United Nations University
WIDER
World Institute for Development Economics Research

Discussion Paper No. 2003/56

Opening the Convergence Black Box

Measurement Problems and Demographic Aspects

Carlos Azzoni, Naercio Menezes-Filho
and Tatiane Menezes*

August 2003

Abstract

In this paper we address the issues involved with the use of microeconomic data, that is, household surveys, to compare the patterns of income growth among different regions instead of the commonly used aggregate data. In particular, we investigate the issues of aggregation of household income to regional income and the problem of demography. As returns to experience generally differ across regions, differences in the patterns of income growth across regions in the same time interval will differ across age groups, which means that convergence or divergence of aggregate income among regions will depend on the age structure of their population. We apply these concepts to the case of the states of Brazil, for which we have repeated cross sections from a rich household survey. We find that patterns of income growth vary a great deal across birth cohorts, depending on the economic returns to experience.

Keywords: Brazil, regional convergence, regional growth, birth cohorts, micro data
JEL classification: O15, R11, R12, R23

---

*Department of Economics, University of São Paulo, Brazil.

This study has been prepared within the UNU/WIDER project on Spatial Disparities in Human Development, directed by Ravi Kanbur and Tony Venables.

UNU/WIDER gratefully acknowledges the financial contribution to the project by the Government of Sweden (Swedish International Development Cooperation Agency—Sida).

Acknowledgements

The authors would like to thank the Interamerican Development Bank, FIPE (Fundação Instituto de Pesquisas Econômicas), and Finep-Pronex, for different forms of support for this research.

The World Institute for Development Economics Research (WIDER) was established by the United Nations University (UNU) as its first research and training centre and started work in Helsinki, Finland in 1985. The Institute undertakes applied research and policy analysis on structural changes affecting the developing and transitional economies, provides a forum for the advocacy of policies leading to robust, equitable and environmentally sustainable growth, and promotes capacity strengthening and training in the field of economic and social policy making. Work is carried out by staff researchers and visiting scholars in Helsinki and through networks of collabourating scholars and institutions around the world.

www.wider.unu.edu *publications@wider.unu.edu*

UNU World Institute for Development Economics Research (UNU/WIDER)
Katajanokanlaituri 6 B, 00160 Helsinki, Finland

Camera-ready typescript prepared by Lorraine Telfer-Taivainen at UNU/WIDER
Printed at UNU/WIDER, Helsinki

ISSN 1609-5774
ISBN 92-9190-498-8 (printed publication)
ISBN 92-9190-499-6 (internet publication)

1 Introduction

The problem of comparing economic growth across regions or countries is on the forefront of economic analysis nowadays (e.g. Hanushek and Kimko 2000; Bils and Klenow 2000). However, most of the papers in the literature use real per capita GDP as the variable whose growth will be analyzed. To the best of our knowledge, no study so far has addressed the issue of convergence using individual level data, that is, comparing the growth of income of individuals that live in different regions.[1] Introducing the concept of convergence with micro data means that we should now be concerned with factors that are correlated with individual income growth and vary systematically across regions. In this paper we concentrate on the effects of demography.

The use of micro data to test economic models has been applied in other fields of economics, especially in consumption and labour supply (e.g. Browning et al. 1985 and Attanasio and Browning 1995). According to Attanasio and Browning (1995:1,119) life cycle models are rejected, under the assumption of a representative consumer, on aggregate time series because of 'aggregation bias and insufficient allowance for the dependence of consumption on demographics'. It is our belief that the conclusions reached by several empirical studies of convergence among countries, like those of Barro and Sala i Martin (1995) Mankiw et al. (1992) and Islam (1995), to cite only a few, are also dependent on the demographic structures of the countries under study, and that introducing demography may provoke substantial changes in the results obtained so far.

The relationship between demography and income has a long tradition in economic literature, starting with the works of Freeman (1979) and Welch (1979), investigating the effect of the baby boom cohort on inequality in the US. This literature has been recently extended by Higgins and Williamson (1997:37) who look at cross-country evidence to find that 'large mature working age cohorts are associated with lower aggregate inequality and large young adult cohorts are associated with higher aggregate inequality'. In this paper we intend to further extend the analysis to look at the effect of demography on income dynamics, that is, we use micro data to examine whether the age structure and returns to experience of different regions has any impact on the convergence of their income.

