change would have a stronger effect on poverty and which would not.

### 3.5 Results

This section reports the results of performing the decomposition described in Section 2 using the estimation strategy outlined in Section 4. The objective is to shed light over the quantitative relevance of the various phenomena discussed in Section 3 on inequality changes during the 1992 - 2005 period.

Before showing the results two points must be clarified. First, the decompositions are path dependent. Hence, results are reported using alternatively  $t$  and  $t'$  as the base year. Second, the simulations are carried out for the whole distribution.

Tables 3.11 to 3.13 show the results both with  $t$  and  $t'$  as base years. Table 3.14 reports the average of these results. A positive number indicates an unequalizing effect. A large number compared to the other figures in the column suggests a significant effect. For instance, the returns-to-education effect on the individual earnings distribution in the 1992 to 1997/98 period is -2.1. This roughly means that the Gini would have decreased -2.1 points, if only the returns to education (*i.e.* the coefficients of the educational dummies in the wage equation) had changed between those years. The number -2.1 tells us two things: (i) since it is a negative number, it implies that the returns-to-education effect was inequality-decreasing, and (ii) since it is large compared to the other numbers in the column, it indicates that the change in returns to education was economically a significant factor affecting inequality. Its effect on equivalent household income distribution is also inequality decreasing, but to a minor degree. Nevertheless, returns to education seem to have a poverty increasing effect. The story here is that changes in return to education are related to income losses, which certainly make income distribution more equal. At the same time, however, lower income groups suffer stronger losses, so some of them fall below the poverty line.

**Table 3.11 Decomposition of the Change in Gini coefficient for Earnings and Equivalent Household Labour Income and Equivalent Labour Household Income Poverty, Paraguay 1992 – 1997/98**

<b>Using 1997/98 coefficients</b>						
<i>Indicator</i>	<b>Earnings</b>		<b>Equivalent income</b>		<b>Poverty</b>	
	Level	Change	Level	Change	Level	Change
1992 observed	58.9		60.6		35.4	
1997/98 observed	55.3	-3.6	58.4	-2.2	27.5	-7.9
<i>Effect</i>						
Returns to education	59.6	0.7	60.2	0.4	37.1	1.7
Gender wage gap	58.4	-0.5	59.5	-1.1	33.6	-1.8
Returns to experience	59.7	0.8	61.0	-0.4	34.6	-0.8
Unobservable factors	61.0	2.1	61.8	1.2	35.1	-0.3
Hours of work	57.6	-1.3	60.2	-0.4	33.9	-1.5
Employment	56.4	-2.5	59.5	-1.1	32.3	-3.1
Education	56.8	-2.1	59.0	-1.6	33.8	-1.6
Other factors	58.1	-0.8		0.8		-0.5
<b>Using 1992 coefficients</b>						
<i>Indicator</i>	<b>Earnings</b>		<b>Equivalent income</b>		<b>Poverty</b>	
	Level	Change	Level	Change	Level	Change
1992 observed	58.9		60.6		35.4	
1997/98 observed	55.3	-3.6	58.4	-2.2	27.5	-7.9
<i>Effect</i>						
Returns to education	59.0	0.1	60.7	0.1	36.2	0.8
Gender wage gap	58.6	-0.3	60.1	-0.5	34.4	-1.0
Returns to experience	58.4	-0.5	60.4	-0.2	34.4	-1.0
Unobservable factors	59.7	0.8	61.1	0.5	34.7	-0.7
Hours of work	58.0	-0.9	60.1	-0.5	33.6	-1.8
Employment	58.2	-0.7	60.3	-0.3	34.4	-1.0
Education	57.3	-1.6	59.7	-0.9	34.0	-1.4
Other factors	58.4	-0.5		-0.4		-1.8
<b>Average changes</b>						
<i>Indicator</i>	<b>Earnings</b>		<b>Equivalent income</b>		<b>Poverty</b>	
1992 - 1997/98 observed	-3.6		-2.2		-7.9	
<i>Effect</i>						
Returns to education	0.4		-0.1		1.3	
Gender wage gap	-0.4		-0.8		-1.4	
Returns to experience	0.2		0.1		-0.9	
Unobservable factors	1.5		0.8		-0.5	
Hours of work	-1.1		-0.4		-1.7	
Employment	-1.6		-0.7		-2.1	
Education	-1.9		-1.3		-1.5	
Other factors	-0.7		0.2		-1.2	

Note: The earnings distribution includes those individuals with  $Y_{it} > 0$  and  $Y_{it}(kt') > 0$ . The equivalent household labour income distribution includes those individuals with  $Y_{it} \geq 0$  and  $Y_{it}(kt') \geq 0$ . Non-labour income is not considered.

Source: Author's calculations based on EPH and EIH of National University and DGEEC.



**Table 3.12 Decomposition of the Change in Gini coefficient for Earnings and Equivalent Household Labour Income and Equivalent Labour Household Income Poverty, Paraguay 1997/98 – 2005**

