Ibero-Amerika Institut für Wirtschaftsforschung
Instituto Ibero-Americano de Investigaciones Económicas
Ibero-America Institute for Economic Research
(IAI)

Georg-August-Universität Göttingen
(founded in 1737)

Diskussionsbeiträge · Documentos de Trabajo · Discussion Papers

Nr. 188

Does Inequality Harm Income Mobility and Growth? An Assessment of the Growth Impact of Income and Education Inequality in Paraguay 1992-2002

Thomas Otter

April 2009

Platz der Göttinger Sieben 3 · 37073 Goettingen · Germany · Phone: +49-(0)551-398172 · Fax: +49-(0)551-398173
e-mail: uwia@gwdg.de · http://www.iai.wiwi.uni-goettingen.de

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 macroeconomic 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.[1]

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

1 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.[2] 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.

2 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[3]. 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

3 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.

error terms. Third, out 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.[4] 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:

$$\ln\ y_{cs,\ t+1} = E\left[\ln y_{cs,\ t+1}\ |\ X_{cs,\ t}\right] + \eta_{c,\ t+1} + \varepsilon_{cs,\ t+1} \qquad (2.1)$$

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:

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

and welfare estimates are based on:

$$\tilde{\mu}_{t+1} = E\left[W_{t+1}\ |\ m_{t},\ \tilde{y}_{h,t+1}\right] \qquad (2.3)$$

4 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.[5]

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[6] 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.

Table2.1 Welfare estimates, Paraguay, Selected Years

Region	Poverty ratio (per capita income) 1992	2002	Poverty change	Inequality change (Gini)	Income growth
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 Capter 1

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

6 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 me 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).[7] 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

7 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[8] 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[9]; 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 hetersokedasticity 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,

8 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.

9 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.

income inequality, human capital inequality, male and female human capital household demographics and employment sector. The model we estimate can be represented as:

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

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. **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.[10] Given this approach, we are limited in our choice of covariates in **X** to what the census has to offer. District fixed effects, represented by α_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:

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

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.:

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

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

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

$$I(y_i) = \ln M(y_i) - M(\ln y_i) \tag{2.7}$$

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

$$\ln M(y_i) = M(\ln y_i) + I(y_i) \tag{2.8}$$

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:

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

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: $Ln(pcy_{92}) – Ln(pcy_{02})/10$	-0.01	0.02	-0.06	0.05
lny92	Log income per capita 1992: $Ln(pcy_{92})$	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: $(lntot_{02} / lntot_{92})$-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: (esp_{02} / esp_{92})-1	0.40	0.29	-0.43	0.63
dhijo	Changes in number of children: $(nhijo_{02} / nhijo_{92})$-1	0.52	0.58	-0.17	6.19
dhedu	Changes in household head's education in number of years: $(hedu_{02} / hedu_{92})$ -1	-0.07	0.06	-0.16	0.04
hage92	Age of household head 1992	46.83	16.14	21.32	82.06
lntot92	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: (est_{02} / est_{92})-1	-0.03	0.05	-0.10	0.03
dginiy	Changes in income inequality: $(giniy_{02} / giniy_{92})$ -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: $Ln(pcy_{92}) – Ln(pcy_{02})/10$	-0.01	0.02	-0.07	0.05
lny92	Log income per capita 1992: $Ln(pcy_{92})$	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: $(lntot_{02} / lntot_{92})-1$	-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 sector 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: $(nhijo_{02} / nhijo_{92})-1$	-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: $(giniy_{02} / giniy_{92}) -1$	-0.02	0.02	-0.08	0.03
ginie92	Education inequality: Gini coefficient wrt $hhedu_{92}$	0.29	0.01	0.27	0.31
dginie	Changes in education inequality: $(ginie_{02} / ginie_{92})-1$	-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: $\mathrm{Ln}(pcy_{92}) - \mathrm{Ln}(pcy_{02})/10$	0.01	0.04	-0.11	0.13
lny92	Log income per capita 1992: $\mathrm{Ln}(pcy_{92})$	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: $(lntot_{02} / lntot_{92})$-1	-0.01	0.09	-0.32	0.48
lntot92	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: $(hedu_{02} / hedu_{92})$ -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: $(giniy_{02} / giniy_{92})$ -1	0.02	0.04	-0.11	0.15
dhana	Changes in percentage of households with at least one analphabetic: $(dhana_{02} / dhana_{92})$-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: $(hfem_{02} / hfem_{92})$-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: $(ginie_{02} / ginie_{92})$-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 $hhedu_{92}$	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: $Ln(pcy_{92}) – Ln(pcy_{02})/10$	0.02	0.04	-0.11	0.16
lny92	Log income per capita 1992: $Ln(pcy_{92})$	10.92	0.55	9.82	12.74
dginiy	Changes in income inequality: $(giniy_{02} / giniy_{92})$ -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: $(lntot_{02} / lntot_{92})$-1	-0.01	0.09	-0.32	0.46
ginie92	Education inequality: Gini coefficient wrt $hhedu_{92}$	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
lntot92	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: $(hecom_{02} / hecom_{92})$-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: $(hecs_{02} / hecs_{92})$-1	0.36	1.42	-1.00	22.19
giniy92	Income inequality: Gini coefficient wrt 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: $(hedu_{02} / hedu_{92})$ -1	0.19	0.24	-0.13	3.61
dhetran	Change of percentage of household head employed in transport and communications sector: $(hetran_{02} / hetran_{92})$-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: $(ginie_{02} / ginie_{92})$-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: $Ln(pcy_{92}) – Ln(pcy_{02})/10$	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 pcy_{92}	0.54	0.02	0.44	0.63
ginie92	Education inequality: Gini coefficient wrt $hhedu_{92}$	0.28	0.02	0.25	0.36
lny92	Log income per capita 1992: $Ln(pcy_{92})$	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: $(giniy_{02} / giniy_{92})$ -1	-0.01	0.05	-0.28	0.17
dginie	Changes in education inequality: $(ginie_{02} / ginie_{92})$-1	0.02	0.05	-0.16	0.14
dlntot	Changes in log total persons in household: $(lntot_{02} / lntot_{92})$-1	-0.02	0.08	-0.32	0.35
dhijo	Changes in number of children: $(nhijo_{02} / nhijo_{92})$-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[11] 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