To investigate these factors we use repeated cross sections of a Brazilian household survey and compare income patterns of individuals living in different states. The use of micro data raises several interesting questions regarding the comparison between individual and aggregate income. First of all, it seems important to compare state GDP with individual income aggregated up to state level, to check how much of what is produced is returned to individuals and is declared in the household surveys. Moreover, it is important to check

[1] The paper that uses a methodology closest to ours is Jalan and Ravallion (2002) that discusses the effect of geographic variables on household consumption growth in China.

whether using the arithmetic or the geometric means make a difference when comparing the levels and the growth of income across states—this we do in Section 2. In Section 3 we compare the results of the traditional cross-regions Barro regressions with the results using micro data. Section 4 confirms that income convergence varies across birth cohorts and explains that this is due to differences in returns to experience across states. The final section concludes and proposes future extensions of this research.

2 Comparing GDP with income

Brazil is well known for its high levels of regional inequality. Being a country with a large territory, that should not be surprising. The Northeast region of Brazil was home to 28 per cent of Brazilian population in the year 2000 and produced only 13 per cent of Brazilian GDP in the year 1998; the rich Southeast region presented 43 per cent of Population and produced 58 per cent of GDP. Per capita income in the Northeast was 54 per cent below the national average, while in the Southeast it was 36 per cent above that level. The poorest state, Piauí, in the Northeast region, had a per capita income level 5.6 times lower than the richest state, São Paulo, in the Southeast region[2]. The above relative figures are not too different from the situation half a century ago, for in 1947 the per capita income relation between Piauí and São Paulo states was five to one. Regional income convergence in Brazil using macro data has been the object of study by some authors, such as Ferreira and Diniz (1995), Schwartsman (1996), Zini (1998), Ferreira (2000), and Azzoni (2001). Micro data were only used in Azzoni et al. (2000) and Azzoni and Servo (2002), with different interests as the ones in this paper.

We now compare the official state GDP figures provided by the Brazilian Census Bureau (macro data) with the income data available in the household surveys conducted by the same Bureau (micro data), using sampling weights to aggregate individual income up to state level. Since the first data set is comparable to data used in convergence studies everywhere, we want to check whether or not our aggregated micro data provides the same results as the standard macro data for Brazilian states.

We deal with 19 of the 25 Brazilian states, since survey data is not available for the unpopulated states of the Amazon Region; Brasília (the Federal District) was also dropped, since the economic dynamics in that area is strongly determined by the federal government salary policy, and thus is not driven by economic factors such as in other areas. In total, only 9 per cent of the Brazilian population of 170 million in 2000 is not included in the study. Table 1 presents a summary of the available data: column (1) shows the traditional GDP macro data, as provided by IBGE; column (2) presents income obtained from the aggregation of micro data. In general, the aggregate income numbers correspond to about 50 per cent of GDP, though for the poorer States (AL, PB, PI and MA) it can get close to

---

[2] See www.ibge.gov.br/ibge/estatistica/economia/contasregionais/, for information on regional income for Brazil, and http://www.ibge.gov.br/ibge/estatistica/populacao/censo2000/ for information on population.

80 per cent. A possible reason for this discrepancy is the under-reporting of income, especially from other sources than labour. For the eight states with the highest GDP, the ranking is the same in both columns; the four lower GDP numbers belong to the same states; some minor modifications in ranking are present for the other middle-size states.

Table 1: Comparing GDP with micro data aggregate income (Brazilian R$ million)

States		GDP Macro Data (1)	Rank	Income Aggregated Micro Data (2)	Rank	Ratio (2)/(1)
São Paulo	SP	306,569	1	143.000	1	.47
Rio de Janeiro	RJ	96,947	2	50.000	2	.52
Minas Gerais	MG	86,527	3	42.000	3	.49
Rio Grande do Sul	RS	68,689	4	31.000	4	.45
Paraná	PR	52,438	5	26.000	5	.50
Bahia	BA	36,735	6	18.900	6	.51
Santa Catarina	SC	31,633	7	15.800	7	.50
Pernambuco	PE	23,261	8	11.500	8	.49
Ceará	CE	17,453	9	9.900	10	.57
Espírito Santo	ES	16,087	10	6.940	11	.43
Goiás	GO	15,906	11	10.900	9	.69
Mato Grosso do Sul	MS	9,219	12	5.020	15	.54
Mato Grosso	MT	9,086	13	6.400	12	.70
Maranhão	MA	7,353	14	5.510	14	.75
Paraíba	PB	6,936	15	5.530	13	.80
Rio Grande do Norte	RN	6,617	16	4.330	16	.65
Alagoas	AL	5,711	17	4.300	17	.75
Sergipe	SE	4,805	18	2.600	19	.54
Piauí	PI	4,192	19	3.050	18	.73
Global		806,163		400.680		.50

Source: See text.