<b>Using 2005 coefficients</b>						
<i>Indicator</i>	Earnings		Equivalent income		Poverty	
	Level	Change	Level	Change	Level	Change
1997/98 observed	55.3		58.4		27.5	
2005 observed	49.0	-6.3	52.8	-5.6	32.6	5.1
<i>Effect</i>						
Returns to education	52.6	-2.7	57.2	-1.2	29.3	1.8
Gender wage gap	55.2	-0.1	58.1	-0.3	27.1	-0.4
Returns to experience	53.9	-1.4	57.6	-0.8	27.7	0.2
Unobservable factors	53.2	-2.1	56.6	-1.8	28.7	1.2
Hours of work	54.3	-1.0	58.0	-0.4	27.8	0.3
Employment	55.4	0.1	57.7	-0.7	28.1	0.6
Education	54.8	-0.5	58.4	0.0	27.3	-0.2
Other factors		1.4		-0.4		1.6
<b>Using 1997/98 coefficients</b>						
<i>Indicator</i>	Earnings		Equivalent income		Poverty	
	Level	Change	Level	Change	Level	Change
1997/98 observed	55.3		58.4		27.5	
2005 observed	49.0	-6.3	52.8	-5.6	32.6	5.1
<i>Effect</i>						
Returns to education	53.1	-2.2	56.6	-1.8	28.9	1.4
Gender wage gap	55.0	-0.3	57.8	-0.6	27.3	-0.2
Returns to experience	54.8	-0.5	58.1	-0.3	28.5	1.0
Unobservable factors	53.9	-1.4	57.3	-1.1	28.5	1.0
Hours of work	53.4	-1.9	57.4	-1.0	28.2	0.7
Employment	55.6	0.3	58.1	-0.3	28.3	0.8
Education	55.5	0.2	58.8	0.4	28.2	0.7
Other factors		-0.5		-0.9		-0.3
<b>Average changes</b>						
<i>Indicator</i>	Earnings		Equivalent income		Poverty	
1997/98 – 2005 observed	-6.3		-5.6		5.1	
<i>Effect</i>						
Returns to education	-2.5		-1.5		1.6	
Gender wage gap	-0.2		-0.4		-0.3	
Returns to experience	-1.0		-0.5		0.6	
Unobservable factors	-1.8		-1.4		1.1	
Hours of work	-1.5		-0.7		0.5	
Employment	0.2		-0.5		0.7	
Education	-0.1		0.2		0.3	
Other factors	0.5		-0.7		0.7	

Note: The earnings distribution includes those individuals with  $Y_{it} > 0$  and  $Y_{it}(kt') > 0$ . The equivalent household labour income distribution includes those individuals with  $Y_{it} \geq 0$  and  $Y_{it}(kt') \geq 0$ . Non-labour income is not considered.

Source: Author's calculations based on EPH and EIH of National University and DGECC.

**Table 3.13 Decomposition of the Change in Gini coefficient for Earnings and Equivalent Household Labour Income and Equivalent Labour Household Income Poverty, Paraguay 1992 – 2005**

<b>Using 2005 coefficients</b>						
<i>Indicator</i>	Earnings		Equivalent income		Poverty	
	Level	Change	Level	Change	Level	Change
1992 observed	58.9		60.6		35.4	
2005 observed	49.0	-9.9	52.8	-7.8	32.6	-2.8
<i>Effect</i>						
Returns to education	56.9	-2.0	59.0	-1.6	38.9	3.5
Gender wage gap	58.3	-0.6	59.2	-1.4	33.2	-2.2
Returns to experience	58.3	-0.6	60.2	-0.4	34.8	-0.6
Unobservable factors	58.9	0.0	60.0	-0.6	36.3	0.9
Hours of work	56.6	-2.3	59.8	-0.8	34.2	-1.2
Employment	56.5	-2.4	58.8	-1.8	32.9	-2.5
Education	56.3	-2.6	59.0	-1.6	33.6	-1.8
Other factors		0.6		0.4		1.1
<b>Using 1992 coefficients</b>						
<i>Indicator</i>	Earnings		Equivalent income		Poverty	
	Level	Change	Level	Change	Level	Change
1992 observed	58.9		60.6		35.4	
2005 observed	49.0	-9.9	52.8	-7.8	32.6	-2.8
<i>Effect</i>						
Returns to education	56.8	-2.1	58.9	-1.7	37.6	2.2
Gender wage gap	58.3	-0.6	59.5	-1.1	34.2	-1.2
Returns to experience	57.9	-1.0	60.1	-0.5	35.4	0.0
Unobservable factors	58.3	-0.6	60.0	-0.6	35.7	0.3
Hours of work	56.1	-2.8	59.1	-1.5	34.3	-1.1
Employment	58.5	-0.4	60.0	-0.6	35.2	-0.2
Education	57.5	-1.4	60.1	-0.5	34.7	-0.7
Other factors		-1.0		-1.3		-2.1
<b>Average changes</b>						
<i>Indicator</i>	Earnings		Equivalent income		Poverty	
1992 - 2005 observed	-9.9		-7.8		-2.8	
<i>Effect</i>						
Returns to education	-2.1		-1.7		2.9	
Gender wage gap	-0.6		-1.3		-1.7	
Returns to experience	-0.8		-0.5		-0.3	
Unobservable factors	-0.3		-0.6		0.6	
Hours of work	-2.6		-1.2		-1.2	
Employment	-1.4		-1.2		-1.4	
Education	-2.0		-1.1		-1.2	
Other factors	-0.2		-0.5		-0.5	

Note: The earnings distribution includes those individuals with  $Y_{it} > 0$  and  $Y_{it}(kt') > 0$ . The equivalent household labour income distribution includes those individuals with  $Y_{it} \geq 0$  and  $Y_{it}(kt') \geq 0$ . Non-labour income is not considered.

Source: Author's calculations based on EPH and EIH of National University and DGEEC.