11 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):

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

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:

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

The β 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 ξ and ω are uncorrelated with the regressors $\tilde{y}$ and **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.[12] Both ξ and ω 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 $\tilde{g}$.[13]

The prediction errors should also be uncorrelated with regressors in **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:

---

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

13 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.

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

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

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

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 $\boldsymbol{\lambda}$ = (β,γ); φ 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:

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

where the model part variance is:

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

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. Finally, 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

Dependent: growth	*coef*	*t-val(adj)*	*t(2SLS)*	*incr(se)*
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

Dependent: growth	*Coef*	*t-val(adj)*	*t(2SLS)*	*incr(se)*
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

Dependent: growth	*Coef*	*t-val(adj)*	*t(2SLS)*	*incr(se)*
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

Dependent: growth	*Coef*	*t-val(adj)*	*t(2SLS)*	*incr(se)*
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
dhecom	0.0015	7.7851	2.7760	2.8044
hecom92	0.1196	10.2076	2.0490	4.9818
dhecs	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

Dependent: growth	*Coef*	*t-val(adj)*	*t(2SLS)*	*incr(se)*
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 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	0.99	0.99	0.60	1.00	1.00
	(11.92)***	(11.73)***	(12.13)***	(17.62)***	(12.30)***	(17.72)***
lny92	-0.05	-0.05	-0.05	-0.05	-0.05	-0.05
	(-16.41)***	(-16.63)***	(-16.98)***	(-15.87)***	(-16.83)***	(-16.04)***
hhedu92	0.003	0.003	0.003	0.003	0.003	0.003
	(7.83)***	(8.15)***	(8.16)***	(7.40)***	(8.10)***	(7.60)***
dlntot	-0.05	-0.05	-0.05	-0.06	-0.05	-0.05
	(-8.54)***	(-8.67)***	(-8.52)***	(-8.59)***	(-8.85)***	(-8.65)***
dsedu	0.003	0.003	0.003	0.003	0.003	0.003
	(4.05)***	(3.98)***	(3.98)***	(3.48)***	(4.06)***	(3.47)***
desp	-0.03	-0.03	-0.03	-0.01	-0.03	-0.03
	(-5.97)***	(-6.22)***	(-6.01)***	(-2.73)***	(-6.36)***	(-4.23)***
dhijo	0.005	0.005	0.005	-0.01	0.005	0.005
	(-4.19)***	(-4.26)***	(-4.22)***	(-4.44)***	(-4.28)***	(-4.46)***
dhedu	-0.03		-0.02	-0.01		
	(-1.97)*		(-1.86)*	(-0.67)		
hage92	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004
	(6.19)***	(6.31)***	(6.41)***	(6.10)***	(6.25)***	(6.16)***
lntot92	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04
	(-5.67)***	(-5.79)***	(-5.79)***	(-5.45)***	(-5.74)***	(-5.52)***