Income and GDP per capita are displayed in Table 2. Again, column (1) presents GDP per capita figures (macro data) and column (2) displays per capita income figures based on the aggregation of individual data. Here we introduce an important modification in the way we calculate income, for we exclude all zero-income cases, and work only with people with positive income. This was necessary because we intend to run all regressions below using a logarithmic specification. As can be seen in Table 2, income values are larger than GDP values, although the numbers are not to far away from each other. That is expected, since our income data includes only households with non-zero income and GDP per capita includes all households.

It is interesting to notice that the ten richest states are the same in both columns, although some modifications in rank are present, the largest being for Rio de Janeiro, with a jump of five positions. However, in percentage terms, the change in Rio de Janeiro was the second smallest. As a whole, poorer states present a higher percentual difference between the two figures, indicating that zero-income cases are more important for those cases than for richer states. Column (3) shows the average of the logarithms of micro data per capita income. These values will be used ahead in the paper to state the importance of the right measurement of per capita values for convergence studies. As numbers show, the figures are always lower (between 2 per cent and 10 per cent) than those in column (2).

Table 2: Per capita GDP and income, 1997 (Brazilian R$/year)

State	Macro data log of GDP per capita (1)	Rank	Micro data log of per capita income (2)	Rank	Average of the log of micro data per capita income (3)	Rank
SP	8,940	1	9.150	1	8.624	1
RJ	8,730	2	8.890	7	8.131	8
RS	8,713	3	8.773	2	8.395	3
SC	8,615	4	8.889	6	8.280	4
PR	8,509	5	8.814	4	8.219	6
ES	8,491	6	8.651	3	8.424	2
MG	8,395	7	8.588	5	8.240	5
MS	8,311	8	8.677	10	8.063	9
MT	8,140	9	8.868	9	8.059	10
PE	7,898	10	8.291	8	8.149	7
GO	7,888	11	8.557	12	7.766	14
SE	7,834	12	8.342	16	7.702	15
BA	7,823	13	8.265	15	7.787	13
RN	7,704	14	8.277	11	7.814	11
CE	7,688	15	8.184	13	7.791	12
AL	7,524	16	8.419	17	7.591	17
PB	7,492	17	8.254	14	7.666	16
PI	7,214	18	7.891	19	7.308	19
MA	7,090	19	7.926	18	7.426	18

Source: See text.

In Table 3 the rates of growth over the period 1981-97 of the different measured state incomes are presented. Again, column (1) refers to macro data GDP per capita, and columns (2) and (3) are our micro data per capita income. Differences in rankings based on growth rates are more pronounced than rankings based on total and per capita income values.[3] In general, the above tables indicate that our set of income and per capita income figures reproduce reasonably well the official GDP per capita figures. Although some differences are present, the numbers as a whole are sufficiently good to be used in the following calculations of regional income convergence in Brazil.

Table 3: Rates of growth in GDP and income per capita, 1981-97

	Macro data		Micro data			
State	Growth in the log of GDP per capita (1)	Rank	Growth in the log of per capita income (2)	Rank	Growth in the average of the log of per capita income (3)	Rank
AL	0.0240	19	0.5169	1	0.4770	2
PB	0.2567	10	0.5140	2	0.3583	5
MA	0.1778	17	0.5057	3	0.5407	1
CE	0.4848	1	0.5044	4	0.3646	4
GO	0.3201	8	0.4774	5	0.3087	8
RN	0.2754	9	0.4428	6	0.2534	13
PR	0.4422	2	0.4399	7	0.1662	18
MG	0.4388	3	0.4311	8	0.3531	6
SP	0.2167	12	0.4114	9	0.3915	3
MT	0.1918	15	0.4038	10	0.2995	11
RJ	0.2099	13	0.3829	11	0.1954	16
PI	0.3959	6	0.3746	12	0.2751	12
ES	0.4206	4	0.3705	13	0.3061	9
MS	0.2519	11	0.3437	14	0.3008	10
SC	0.3884	7	0.3218	15	0.3421	7
PE	0.0641	18	0.2584	16	0.2015	15
RS	0.4012	5	0.2502	17	0.2351	14
BA	0.1791	16	0.2417	18	0.1954	17
SE	0.1992	14	0.1449	19	0.0538	19

Source: See text.