**Table 3.14 Decomposition of the Change in the Gini Coefficient and Equivalent Household Income Poverty Rates Changing the Base Year, Paraguay, Selected Periods**

	Earnings			Equivalent household income			Poverty		
	1992 - 1997/98	1997/98 - 2005	1992 - 2005	1992 - 1997/98	1997/98 - 2005	1992 - 2005	1992 - 1997/98	1997/98 - 2005	1992 - 2005
Observed	-3.6	-6.3	-9.9	-2.2	-5.6	-7.8	-7.9	5.1	-2.8
Effect									
Returns to education	0.4	-2.5	-2.1	-0.1	-1.5	-1.7	1.3	1.6	2.9
Gender wage gap	-0.4	-0.2	-0.6	-0.8	-0.4	-1.2	-1.4	-0.3	-1.7
Returns to experience	0.2	-0.9	-0.7	0.1	-0.5	-0.4	-0.9	0.6	-0.3
Unobservable factors	1.5	-1.8	-0.3	0.8	-1.4	-0.6	-0.5	1.1	0.6
Hours of work	-1.1	-1.5	-2.6	-0.4	-0.7	-1.2	-1.7	0.5	-1.2
Employment	-1.5	0.2	-1.3	-0.7	-0.5	-1.2	-2.1	0.7	-1.4
Education	-1.9	-0.1	-2.0	-1.3	0.2	-1.1	-1.5	0.3	-1.2
Other factors	-0.8	0.5	-0.3	0.3	-0.8	-0.5	-1.1	0.6	-0.5

Note: The earnings distribution includes those individuals with  $Y_{it} > 0$  and  $Y_{it}(kt') > 0$ . The equivalent household labour income distribution includes those individuals with  $Y_{it} \geq 0$  and  $Y_{it}(kt') \geq 0$ . Non-labour income is not considered.

Source: Author's calculations based on EPH and EIH of National University and DGEEC.

### *Returns to education*

Table 3.14 confirms the assumption of Section 3. Changes in the returns to education had an unequalizing effect on the individual earnings distribution between 1992 and 1997/98 and a strong equalizing effect in the next seven years. The effects on the equivalent income distribution were similar. Over the whole period 1992-2005, changes in the returns to education (in terms of hourly wages) represented an important inequality-decreasing factor.

### *Gender wage gap*

As expected, changes in the gender parameter of the wage equation implied an equalizing effect on the individual earnings distribution. During the last decade the gender gap has substantially reduced in size. Given that women earn less than men, that movement had an unambiguous inequality - decreasing effect on the earnings distribution. It is interesting to notice that the gender effect becomes more important in the equivalent household labour income distribution. Two factors combine to generate this result. First, female workers are more concentrated in the lower part of the distribution than men (mainly in rural area) and therefore a relative wage change implies a decrease in household income inequality. Second, a proportional wage increase for all females is more relevant in low-income families since women's earnings are a more significant part of the

total resources in those households than in rich families. A great example is the disproportionate number of poor households headed by working women. Consequently, its effect on poverty is a reduction.

### *Returns to experience*

Changes in the returns to experience (age) implied an unequalizing effect on the earnings distribution during the period 1992-1997/98 and an equalizing effect in the next seven years. A brief explanation follows. The effects of changes in the returns to experience did not have direction and impact for all education levels. There was a clear equalizing effect for low education levels, unclear effects for medium education levels and an unequalizing effect for high educational levels. The empirical evidence shows that the increase of inequality in higher education levels outpaced equalizing effects for lower levels. Nevertheless, the equalizing effects for lower education levels do exist and are associated to higher incomes, so that poverty tends to reduce. Between 1997/98 and 2005 there is a more generalized equalizing effect, however, associated with income losses. So inequality decreased meanwhile poverty increased. Results for equivalent household income show the same patterns.

### *Unobservable Factors*

Changes in endowments and returns to unobservable factors have implied unequalizing changes in wages in the 1992 – 1997/98 period, associated with poverty reduction and opposite effects over the next seven years, for both, earnings and equivalent household income.

### *Hours of Work*

We carried out three simulations to assess the relevance of employment changes on inequality. In all of these the distribution in the base year is simulated using the parameters of the Tobit employment equation of the other year. In the *employment* and *participation* effects, people with non-positive simulated hours of work are assigned zero earnings, so they remain included in the data set. People who work in the simulation are assigned the actual base year wage and the simulated worked hours in the *employment* effect and the actual worked hours in the *participation* effect. The third simulation is intended to single out the impact of changes in hours worked. People who change labour status (*i.e.* we kept their current earnings) and change hours of work to individuals who work both in the base year and in the simulation, were ignored.

An equalizing employment effect shows up in the individual earnings and equivalent household income distribution for the whole period. Nevertheless, in the first period it is associated with poverty reduction (income increase for lower

earnings) and with a poverty increase in the second period. Notice that since we exclude those individuals with zero earnings from that distribution, the employment effect is basically the result of relative changes in the number of hours of work. The figures for the hours-of-work and participation effects confirm this assertion. As discussed in Section 3, the nineties witnessed a substantial increase in hours of work in general.

### *Employment*

Labour participation grew fast after 1992, a period of economic growth and creation of new and additional labour opportunities. Consequently, inequality and poverty decreased during the first period, for earnings and equivalent household income. Part of these gains were lost after 1998, when the economy entered a period of recession. Middle classes rank first regarding income losses, so inequality and poverty increased after 1998, for both, earnings and equivalent household income. Nevertheless, losses of the second period were not as strong as the gains of the first, so the overall effect was an inequality and poverty reduction.

### *Education*

Paraguay has witnessed important changes in the educational composition of its population since the implementation of educational reform was started in 1994. An inequality and poverty reduction for earnings and households, together with poverty reduction was the result for the first period. In the second period earnings distribution keeps getting better for labour income, but equivalent household income distribution unequalizes and poverty increases. This might be the result of the increase in workers with university level education at the same time as income losses for primary education.

### *Other factors and Interactions*

The last row in Table 3.14 is calculated as a residual. It encompasses the effects of interaction terms and many factors not considered in the analysis. According to Table 3.14, in general, this term is lower than the mean of the other terms in the decomposition, implying either that the factors not considered in the analysis are not extremely important or that they tend to compensate each other.