nhijo92	0.01	0.01	0.01	0.01	0.01	0.01
	(3.27)***	(3.36)***	(3.31)***	(3.04)***	(3.36)***	(3.09)***
hsedu92	0.002	0.001	0.002	0.002	0.001	0.001
	(2.83)***	(2.68)**	(2.81)**	(2.75)**	(2.61)**	(2.71)**
hfem92	0.02	0.02	0.02	0.02	0.02	0.02
	(3.18)***	(3.26)***	(3.35)***	(3.04)***	(3.15)***	(3.08)***
dest	0.15	0.11	0.12	-0.02	0.14	0.14
	(3.73)***	(3.17)***	(4.41)***	(-0.88)	(6.04)***	(-1.33)
dginiy	0.03	-0.01		-0.02		-0.78
	(1.12)	(-0.93)		(-0.94)		(-2.63)***
giniy92	-0.80	-0.73	-0.74		-0.78	
	(-5.28)***	(-4.93)***	(-5.25)***		(-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) looses quality. Adjusted R squared decreases from 0.909 (Table 2.14) to 0.836 (Table 2.9), and the models 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) looses quality. Adjusted R squared decreases from 0.867 (Table 2.15) to 0.467 (Table 2.10), and the models 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 Guarani (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) looses almost all its quality. Adjusted R squared decreases from 0.601 (Table 2.16) to 0.087 (Table 2.11), and the models 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	0.80	0.88	1.45	1.10	1.44	0.73	0.80	0.83
	(3.20)***	(2.54)**	(2.76)**	(5.43)***	(3.29)***	(5.38)***	(2.30)**	(45.17)***	(47.75)***
lny92	-0.07	-0.07	-0.07	-0.07	-0.07	-0.07	-0.07	-0.07	-0.07
	(-39.15)***	(-38.73)***	(-46.45)***	(-38.95)***	(-37.98)***	(-38.96)***	(-46.44)***	(-46.51)***	(-55.22)***
hhedu92	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004
	(23.96)***	(23.56)***	(25.18)***	(23.83)***	(23.04)***	(23.84)***	(25.07)***	(25.18)***	(27.66)***
dlntot	-0.05	-0.05	-0.05	-0.05	-0.05	-0.05	-0.05	-0.05	-0.05
	(-9.25)***	(-9.04)***	(-9.10)***	(-9.27)***	(-8.61)***	(-9.27)***	(-8.96)***	(-9.01)***	(-8.57)***
lntot92	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03
	(-14.89)***	(-14.63)***	(-14.37)***	(-14.86)***	(-14.28)***	(-14.87)***	(-14.29)***	(-14.34)***	(-14.58)***
hage92	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	(10.84)***	(10.58)***	(10.19)***	(10.69)***	(10.84)***	(10.70)***	(10.20)***	(10.20)***	(12.27)***
hetran92	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04
	(-5.29)***	(-5.17)***	(-5.06)***	(-5.22)***	(-4.80)***	(-5.22)***	(-5.01)***	(-5.01)***	(-4.79)***
heagro92	0.05	0.05	0.04	0.05	0.05	0.05	0.04	0.04	0.04
	(5.20)***	(5.01)***	(4.55)***	(5.20)***	(4.53)***	(5.21)***	(4.53)***	(4.55)***	(4.50)***
dhijo	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(-3.35)***	(-3.45)***	(-3.78)***	(-3.34)***	(-3.60)***	(-3.35)***	(-3.79)***	(-3.79)***	(-3.69)***
giniy92	-5.06		-3.84	-2.53	2.00	-2.59			
	(-3.27)***		(-2.53)**	(-2.25)**	(1.57)	(-2.30)**			
dginiy	-0.08	-0.07		-0.08	-0.01	-0.08			
	(-3.48)***	(-2.79)**		(-3.17)***	(-0.41)	(-3.18)***			
ginie92	7.07	0.36	5.62		-5.07	0.17	0.50		
	(2.37)**	(0.16)	(1.89)*		(-1.94)*	(3.91)***	(0.23)		
dginie	0.19	0.14	0.15	0.16		0.16	0.12	0.12	
	(7.48)***	(6.87)***	(6.57)***	(7.35)***		(7.39)***	(6.25)***	(6.39)***	
ginie92 sqrd	-12.05	-0.30	-9.52	0.30	9.14		-0.55	0.31	0.28
	(-2.31)**	(-0.08)	(-1.83)*	(3.87)***	(2.01)**		(-0.14)	(4.12)***	(3.61)***

