Thirty Years of Earnings Inequality in Brazil (1981 to 2011): Age, Period and Cohort Effects.

(Paper apresentado no XVI Congresso Brasileiro de Sociologia)

Rogério Jerônimo Barbosa (CEM/Cebrap, USP)
Carlos A. Costa-Ribeiro (IESP-UERJ)
Flavio Carvalhaes (IESP-UERJ)

Draft 2

Setembro, 2013

1 - Introduction

It is widely known that Brazilian income inequality indexes is among the highest in the world – the eight higher in the 1990's. During the 1980's and 1990's, for example, the average Gini index for family *per capita* income was 80% larger than the average for OCDE countries and 20% larger than the average for other Latin American countries (Ferreira et al, 2006). By the end of the 1980's in Latin America, only Guatemala and Honduras had a scenario similar to Brazil's (Barros et al, 1997). Fortunately, after a period of steady growth (which lasted until 1989), those indexes started to fall and the picture started to change. Specifically during the last decade, income inequality entered a rapid declining path. The present indexes are in the lowest levels ever registered in the country.

Given the high quality of the scholarship on Brazilian income inequality, we would like to contribute adding the dimension of age, period and cohort effects that were previously discussed in the literature (Menezes-Filho, Fernandes e Picchetti, 2006), but never properly analyzed. For that we will use new models recently proposed for studying income levels and inequality across those three time-related dimensions (Yang & Land, 2006; Yang & Land, 2008; Zheng, Yang & Land, 2011). Amongst the changes that affected Brazilian society in the last thirty years, population changes play an important role determining trends of income inequality. The country has experienced an intense and fast demographic transition that altered its population composition (Chaimowicz, 1997). We will focus on how the trends of income inequality are related to the population dynamics. Specially, we will analyze how cohort and periods interacted during the last thirty years and how they are related to the rise and fall of income inequality in the country.

While period is important because macro-economic changes affect the whole working population at that moment, cohort seems to be relevant in Brazil because each generation experienced very different structures of opportunity related to the fast social changes taking place in the second half of the 20th century. Among those changes, it is important to highlight how the supply for educational opportunities affected each cohort, boosting intergenerational income inequalities based on differentials in labor market returns to education. In other words, effects on a certain period can be related to aspects of people's characteristics that were achieved in earlier phases of the life cycle. Another important issue in our analysis is the comparison between men and women. This will allow us to understand, on one side, the levels of gender earnings inequality and how they relate to age, period and cohort, and, on the other side, how the incremental participation of women in the labor market affect the levels of inequality and of mean earnings of men.

The paper is divided in seven sections including this introduction. The second presents some basic characteristics of income inequality in Brazil and the importance of including the age, period and cohort dimensions to the analysis given the history of social change in the country. Sections three and four present our modeling strategy and data. Parts five and six present the results for the whole population and split by gender respectively. The final part sums up the main conclusions.

2 - Brazilian Inequality Studies: Thirty Years of Changes and the Effects of Age, Period and Cohort on Income Inequality

The academic debate on income inequality in Brazil started in the early 1970's, when reliable data became available to inform the pioneer works of Hoffman and Duarte (1972), Fishlow (1972), and Langoni (1973). These authors tried to understand the mechanisms that affected inequality. Fishlow (1972) highlighted the importance of labor market and political factors linked to the military government of 1964. According to him, the growth of inequality during the 1960´s was related to repressive mechanisms created by the Military Dictatorship (1964 to 1984) to control the labor force, on one side, and to monetary policies which allowed for the country's growth at the expense of high inflation rates, on the other side. In contrast, Langoni (1973) argued that in a context of accelerated industrial and economic growth combined with high educational heterogeneity, those with higher levels of education received abnormally high returns to their qualifications. In other words, given the unequal returns to educational levels (which were increasing rapidly) the economic development process introduced economic heterogeneity based on the population's unequal educational achievement. In fact, with the steady economic growth demanding for qualified work force and there existing low percentages of people with higher (or adequate) educational degrees, the returns to education were extremely high, leading to increasing inequality during that period.