## **3.6 Discussion**

The results of the paper suggest that the smaller change in inequality between 1992 and 1997/98 is mainly the result of employment (including hours of work) and education effects, characterized by a primary schooling expansion. The stronger inequality reduction effect after 1997 is due to returns to education, hours of work (since unemployment increased) and unobservable factors. Maybe

the most interesting finding of the paper is, that the general trend of an inequality decrease (interrupted by the 2000 – 2002 economic crises) held over the observation period, finding a way to reduce inequality even during periods of poverty increase. Comparing the post-2002 period with 1997/98, we can see that in 2005, even if poverty was higher, inequality was lower. As shown above, labour market conditions in a mix-up of participation rates, unemployment, hours of work and returns to education are the mechanisms which helped to decrease inequality, as well as unobservable factors in the 1997/98 to 2005 period, however, poverty increased at the same time. Labour income in 2005, in general, was lower than in 1992 and income was lost over the whole distribution and, in a higher level for higher income groups, this is why inequality decreased. However, since income also decreased for the poor, some former non-poor workers of households fell below the poverty line, and are now what are known as “new-poor” households. Good inequality reduction policies should search the opposite output, inequality and poverty reduction at the same time.

One of the surprising findings of this paper is the extremely high returns to education in 1997/98. Education reform started in 1994 with primary education, so education reform results could not yet have had impact labour market in 1997/98. However, the decrease of returns to education after 1998 can be observed in relation to education reform, at least for secondary education. As labour force increases its human capital at a massive rate, returns to education tend to decrease. We checked for returns to education in 1999 and 2000/01 surveys. In both cases, returns to education are surprisingly high, although slightly lower than 1997/98 results. Consequently, there seems to be no measurement error. Returns to education fall sharply in 2002, just at the time of a deepening in the economic crisis. Thus, decreasing returns to education seem to be a mix of lower remuneration levels in all the economy and the results of education reform.

Changes of inequality at the equivalent household income level are difficult to understand. Nevertheless, they are included in this paper just to show that even if inequality changes related, for instance to labour participation, could be important at an individual level, their impact at the household level does not necessarily have to be the same. Interestingly, gender wage gap reductions have a strong poverty reduction impact at a household level. The chain of effects seems to be that the additional income for women, leaving everything else constant, also benefits female headed households, most of which live below the poverty line.

The same factor which explained inequality changes, employment, hours of work and education factors have the main impact on changes on income poverty



levels, as should be expected. Once more, the signs and the “rank” of these poverty estimates should be considered and not necessarily the magnitude of simulated changes in poverty, since their estimation method was not very sophisticated.

### 3.7 Conclusions

The decomposition methodology used in this paper can describe more completely the reasons for changes in aggregate income inequality within particular economies. A country may experience relatively little change in the overall level of income inequality despite significant changes in the composition of that inequality. Analyzing several countries during the same period of time, using this kind of inequality decomposition will produce more detailed results than cross country comparisons and might show how despite similar economic crisis and common trends in a given region and period of time, overall levels of and changes in income inequality remain distinct by country (Bourguignon, Ferreira, Lustig, 2005). Dion (2007) concludes from their comparison of several countries show that it seems likely that differences in inequality outcomes may reflect differences not only in endowments, prices and occupation effects, but also in policy decisions and priorities of different governments.

This paper contributes to an upcoming political discussion in Paraguayan development politics, which are starting to shift their focus somewhat away from poverty reduction politics towards inequality reduction politics, now understanding poverty in part, as a consequence of inequality. This contribution is appreciated by showing the results of a microeconometric decompositions methodology. This technique allows the assessment of the relevance of various factors that affected inequality during a period of 13 years, between 1992 and 2005.

This paper is not on Paraguayan poverty or inequality reduction politics. Nevertheless, some concluding comments on these can help to better understand the acquired results. The story we can tell, knowing Paraguayan politics, is that the impacts of market forces and business cycles have a stronger impact on Paraguayan inequality and poverty, than special policies do. This is, on the one hand, because there are very few of such policies and, on the other hand, most of these lack scale, so even if the political concept is adequate, impact cannot be created for problems of scale. Educational politics is one of the exceptions.

There are also structural problems in Paraguayan economy, with an informal sector of about 50% of the labour force, so any initiative taken by the government, for instance on legal minimum wage, will not have any impact on half of the labour force. These kinds of problems are strongly limiting possibilities for

policies impact; just an example to better understand the importance of these kinds of phenomenon. In 2005, only 10% of the labour force had a labour income equal or above the legal minimum wage.

Consequently, if regional market forces and business cycles tell almost the whole story of inequality and poverty changes, we should better understand how this works. Labour income in Paraguay is much more than monetary income. It includes monetarized values for self consumption of agricultural products cultivated by farmers. In 1997, agricultural GDP growth was 2.2 times bigger than overall GDP growth. More than 35% of the labour force works in the agricultural sector. Cultivating land is almost the same amount of work (in hours) year after year, but if the harvest is good and prices are even better, for a small period of years, the returns to education (even for low educated small farmers) will be high for these years. Returns to education level benefit from an open economy in “good times”, but in “bad times” external shocks such as the Brazilian devaluation and the Argentine default strike even harder.

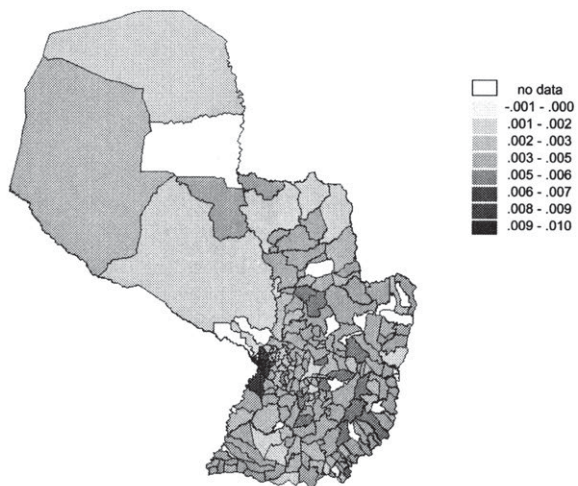
Thus, if social politics are necessary to reduce inequality, but economic and market forces are stronger in their negative impacts than positive impacts that could be generated by social politics, maybe protection mechanisms for vulnerable groups would be the necessary complement to social politics and research should focus on these issues.