giniy92 sqrd	5.22	-0.22	4.02	2.50	-2.25	2.56	-0.12	-0.12	-0.06
	(3.14)***	(-4.83)***	(2.45)**	(2.07)**	(-1.62)	(2.12)**	(-4.11)***	(-4.18)***	(-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	gry	gry	gry	gry	gry	gry	gry
(Constant)	0.91	0.88	0.90	0.82	0.93	0.81	0.80
	(27.39)***	(25.34)***	(29.73)***	(28.41)***	(29.09)***	(31.89)***	(24.80)***
lny92	-0.07	-0.07	-0.07	-0.06	-0.07	-0.06	-0.07
	(-31.05)***	(-30.96)***	(-31.48)***	(-31.00)***	(-32.92)***	(-31.51)***	(-29.82)***
D13	0.03	0.03	0.03	0.03	0.03	0.03	0.03
	(12.87)***	(17.70)***	(13.25)***	(14.03)***	(17.55)***	(14.44)***	(16.55)***
dlntot	-0.08	-0.08	-0.08	-0.08	-0.08	-0.08	-0.09
	(-13.99)***	(-14.27)***	(-14.30)***	(-14.26)***	(-13.84)***	(-14.57)***	(-14.27)***
lntot92	-0.05	-0.06	-0.05	-0.04	-0.06	-0.04	-0.06
	(-6.50)***	(-7.71)***	(-6.52)***	(-5.88)***	(-7.83)***	(-5.90)***	(-7.55)***
hsedu92	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	(13.75)***	(13.46)***	(13.82)***	(12.75)***	(14.30)***	(12.83)***	(12.58)***
dsedu	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	(9.14)***	(8.46)***	(9.12)***	(9.40)***	(8.75)***	(9.39)***	(8.65)***
D15	0.01	0.02	0.01	0.01	0.03	0.01	0.03
	(2.73)**	(6.28)***	(2.70)**	(3.23)***	(7.38)***	(3.20)***	(6.95)***
D6	-0.04	-0.03	-0.04	-0.04	-0.03	-0.04	-0.04
	(-6.65)***	(-5.73)***	(-6.64)***	(-7.13)***	(-5.63)***	(-7.12)***	(-6.31)***
hana92	-0.05	-0.04	-0.05	-0.05	-0.05	-0.05	-0.05
	(-10.34)***	(-9.22)***	(-10.33)***	(-10.67)***	(-9.43)***	(-10.67)***	(-10.04)***
dhedu	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	(9.03)***	(8.52)***	(9.04)***	(9.37)***	(8.39)***	(9.39)***	(9.00)***
D10	-0.02	-0.01	-0.02	-0.02	-0.01	-0.02	-0.01
	(-8.96)***	(-4.97)***	(-9.11)***	(-9.22)***	(-3.86)***	(-9.36)***	(-3.10)***
dginiy	-0.16		-0.16	-0.15		-0.15	
	(-8.93)***		(-9.61)***	(-8.21)***		(-8.84)***	
dhana	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001
	(-4.49)***	(-4.18)***	(-4.50)***	(-4.49)***	(-4.10)***	(-4.50)***	(-3.70)***

D8	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	(3.54)***	(4.67)***	(3.57)***	(3.22)***	(4.73)***	(3.25)***	(3.87)***
dhfem	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(-1.92)**	(-2.37)**	(-1.93)**	(-1.62)*	(-2.37)**	(-1.63)*	(-2.53)**
hage92	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	(5.73)***	(6.30)***	(5.69)***	(4.85)***	(6.97)***	(4.81)***	(6.21)***
giniy92	-0.02	0.09		-0.02	-0.11		0.04
	(-0.66)	(3.41)***		(-0.58)	(-3.72)***		(1.54)
dginie	-0.14	-0.11	-0.14				-0.03
	(-5.19)***	(-3.90)***	(-5.19)***				(-1.16)
nhijo92	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	(3.15)***	(3.79)***	(3.16)***	(2.94)***	(3.89)***	(2.95)***	(3.89)***
hfem92	0.02	0.01	0.02	0.02	0.00	0.02	-0.01
	(2.62)**	(0.55)	(2.55)**	(2.29)**	(0.75)	(2.23)*	(-0.45)
ginie92	-0.33	-0.29	-0.33	-0.20	-0.23	-0.21	
	(-6.74)***	(-5.69)***	(-7.17)***	(-4.67)***	(-4.78)***	(-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