3 Macro data traditional convergence regressions

In this section we use the macro data for GDP and the income aggregated up to state level presented in the previous section, to estimate the traditional Barro-type convergence regressions, in which income growth over a period of time is regressed against the initial level of income. The objective here is to verify if the two sets of macro data will lead to

[3] This may happen because of changes over time in the proportion of young non-working people (excluded from the micro sample) across states, because of changes in the misreporting of income in our surveys or because of specific factors that affected GDP but not income for some states in 1981.

similar results. We will also verify if using the average of the logarithms will lead to similar results as for the logarithms of the average. Based on the Barro and Sala-i-Martin (1995) approach, we estimate the equation:

$$(\ln y_T - \ln y_0) = \alpha + \lambda \ln y_0 \tag{1}$$

in which y_0 is income in the beginning of the period, and $\lambda = \left(1 - e^{-\beta T}\right)$. The time necessary to reduce inequality by the half is given by $e^{-\beta t} = 1/2$.

Regression results are presented in Table 4. In column (1) traditional GDP (macro data) is used; column (2) presents results obtained with micro data per capita income. In both regressions absolute convergence is present, and at similar speeds: the calculated speed of convergence is only 13 per cent higher with micro data, with 71 years to half convergence, while 81 years would be needed with traditional GDP data. A better fit, as shown by the R2 is obtained with micro data. Thus, this first set of results indicates that the micro data we use, when transformed into macro data, lead to quite similar convergence results as the traditional macro data, GDP, data set. This is an important point to stress, for we will argue further in the paper that macro data type results might be misleading. And we will do that by exploring the richness of information present in the micro data available. Thus, it is important to show that the available micro data is equivalent to the available macro data, so that the observed differences in results will be due to factors to be stressed in the appropriate sections of the paper, and not to differences in the data sets.

Table 4: Traditional macro data convergence regressions

	Macro data	Micro data	
	Log of per capita GDP (1)	Log of per capita income (2)	Average of the logs of per capita income (3)
Constant	1.269 (0.3614)	1.556 (0.406)	1.271 (0.537)
λ	-0.127* (0.048)	-0.144* (0.050)	-0.127** (0.070)
β	0.00854	0.0097	0.0085
Years to half convergence	81	71	81
R2	0.15	0.23	0.19
Sample size	19	19	19

Note: * significant at 5%; ** significant at 10%.
Source: See text.

We now move to the last point to make in this section, which is the form of calculation of average income. When one thinks in terms of using household level income aggregated over cohorts to try and replicate the usual cross-country GDP per capita regressions, one has to address the issue of how to construct cohort income averages. As Attanasio and Weber (1993) point out, while national accounts data are only available as the arithmetic mean, we can aggregate the household income data in different ways. For example, we can use the logarithm of the arithmetic mean and the mean of the logarithm (i.e. the logarithm of the geometric mean). The difference between the two is Theil's entropy measure. If this measure differs across states and over time, the results of the usual cross-country regressions will depend on the way cohort income is generated. As column (3) indicates, results differ slightly, for the significance of the estimated coefficient for the initial income level is only marginally significant, suggesting that this is an immaterial practical point. The coefficient of the initial income in this case is almost the same as the one in the first column. In the remaining of this paper, the properly calculated average of the logarithms will be used.[4]

4 Exploring the richness of micro data using birth cohorts

4.1 Constructing birth cohorts

Micro data have not been used so far in the literature to examine issues of convergence. It is well known in the consumption and labour supply literature that with repeated cross sections it is possible to construct demographic cohorts based on date of birth, and calculate cohort-year means for all variables of interest, including income, education, labour force participation and living conditions; see Browning, Deaton and Irish (1985) and Attanasio and Browning (1995). We propose to extend this methodology to include the state of residence as another grouping variable and derive state-cohort-year means for the variables of interest. For example, income for a cohort *c* in state *s* in year *t* is:

$$\overline{y}_{cst} = \frac{\sum_{i}^{n_{cst}} \ln y_i}{n_{cst}} \tag{2}$$

Where n_{cst} is the number of household heads born in an interval of determined years (e.g. 1940 to 1945), living in state s in period t. Ten cohorts were constructed for each state in each year. The same procedure was applied to all variables included in the analysis, so that we have, for example, the average number of years of education for the household heads included in each cohort. The same holds for all other variables.

[4] That is another important reason for not using zero-income cases in our study.

The micro data we use come from the rich yearly household survey PNAD (Pesquisa Nacional por Amostra de Domicílios, carried out by the Brazilian Census Bureau, IBGE). This data can be used as a pseudo-panel, by constructing a model that looks like an individual-level model but is for cohorts (see Ravallion 1998). For each state, in each year, we have a sample of households based on the head's year of birth. The youngest cohort is the one with the household heads born in the 5-year period centered in 1972, while the oldest is formed by the household whose heads were born around the year 1922. For each cohort we calculated average levels of income, age, education, public infrastructure etc. We can then follow these cohorts over time (as they grow older) and analyze the influence of household variables on income growth. Although these are in fact different 'households', they can be considered as a good representation of their cohort, provided the samples are big enough.