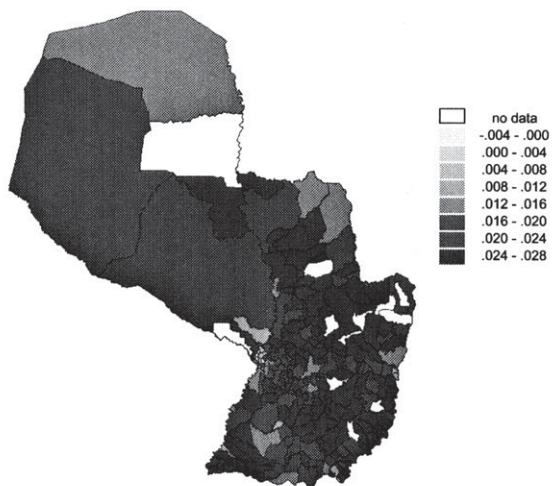
## A - Annex to Chapter 1

**Figure A1** Structured error per capita income estimates 1992  
at district level



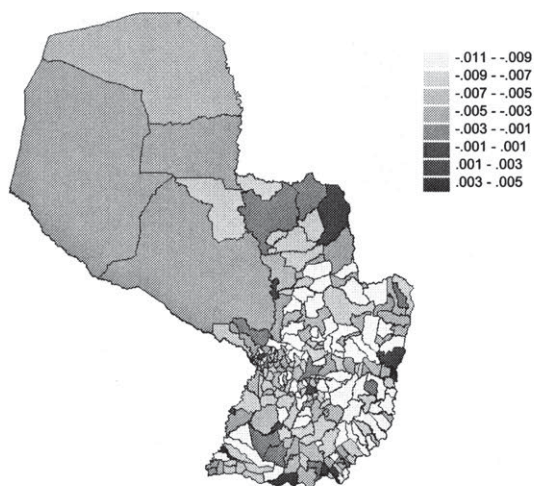
Source: Author's calculations based on ECV 1992 and CNPV 1992

**Figure A2** Unstructured error per capita income estimates 1992  
at district level



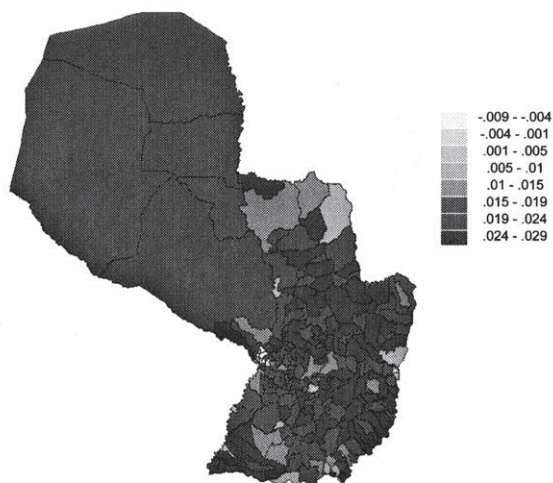
Source: Author's calculations based on ECV 1992 and CNPV 1992

**Figure A3 Structured error per capita income estimates 2002 at district level**



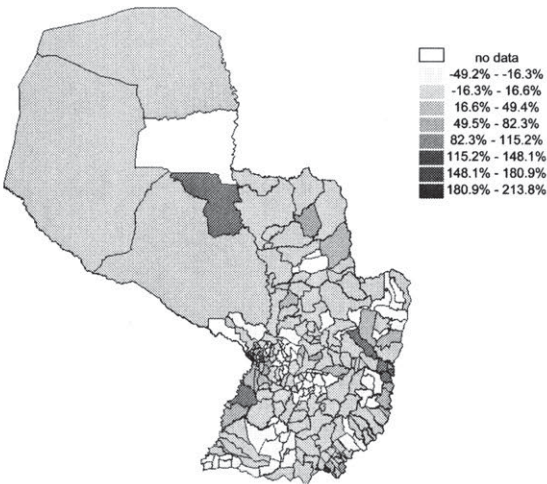
Source: Author's calculations based on EPH 2002 and CNPV 2002

**Figure A4 Unstructured error per capita income estimates 2002 at district level**



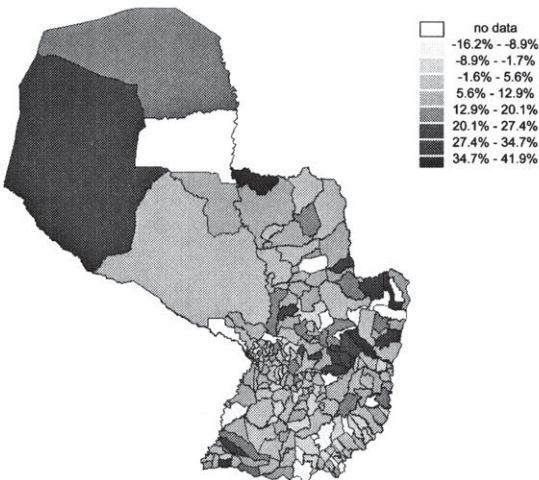
Source: Author's calculations based on EPH 2002 and CNPV 2002

**Figure A5** Relative change in FGT0 per capita income – period 1992 – 2002 at district level



Source: Author’s calculations based on ECV 1992, CNPV 1992, EPH 2002 and CNPV 2002