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

hsedu92	0.003	0.003	0.002	0.003	0.003	0.003	0.003	0.004
	(2.74)**	(3.68)***	(1.87)*	(3.03)***	(3.20)***	(2.81)***	(3.55)***	(4.46)***
hfem92	0.02	0.03	0.03	0.02	0.02	0.04	0.02	0.03
	(3.52)***	(4.99)***	(5.39)***	(4.22)***	(3.68)***	(6.78)***	(4.75)***	(5.32)***
dhedu	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(-3.73)***	(-4.48)***	(-2.11)**	(-3.87)***	(-3.92)***	(-2.90)***	(-4.42)***	(-4.80)***
dhetran	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	(2.83)***	(2.70)**	(2.68)**	(2.85)***	(2.82)***	(2.56)**	(2.68)**	(2.67)**
dsedu	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	(1.48)	(1.15)	(1.34)	(1.57)	(1.74)*	(1.03)	(1.09)	(1.53)
dginie	-0.02	-0.04	-0.06	-0.02		-0.07	-0.04	
	(-3.22)***	(-5.09)***	(-7.41)***	(-3.31)***		(-9.20)***	(-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	0.48	1.01	0.84	0.90	0.49	0.94	0.55	0.82
	(5.61)***	(4.88)***	(7.65)***	(5.04)***	(5.93)***	(5.02)***	(7.07)***	(18.88)***	(4.99)***
hage92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(7.05)***	(6.60)***	(7.12)***	(7.73)***	(7.17)***	(6.63)***	(7.09)***	(6.68)***	(7.70)***
hhedu92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(10.81)***	(10.71)***	(10.79)***	(10.55)***	(10.81)***	(10.72)***	(10.81)***	(10.73)***	(10.55)***
giniy92	-1.49		-1.37	-1.22	-1.47	-0.06	-1.42	-0.06	-1.23
	(-3.23)***		(-2.99)***	(-2.42)**	(-3.18)***	(-2.75)**	(-3.11)***	(-2.74)**	(-2.45)**
ginie92	1.26	0.99		0.91	1.04	1.00	0.59	0.61	1.03
	(1.96)*	(1.54)		(1.29)	(1.66)	(1.56)	(17.53)***	(18.15)***	(1.51)
lny92	-0.06	-0.06	-0.06	-0.07	-0.06	-0.06	-0.06	-0.06	-0.06
	(-29.15)***	(-29.17)***	(-29.25)***	(-27.36)***	(-30.26)***	(-29.15)***	(-29.25)***	(-29.32)***	(-27.91)***
nhijo92	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(-8.94)***	(-8.60)***	(-9.12)***	(-10.60)***	(-9.11)***	(-8.63)***	(-9.06)***	(-8.74)***	(-10.58)***
dginiy	-0.16	-0.15	-0.15		-0.15	-0.15	-0.15	-0.15	
	(-15.54)***	(-15.37)***	(-15.46)***		(-15.49)***	(-15.40)***	(-15.51)***	(-15.39)***	
dginie	0.02	0.02	0.01	-0.01		0.02	0.01	0.01	
	(1.43)	(1.32)	(0.98)	(-0.72)		(1.32)	(1.22)	(1.21)	
dlntot	-0.07	-0.08	-0.07	-0.08	-0.07	-0.08	-0.07	-0.07	-0.09
	(-8.64)***	(-8.73)***	(-8.46)***	(-8.98)***	(-8.55)***	(-8.69)***	(-8.58)***	(-8.68)***	(-9.09)***
dhijo	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(-2.16)**	(-1.95)*	(-2.70)**	(-1.82)*	(-2.24)**	(-1.97)*	(-2.49)**	(-2.20)**	(-1.78)*
giniy92 sqrd	1.33	-0.05	1.22	1.12	1.31		1.26		1.13
	(3.10)***	(-2.59)***	(2.86)***	(2.40)***	(3.05)***		(2.98)***		(2.42)***
ginie92 sqrd	-1.16	-0.66	1.02	-0.54	-0.80	-0.68			-0.74
	(-1.04)	(-0.59)	(17.43)***	(-0.44)	(-0.74)	(-0.61)			(-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 1