The average number of households per cohort is 269. Due to the small number of observations in some cohorts, only 2,470 of them were considered in the analysis, instead of the 3,040 possible cases (10 cohorts x 19 states x 16 years). Only cohorts with at least 26 households were included in the sample. The small cohorts were located among the youngest in the first years of the period and among the oldest in the more recent years. This means that we are dealing in fact with an unbalanced panel. The total number of households considered in different years ranged from 49,514 to 90,776, with an average of 58,328. Table 5 provides descriptive information about the cohorts.

Table 5: Cohort description

Cohort	Date of birth	Average age in 1981	Average age in 1996	Minimum number of households per cohort	Maximum number of households per cohort
1	1922	59	75	26	1715
2	1927	54	70	70	2298
3	1932	49	65	80	2646
4	1937	44	60	99	3171
5	1942	39	55	132	3569
6	1947	34	50	144	4409
7	1952	29	45	182	5439
8	1957	24	40	237	6317
9	1962	19	35	293	6681
10	1967	14	30	317	7126

Source: See text.

The advantages of using cohort level data are many-fold. First and most importantly, the use of micro data allows us to control for changes in the composition of population in each state, something that cannot be done with aggregate data. Second, we can control for life cycle and generation effects, which means that we are really analyzing income growth within generations, or for a population with the same age. Third, it is possible to identify state fixed effects without having to rely only on the time component of the series, since we have various observations for a given state in a given year (10 in our case). Finally, one can rely on the differences across generations within a state-year group to identify the effects of human capital on growth, for example, which are not readily identified using aggregate data.

The main disadvantage of using cohort level data is that if there are measurement errors at the household level they are likely to be carried out to the cohort means, unless the cell sizes are big. Another possible problem with this methodology is related to migration across states, for it may cause the composition of the cohorts to change over time. If this change is driven by observed variables (e.g. education), then by including these variables in the convergence equation we avoid the problem. If it is driven by unobserved components, however, it may mean that we are not in effect controlling for household fixed effect. It should be said, however, that during the period analyzed migration flows were not strong in Brazil. The period involves the so-called 'lost decade', for the very bad macroeconomic performance of the country's economy, when unemployment rates rose in the country as a whole but especially in the rich areas, providing very few incentives to potential migrants.[5] Macroeconomic stabilization came in 1994 but even this could not provide enough incentives to migration for, at first, recovery was spread across the country as a whole, and second, recovery was not stable over time, with important oscillations in GDP growth and unemployment rates. Finally, the period under stabilization is short (late 1994 to 1997) compared to the years of almost stagnation that preceded it.

5 Regression results

Using the cohort data described above, we ran cross section and panel convergence regressions, with results presented in Table 6. Column (1) repeats column (3) from Table 4, for easy of comparison. Column (2) presents the results for a regression relating per capita income growth in a cohort/state against the initial level of income in that cohort/state, with a total number of 190 observations (cohort/states). As compared to the result presented in column (1), the speed of convergence increases by 747 per cent, leading to a half convergence period of only 10 years.

This change in speed is important enough to show what sort of difference in results could be obtained in using more appropriate micro data to study convergence. However, having cohort data allows us to control for other aspects involved in the convergence process. One important aspect to control for is the life cycle effect, for aggregate data deal equally with young and old people. Young people are in the uprising part of their income life cycle and old people are already in the declining part of it; over time, it is expected that young cohorts get richer and old cohorts get poorer, regardless of region or income level. Thus,

[5] See Baer (2002) for more details on the Brazilian economy during the period considered in this study.

the study of the evolution of income over a time period should control for that, and we do that by including cohort dummies into the regressions. Results are shown in column (3). Comparing to column (2), we observe that convergence occurs at a speed 68 per cent lower, although still higher (168 per cent) than in the case of macro data in column (1). Thus, the control for life cycle effects is important and changes significantly the calculated speed of convergence. When considering convergence within the same age range, what is in fact done in column (3), absolute convergence is still present, but at a much lower speed, around one third of the previous calculated speed.