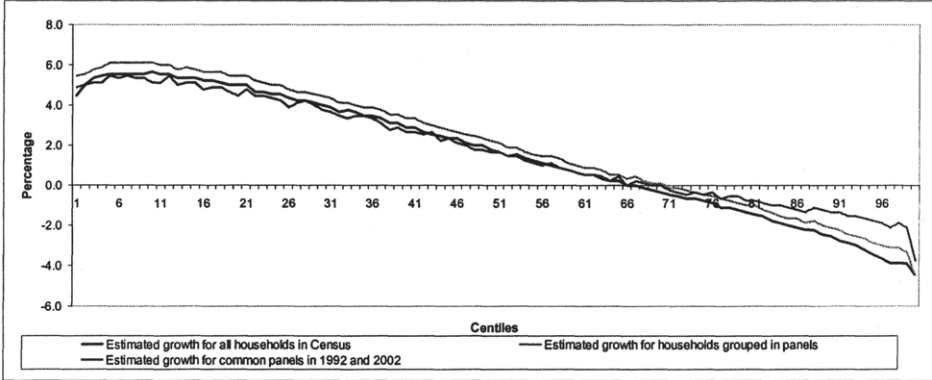
**Figure A6** Relative change in Gini per capita income – period 1992 – 2002 at district level



Source: Author’s calculations based on ECV 1992, CNPV 1992, EPH 2002 and CNPV 2002

**B – Annex to Chapter 2**

**Figure A.1 Growth Incidence Curves Paraguay 1992 – 2002**



Source: Author's calculations based on results in Chapter 1

## Bibliography

- Aghion, P., Caroli, E. and Garcia-Penalosa, C. (1999). Inequality and economic growth: The perspective of the new growth theories, *Journal of Economic Literature*, **37** (4), 1615-60.
- Alderman, H., M. Babita, G. Demombynes, N. Makhatha, and B. Ozler (2002) How Low Can You Go?: Combining Census and Survey Data for Mapping Poverty in South Africa, *Journal of African Economics*, vol. **11**, pp. 169-200.
- Altimir, O., Beccaria, L. and González Rozada, M. (2000). La evolución de la distribución del ingreso familiar en la Argentina. Documento de trabajo de la Maestría en Finanzas Públicas Provinciales y Municipales, UNLP.
- Amemiya, T. (1985). *Advanced Econometrics*, Harvard University Press.
- Andersen, Lykke E. (2001). Social Mobility in Latin America. Institute for Socio-Economic Research, Universidad Católica Boliviana. Working Paper No. 03/2000.
- Atkinson, A. B. and Brandolini, A. (2001). Promise and pitfalls in the use of "secondary" data-sets: Income inequality in OECD countries as a case study, *Journal of Economic Literature*, **39** (3), 771-799.
- Atkinson, A.B., F. Bourguignon, and C. Morrisson, (1992). Empirical Studies of Earnings Mobility. (Chur, Switzerland: Harwood).
- Balisacan, A. M. and Fuwa, N. (2003). Growth, inequality and politics revisited: A developing-country case, *Economics Letters*, **79** 53-58.
- Bigman, D. and Fofack, H. (2000). Geographical Targeting for Poverty Alleviation: An Introduction to the Special Issue, *The World Bank Economic Review*, **14**(1), pp. 129-145.
- Birdsall, N. and Londono, J. L. (1997). Asset inequality matters: An assessment of the World Bank's approach to poverty reduction, *American Economic Review, Papers and Proceedings*, **87** (2), 32-37.
- Bouillon, C., Legovini, A. and Lustig, N. (1998). Rising inequality in Mexico: returns to household characteristics and the "Chiapas effect". Mimeo.
- Bourguignon, F., Ferreira, F. and Lustig, N.(2005). The Microeconomics of Income Distribution Dynamics in East Asia and Latin America, Worldbank, Washington.
- Bourguignon, F. and Ferreira, F. (2005). Decomposing Changes in the Distribution of Household Incomes: Methodological Aspects, in: Bourguignon, F., Ferreira, F. and Lustig, N.(2005). The Microeconomics of Income Distribution Dynamics in East Asia and Latin America, Worldbank, Washington, p. 17-46.

- Bourguignon, F., Fournier, M., and Gurgand, M. (1999). Fast development with an stable income distribution: Taiwan, 1979-1994. Mimeo.
- Buhmann, B., Rainwater, G. Schmaus, G. y Smeeding, T. (1988). Equivalence scales, well being, inequality and poverty: sensitivity estimates across ten countries using the Luxembourg Income Study database. *Review of Income and Wealth* 34, 115-142.
- Brock, W. A. and Durlauf, S. N. (2001). Growth empirics and reality, *World Bank Economic Review*, 15 (2), 229-272.
- Caselli, F., Esquivel, G. and Lefort, F. (1996). Reopening the convergence debate: A new look at cross-country growth empirics, *Journal of Economic Growth*, 1 (3), 363-390.
- Castello, A. and Domenech, R. (2002), Human capital inequality and economic growth: Some new evidence, *Economic Journal*, 112 (478), C187-C200.
- Deininger, K. and Okidi, J. (2003). Growth and poverty reduction in Uganda, 1999-2000: Panel data evidence, *Development Policy Review*, 21 (4), 481-509.
- Deininger, K. and Squire, L. (1996), A new data set measuring income inequality, *World Bank Economic Review*, 10 (3), 565-91.
- Demombynes, G., Elbers, C., Lanjouw, J., Lanjouw, P., Mistiaen, J. and Özler, B. (2002). Producing an Improved Geographic Profile of Poverty: Methodology and Evidence from Three Developing Countries. World Institute for Development Economics Research, United Nations University, discussion paper no. 2002/39. [
- Dion, M., (2007). Poverty, inequality, policy and politics in Latin America. *Latin American Research Review*, 42/1, 186-195.
- Elbers, C., Lanjouw, J. O. and Lanjouw, P. (2005), Imputed welfare estimates in regression analysis, *Journal of Economic Geography*, forthcoming.
- Elbers, C. and Gunning, J. W. (2004). "Transitional growth and income inequality: Anything goes", mimeo, Vrije Universiteit, Amsterdam.
- Elbers, C., Lanjouw, J. O. and Lanjouw, P. (2003). Micro-level estimation of poverty and inequality, *Econometrica*, 71 (1), 355-364.
- Elbers, C., Lanjouw, J. O., Lanjouw P. and Leite P. G. (2004) Poverty and Inequality in Brazil: New Estimates from Combined PPV-PNAD Data, Washington.
- Ferreira, Francisco y Paes de Barros, Ricardo (1999). Climbing a moving mountain: explaining the decline in income inequality in Brazil from 1976 to 1996. Mimeo.
- Fields, G., Duval, R., Feije, S. and Sanchez, M. (2006) Inter-generational Income Mobility in Latin America, Cornell University, processed.
- Forbes, K. J. (2000). A reassessment of the relationship between inequality and growth, *American Economic Review*, 90 (4), 869-887.