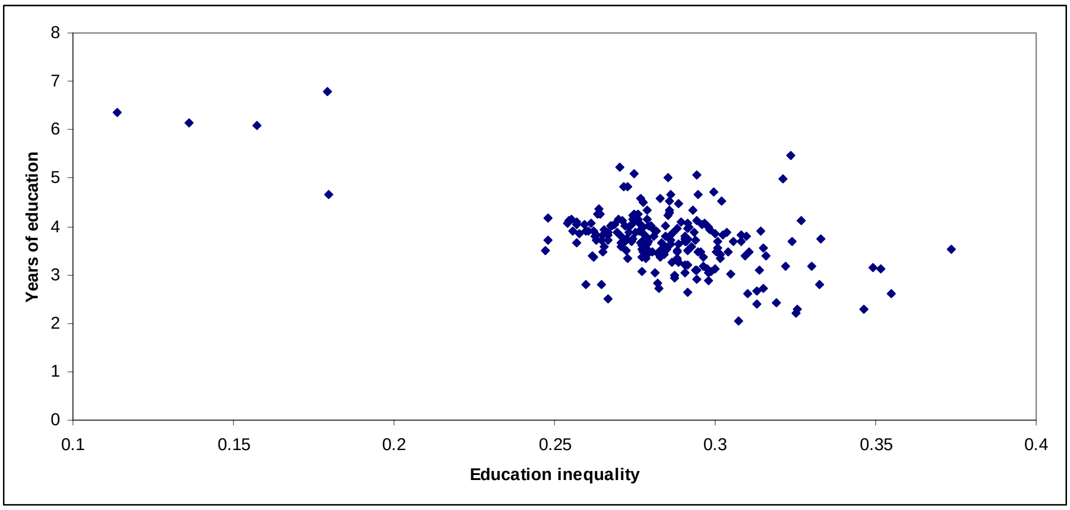
2.7 Discussion

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 then 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 observation 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).

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.

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.

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.

Bourguignon, F. (2004) "The poverty-growth-inequality triangle", mimeo, Paper presented at the Indian Council for Research on International Economic Relations, New Delhi.

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.

Deininger, K. and Okidi, J. (2003), Growth and poverty reduction in Uganda, 1999-2000: Panel data evidence, *Development Policy Review,* **21** (4)**,** 481-509.

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. and Lanjouw, P. (2005), Imputed welfare estimates in regression analysis, *Journal of Economic Geography,* **forthcoming**.

Forbes, K. J. (2000), A reassessment of the relationship between inequality and growth, *American Economic Review*, **90** (4)**,** 869-887.

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.

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.

Hassler, John & José V. Rodríguez Mora (1998), IQ, Social Mobility and Growth. Seminar Papers No. 635, Institute for International Economic Studies, Stockholm University. January.

Hentschel, J., Lanjouw, J. O., Lanjouw, P. and Poggi, J. (2000), Combining census and survey data to trace the spatial dimensions of poverty: A case study of Ecuador, *World Bank Economic Review,* **14** (1)**,** 147-65.

Jarvis, S. and Jenkins S.P. (1996), Changing Places: Income Mobility and Poverty Dynamics in Britain. University of Essex Paper Number 96-19, processed.

Krugman, P. (1992), The Rich, the Right, and the Facts. *The American Prospect*. **11** 19-31.

Lundberg, M. and Squire, L. (2003), The simultaneous evolution of growth and inequality, *Economic Journal*, **113** (April)**,** 326-344.

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.

Murphy, Kevin M., Andrei Scheifer and Robert W. Vishny, 1991. "The Allocation of Talent: Implications for Growth", *Quarterly Journal of Economics*, Vol 106(2): 503-530.

Otter, Th., (2006), Micro level estimation of welfare in Paraguay. Household income estimates at district level 1992 and 2002, mimeo.

Otter, Th., (2004) Migración Interna, Pobreza y Capital Social en Paraguay entre 1985 y 2002 – ¿Desafío, Obstáculo u Oportunidad para Políticas Públicas de Reducción de Pobreza?, paper for DGEEC Paraguay, mimeo.

Otter, Th., (2002), Activos y Pobreza en Paraguay, paper for DGEEC Paraguay, mimeo.

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.

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.

Slemrod, J. (1992), Taxation and Inequality: A Time-Exposure Perspective, in: J.M. Poterba, ed., Tax Policy and the Economy, Vol. 6 (Cambridge, MA: MIT Press for the NBER) 105-127.

Thorbecke, E. and Charumilind, C. (2002), Economic inequality and its socioeconomic impact, *World Development,* **30** (9)**,** 1477-1495.