Table 6: Micro data convergence regressions

	Macro data	Complete period (1981-97)		13 four-year rolling periods			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	1.271	5.511	2.534	2.454	1.656	3,208	5,494
	(0,537)	(0.347	(0.317	(0.099)	(0.067)	(0.096)	(0.127)
λ	-0.127	-0.684	-0.306	-0.291	-0.170	-0.443	-0.737
	(0.070)	(0.044)	(0.039)	(0.014)	(0.009)	(0.013)	(0.017)
β	0.0085	0.0720	0.0228	0.0215	0.0116	0.0366	0.0835
Years to half convergence	81	10	30	32	60	19	8
Education						0.134	0.173
	-	-	-	-	-	(0.004)	(0.008)
Cohort dummy	no	no	yes	yes	yes	yes	yes
State dummy	no	no	no	no	no	no	yes
Time dummy	no	no	no	no	yes	yes	yes
R2	.197	0.60	0.88	0.27	0.72	0.80	0.85
Sample Size	19	190	190	2,470	2,470	2,470	2,470

In columns (4)-(7) we present results of panel regressions, with four-year rolling periods (e.g. 1981-4), leading to 2470 observations. Column (4) is similar to column (3) and the results are quite similar, with a small decrease (5.7 per cent) in the speed of convergence. In column (5) we include time dummies to the previous regression, to control for shocks occurring in different years. The estimated speed of convergence decreases by 46 per cent in relation to column (4), indicating that controlling for time-related shocks is also important for convergence results.[6]

In the last two columns we move into conditional convergence, introducing education as another explanatory variables in the regressions. Thus, we are not any more considering that states are converging to a common steady state level of per capita income but, instead, that each state is converging to its own steady state level. In column (6) we control for

[6] This point was made by Temple (1999) in his thoroughly review of convergence studies

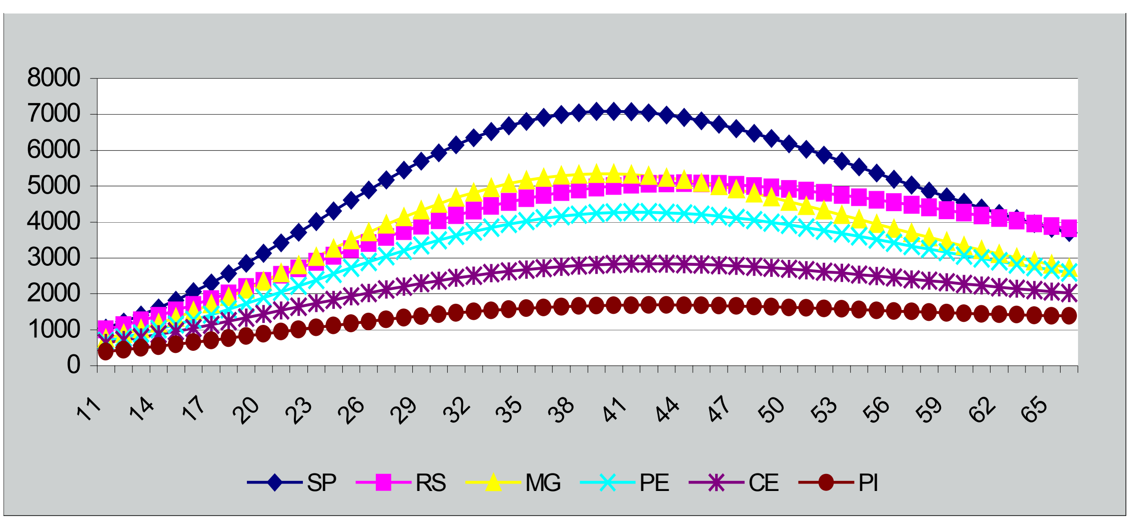
cohort and time dummies, as in column (5), and observe, as expected, a much higher speed of convergence (215 per cent). Education appears as positive and significant, as expected. Finally, in column (7) we introduce state dummies, to control for their specific characteristics, obtaining an additional 128 per cent increase in the speed of convergence.

6 Interpretation of the results

Why do the results look so different when we use micro instead of aggregate data? One possible explanation is that the aggregated results are affected by a compositional bias. It is well known from the labour literature that earnings tend to rise over the life cycle, at a decreasing rate. Figure 1 shows that this is also the case in Brazil, but that the earnings profile differs markedly across states. Returns to experience are more pronounced in São Paulo than in Rio Grande do Sul, Minas Gerais and Pernambuco, and much higher than in Ceará, for example. This means that those living in São Paulo throughout their life will experience a rise in relative earnings in the early phases of his life cycle (as compared to those living in Ceará, for example), followed by a drop in relative earnings later on.