- Foster, J., Greer, J., and Thorbecke, E. (1984). A Class of Decomposable Poverty Measures, *Econometrica*, **52**, pp. 761-66.
- Galor, O. (2000). Income distribution and the process of development, *European Economic Review*, **44** 706-712.
- Ganuza, E., Morley, S., Robinson, S. and Vos, R. (Editors) (2004). ¿Quién se beneficia del libre comercio? Promoción de exportaciones y pobreza en América Latina y el Caribe en los 90, Bogotá, Alfaomega Colombiana.
- Ganuza, E., Paes de Barros, R., Taylor, L. and Vos, R. (Editors) (2001). Liberalización, desigualdad y pobreza: América Latina en los 90, Buenos Aires, Editorial Universitaria EUDEBA.
- Ganuza, E., Taylor, L. and Morley, S., (Editors) (1998). Política macroeconómica y pobreza en América Latina y el Caribe, Madrid, Mundi Prensa Libros.
- Gasparini, L., Marchionni, M. and Sosa, W. (2005). Characterization of Inequality Changes through Microeconomic Decompositions: The case of Greater Buenos Aires, in: Bourguignon, F., Ferreira, F. and Lustig, N.(2005). The Microeconomics of Income Distribution Dynamics in East Asia and Latin America, Worldbank, Washington, p. 47-82.
- Gibson, B., Molinas, J. and Molinas, M. (2001). Liberalización del Sector Externo en Paraguay. Efectos sobre el crecimiento, la distribución del ingreso y la pobreza, in: Ganuza, E., Paes de Barros, R., Taylor, L. and Vos, R. (Editors) (2001), Liberalización, desigualdad y pobreza: América Latina en los 90, Buenos Aires, Editorial Universitaria EUDEBA, p. 815-876.
- Ghosh, M. and Rao, J.N.K. (1994). Small Area Estimation: An Appraisal. *Statistical Science*. **9**(1): 55-93
- Hassler, John & José V. Rodríguez Mora (1998). IQ, Social Mobility and Growth. Seminar Papers No. 635, Institute for International Economic Studies, Stockholm University. January.
- Heckman, J. (1974). Shadow Prices, Market Wages, and Labour Supply. *Econometrica*. **42**, 4.
- Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*. **47**, pp. 153-161.
- Henninger, N. (1998). Mapping and Geographic Analysis of Human Welfare and Poverty – Review and Assessment, World Resources Institute, Washington, D.C.
- Hentschel, J., Lanjouw, J., Lanjouw, P. and Poggi, J. (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), pp. 147-165.
- Jarvis, S. and Jenkins S.P. (1996). Changing Places: Income Mobility and Poverty Dynamics in Britain. University of Essex Paper Number 96-19, processed.



- Joeng, H. (2001) An Assessment of Relationship Between Growth and Inequality Using Micro Data from Thailand. University of Southern California.
- Juhn, C, Murphy, K. y Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy* 101 (3), 410-442.
- Klasen, S. (2004). In Search of the Holy Grail: How to Achieve Pro Poor Growth?, in Tungodden, B., N. Stern, and I. Kolstad (eds). *Toward Pro Poor Policies-Aid, Institutions, and Globalization*. New York: Oxford University Press. Reprinted in Krakowski, M. (ed.) *Attacking Poverty: What makes growth pro-poor?* Baden-Baden: Nomos (2004).
- Krugman, P. (1992). The Rich, the Right, and the Facts. *The American Prospect*. 11 19-31.
- Kuznets, S. (1955). Economic growth and income inequality. *American Economic Review*, 45, 1-28.
- Lee, H. (1999). Article on income distribution in Argentina in the Argentina Poverty Assessment. The World Bank.
- Legovini, A., Bouillon, C. and Lustig, N. (1998). Can Education Explain Changes in Income Inequality in Mexico?, in Bourguignon, F., Ferreira, F. and Lustig, N. (2005) *The Microeconomics of Income Distribution Dynamics in East Asia and Latin America*, World Bank, Washington, pp. 275-312.
- Llach, J. y Montoya, S. (1999). *En pos de la equidad. Pobreza y distribución del ingreso en la Argentina*.
- Lundberg, M. and Squire, L. (2003). The simultaneous evolution of growth and inequality, *Economic Journal*, 113 (April), 326-344.
- Machado, J. y Mata, J. (1998). Sources of increased inequality. Mimeo. Universidade Nova de Lisboa.
- Mincer, J. (1974). *Schooling, Experience and Earnings*. New York, Columbia University Press for NBER.
- Molinas, J., Cabello, C. and Otter, Th. (2004). Paraguay: Crecimiento, exportaciones y reducción de pobreza y desigualdad, in: Ganuza, E., Morley, S., Robinson, S. and Vos, R. (Editors) (2004), *¿Quién se beneficia del libre comercio? Promoción de exportaciones y pobreza en América Latina y el Caribe en los 90*, Bogotá, Alfaomega Colombiana, p. 423-452.
- Morley, S. and Vos, R. (1998). Pobreza y crecimiento dual en Paraguay, in: Ganuza, E., Taylor, L. and Morley, S., (Editors) (1998), *Política macroeconómica y pobreza en América Latina y el Caribe*, Madrid, Mundi Prensa Libros, p. 669-712.
- Otter, Th., (2007). Migración interna, pobreza y capital social en Paraguay entre 1985 y 2002, *Cuadernos de Economía y Administración* 1(1), pp 109-158.