The results of Figure 1 imply that if we follow a generation over time (and therefore over its life cycle) and compare mean earnings of those living in states with high returns to experience to those living in states with flat age-earnings profiles, we will first observe a divergence in their mean earnings, until both reach the peak of their earnings profile, and then a convergence process. Therefore, whether we will observe income convergence or divergence among cohorts living in different states will depend on the stage of their life cycle.

Figure 1: Returns to experience, 1997



To confirm that this is indeed the case, Figure 2 presents the results of traditional convergence regressions, run separately for each cohort. The results are quite striking. One can note that for the youngest cohorts (9 and 8) we observe divergence in income, whereas for the older cohorts the opposite occurs. Moreover, the speed of convergence rises continuously across generations. As the various generations live together at any moment in time, whether we will observe convergence or not in the aggregate will depend on the relative size of each cohort in each state. For example, as the population in both São Paulo and in Ceará are predominantly young, the demographic effect will act as a force conducting to divergence of their average incomes over time, as the population grow older. This process may even out-weight any economic tendencies leading to income convergence. Once the majority of the population reaches their forties, the reverse will occur. We expect that this effect will be even stronger when different countries are compared, for then the differences in the demographic composition will be higher, which will make the final convergence result even more dependent on the age structure and on the returns to experience prevailing in each of the countries involved in the study.

Figure 2: Convergence across cohorts

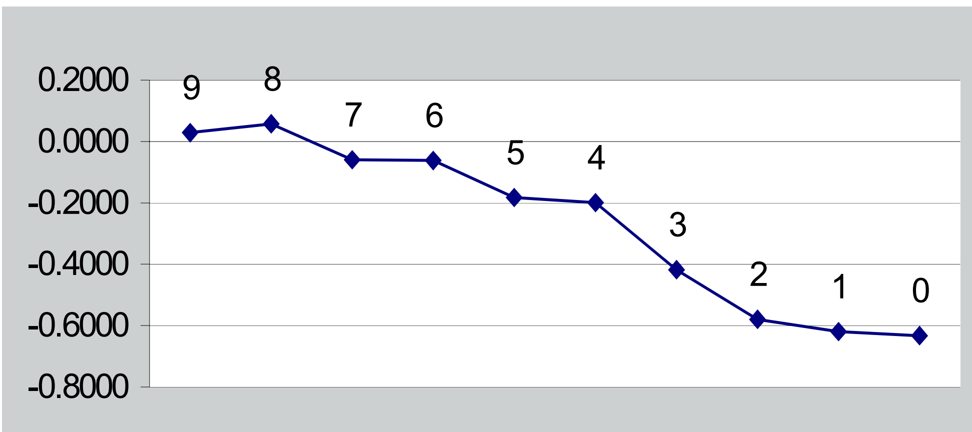


Figure 2 and the discussion above implies that the estimated convergence coefficients presented in Table 5 are in fact weighted averages of the population coefficient for each cohort, with weights given by the conditional variance of initial income at each cohort (see Angrist and Krueger, 1999 for a similar discussion in the context of returns to schooling).

7 Conclusions and future research

In this paper we emphasized the demographic effects involved in the traditional convergence literature. We used micro data, averaged up to cohort level, to investigate the impact of returns to experience and of the age structure of the population on the estimated convergence result. We started by showing that our household level data is compatible with the national accounts data, traditionally used in the literature, accounting for between 50 per cent and 80 per cent of GDP per capita, depending on the state. We also showed that the cross sectional convergence results using aggregated micro data yield similar results to using national accounts data, under different methods of aggregation.

The main results of the paper showed that the use of micro data provokes an increase in the speed of convergence, mainly because it incorporates convergence of income across different generations within the same state. Once we controlled for cohort effects, the results approached the aggregated ones. The use of panel data did not qualitatively alter the main results, despite the rise in the speed of convergence that occurred when the more frequent data were used. Moreover, the introduction of time and state dummies raised the speed of convergence considerably, as expected.

In explaining the differences in the results using micro as compared to more aggregate data, we emphasized the fact that the speed of convergence varies considerably across cohorts, so that the aggregate results depend a great deal on the composition of the population and on the returns to experience, that differ a lot across the States of Brazil. We speculated that this dependency will be even higher when comparing income growth across countries, with very different demographic characteristics.[7].

It is our hope that this paper will establish a new line of research in the convergence literature, and therefore there are several things we intend to do in the next steps of this research. First, we intend to simulate what would happen with the aggregate convergence results once we give different weights to each cohort within each state, to compute aggregate income, perhaps using the projections of the state population provided by the Brazilian Census Bureau. These weights can also be used directly in the micro regressions. We will also try and isolate the effects of returns to experience from the overall process of economic convergence, by aggregating the residuals of state specific regressions of earnings on age, instead of using earnings directly to compute average income at cohort and state level. We also intend to treat the returns to experience themselves as endogenous, by making them depend on the relative cohort sizes, as in Higgins and Williamson (1999) and instrument initial income. In order to do that, we will have to establish a framework to deal with migration, since it is clear that people move across states over their life cycle in a non-random way, and this affects the behavior of the returns to experience.

References

Angrist. J.D. and A.B. Krueger (1999) 'Empirical Strategies in Labour Economics', in O. Ashenfelter and D. Cards (eds) *Handbook of Labour Economics*, North Holland: Amsterdam.

Attanasio, O and M. Browning (1995) 'Consumption over the Life Cycle and the Business Cycle', *American Economic Review* 85:1,118-36.

[7] Although it is important to note that the key variables under study may not be compatible internationally, and that it may difficult to control for differences in institutions and culture, that may be related to the income or growth.

Attanasio, O., and G. Weber (1993) 'Consumption Growth, the Interest Rate and Aggregation', *Review of Economic Studies* 60:631-49

Azzoni, C. (2001) 'Economic Growth and Regional Income Inequalities in Brazil', *Annals of Regional Science* 35(1):133-52.

Azzoni, C., N. Menezes-Filho, T. Menezes and R. Silveira-Neto (2000) 'Geography and Regional Income Inequality in Brazil', *Research Network Working Paper R-395, Inter-American Development Bank, Research Department.*

Azzoni, C. and L.M. Servo (2002) Education, Cost of Living and Regional Labour Income Inequality in Brazil', *Papers in Regional Science* 81(2):57-75.

Baer, W. (2002) *The Brazilian Economy*, Praeger: Westport CT.

Barro, R. and X. Sala-I-Martin (1995) *Economic Growth*, McGraw-Hill: New York.

Bils, M. and P.J. Klenow (2000) *Does Schooling Cause Growth*?, *American Economic Review* 905:1,160-83.

Browning, M., A. Deaton and M. Irish (1985) 'A Profitable Approach to Labour Supply and Commodity Demands Over the Life Cycle', *Econometrica* 53(3):503-44

Ferreira, A. (2000) 'Convergence in Brazil: Recent Trends and Long-Run Prospects', *Journal of Applied Economics* 3(4):479-89.

Ferreira A.H. and C.C. Diniz (1995) 'Convergencia entre las rentas per capita estaduales en Brasil', *EURE-Revista Latioamericana de Estudios Urbano Regionales* XXI(62).

Freeman, R. (1979) 'The Effects of Demographic Factors on Age-Earnings Profiles', *Journal of Human Resources* 14:289-318.

Hanushek, E.A. and D.D. Kimko (2000) 'Schooling, Labour-Force Quality and the Growth of Nations', *American Economic Review* 90(5):1,184-208.

Higgins, M. and J.G. Williamson (1997) Age Structure Dynamics in Asia and Dependence on Foreign Capital, *Population Development Review* 23:(2):261-93

Higgins, M. and J.G. Williamson (1999) Explaining Inequality the World Round: Cohort Size, Kuznetz Curves, and Openness', *NBER Working Papers* 7,224, July National Bureau of Economic Research: Cambridge MA.

Islam, N. (1995) 'Growth Empirics: A Panel Data Approach', *Quarterly Journal of Economics* 110(4):1,127-70.

Jalan, J. and M. Ravallion (2002) 'Geographic Poverty Traps? A Micro Model of Consumption Growth in Rural China', *Journal of Applied Econometrics* 17(4):329-46.

Mankiw, G.N., D. Romer and D.N. Weil (1992) A Contribution to the Empirics of Economic Growth, *Quarterly Journal of Economics* 107(2):407-37.

Ravallion, M. (1998) ‘Poor Areas’, in A. Ullah and D. Giles (eds) *Handbook of Applied Economic Statistics,* Marcel Dekkar: New York.

Schwartsman, A. (1996) ‘Convergence Across Brazilian States’, *IPE Discussion Papers* 02/96, IPE-Universidade de São Paulo: São Paulo.

Temple, J. (1999) The New Growth Evidence, *Journal of Economic Literature* XXXVII:112-56.

Welch, F. (1979) ‘Effects of Cohort Size on Earnings: the Baby Boom Babies Financial Boost’, *Journal of Political Economy* 87:65-97

Zini A.A. (1998) ‘Regional Income Convergence in Brazil and Its Socioeconomic Determinants’, *Economia Aplicada* 2(2):383-411.