- Pradhan, M., Suryahadi, A., Sumarto, S. and Pritchett, L. (2001). Eating Like Which 'Joneses'? An Iterative Solution to the Choice of Poverty Line Reference Group, *The Review of Income and Wealth*, **47**(4), pp. 473-487.
- Pritchett, L., Sumarto, S. and Suryahadi, A. (2000). Quantifying Vulnerability to Poverty: A Proposed Measure, Applied to Indonesia, *World Bank Policy Research Working Paper No. 2437*, The World Bank, Washington, DC.
- Raut, Lakshmi K. (1996). Signalling equilibrium, Intergenerational mobility and long-run growth. University of Hawaii-Manoa. Draft. February.
- Ravallion, M. (1997). Can high-inequality developing countries escape absolute poverty? *Economics Letters*, **56** 51-57.
- Ravallion, M. (1998). Does aggregation hide the harmful effects of inequality on growth? *Economics Letters*, **61** 73-77.
- Rao, J. N. K. (1999). Some Recent Advances in Model-Based Small Area Estimation, *Survey Methodology*, **25**, pp. 175-86.
- Robles, M. (2000). Indicadores para la focalizacion, DGEEC, Fernando de la Mora.
- Robles, M. and Santander, H. (2004). Mapas de Pobreza para Paraguay 2004, DGEEC, Fernando de la Mora, mimeo.
- Schipper, Y. and Hoogeveen, J. (2004). Does Inequality harm growth? An assessment on the growth impact of expenditure and education inequality in Uganda, mimeo.
- Shorrocks, A. F. (1978). Income Inequality and Income Mobility, *Journal of Economic Theory*, **19** 376-393.
- Scott, Christopher D. and Julie Litchfield (1994) Inequality, Mobility and the Determinants of Income Among the Rural Poor in Chile, 1968-1986, Development Economics Research Programme, STICERD, London School of Economics and Political Science, Discussion Paper No. 53.
- Thorbecke, E. and Charumilind, C. (2002). Economic inequality and its socio-economic impact, *World Development*, **30** (9), 1477-1495.
- van de Walle, D. (1998). Targeting Revisited, *The World Bank Research Observer*, **13**(2), pp. 231-248.

**Göttlinger Studien zur Entwicklungsökonomik  
Göttingen Studies in Development Economics**

Herausgegeben von/Edited by Hermann Sautter  
und/and Stephan Klasen

Die Bände 1-8 sind über die Vervuert Verlagsgesellschaft (Frankfurt/M.) zu beziehen.

- Bd./Vol. 9 Hermann Sautter / Rolf Schinke (eds.): Social Justice in a Market Economy. 2001.
- Bd./Vol.10 Philipp Albert Theodor Kircher: Poverty Reduction Strategies. A comparative study applied to empirical research. 2002.
- Bd./Vol.11 Matthias Blum: Weltmarktintegration, Wachstum und Innovationsverhalten in Schwellenländern. Eine theoretische Diskussion mit einer Fallstudie über „Argentinien 1990-1999“. 2003.
- Bd./Vol.12 Jan Müller-Schaeßel: Die Privatisierung und Regulierung des Wassersektors. Das Beispiel Buenos Aires/Argentinien. 2003.
- Bd./Vol.13 Ludger J. Löning: Economic Growth, Biodiversity Conservation, and the Formation of Human Capital in a Developing Country. 2004.
- Bd./Vol.14 Silke Woltermann: Transitions in Segmented Labor Markets. The Case of Brazil. 2004.
- Bd./Vol.15 Jörg Stosberg: Political Risk and the Institutional Environment for Foreign Direct Investment in Latin America. An Empirical Analysis with a Case Study on Mexico. 2005.
- Bd./Vol.16 Derk Bienen: Die politische Ökonomie von Arbeitsmarktreformen in Argentinien. 2005.
- Bd./Vol.17 Dierk Herzer: Exportdiversifizierung und Wirtschaftswachstum. Das Fallbeispiel Chile. 2006.
- Bd./Vol.18 Jann Lay: Poverty and Distributional Impact of Economic Policies and External Shocks. Three Case Studies from Latin America Combining Macro and Micro Approaches. 2007.
- Bd./Vol.19 Kenneth Hartgen: Empirical Analysis of Determinants, Distribution and Dynamics of Poverty. 2007.
- Bd./Vol. 20 Stephan Klasen / Felicitas Nowak-Lehmann: Poverty, Inequality and Migration in Latin America. 2008.
- Bd./Vol. 21 Isabel Günther: Empirical Analysis of Poverty Dynamics. With Case Studies from Sub-Saharan Africa. 2007.
- Bd./Vol. 22 Peter Dung: Malaysia und Indonesien: Wirtschaftliche Entwicklungsstrategien in zwei Vielvölkerstaaten. 2008.
- Bd./Vol. 23 Thomas Otter: Poverty, Income Growth and Inequality in Paraguay During the 1990s. Spatial Aspects, Growth Determinants and Inequality Decomposition. 2008.