The 1970's debate ended with Langoni's position as being considered the correct one. Since these initial studies, an extensive literature has been produced confirming that the steady and massive economic development during the 1960's and 1970's was followed by an expressive take off of income inequality. Between 1960 and 1980, the proportion of income of the richest 20% of the population raised from 54% to 63% of the total, while the appropriation of income of the poorest 40% of the population dropped from 12% to 10% of the total (Barros et al, 1997). Income inequality continued to grow during the 1980's – when the country experienced several cycles of economic crisis. All indexes peak in 1989, and from this year onward there has been a significant decline, with a more expressive drop from 2001 onward (Ferreira et al, 2006). Studies dedicated to the analysis of Brazilian inequality trends are unequivocal in pointing that it has diminished independently of what variable is analyzed: family income (total or *per capita*), total individual income, or earnings (salary or wage). Among these, the latter is of special interest to us because of its heavy weight on inequality as a whole[1]. In this paper we will analyze inequality in earnings (monthly salary) from 1981 to 2011[2].

From the many aspects related to the trend of rising income inequality until 1994 and declining afterwards, we will highlight three in this paper: period, cohort and age.[3] Disentangling these three effects is of special importance in Brazil, given the country's history of fast social change from the 1960's onward. Therefore, in order to better understand period and cohort effects it is necessary to briefly explain a bit more some stylized characteristics of the Brazilian economic and demographic development during the last century.

From the mid 1960's to the end of 1970's, Brazil experienced intense economic growth, a period known as the "Brazilian economic miracle". This phase of fast development and industrialization happened simultaneously to an also fast rural-urban transition – until 1960, more than half of the population lived in rural areas, while by

---

[1] The labor market earnings account for 76% of the total family income (Barros et al, 2006b).

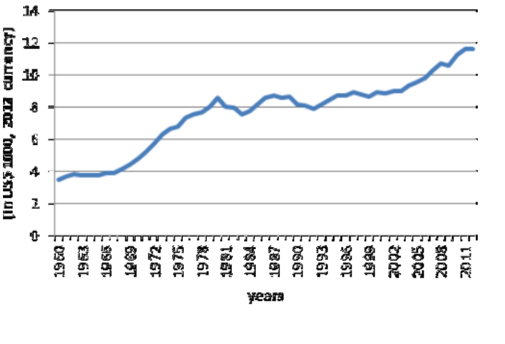
[2] In Brazil, employed people most commonly get paid monthly (not hourly). It is not usual, for employees or even for employers, to calculate earnings in terms of hours. Additionally it is important to mention that the country labor market has high levels of informality (unregulated labor relationships, normally without a contract). The most common kind of informal worker are the so called "self-employed", who are generally low paid, usually allocated in the lower service sectors and do not contribute to or receive any kind of social insurance or welfare benefits. Those "self-employed" normally do not work full time – and it seems not appropriate to interpret this fact as a consequence of a rational economic decision. In other words, the amount of time spent working is correlated to the formality or informality of jobs. We could say there is a kind of selection bias (linked to the duality between formal and informal) that is expressed in the average time a person works. So the monthly salary can be considered a net result of two inequalities: a) one related to how much people receives by hour; b) another related to how many hours people can/may work.

[3] Other important features of the drop since 1994 would be a relative homogenization of the age composition of the population, which helped to diminish differences in returns to experience (Barros et al, 2006c) and a convergence between rural and urban wages (Ferreira et al, 2006). More recently, starting in 2002, there is also the implementation of "Bolsa Família", a direct cash transfer program (very similar to Mexico's Oportunidades) that also had a significant effect on the reduction of inequality, specially trough poverty alleviation (Soares, 2006; Barros, et al, 2006b; Hoffmann, 2007).

2000 this proportion was less than 20%. In contrast, during the 1980's and 1990's there was practically no economic growth (see Figure 1) – so those twenty years were known as the "lost two decades". This period of economic stagnation began with the debt crisis of the early 1980's, which affected the whole region of Latin America. In addition, during the 1980's until 1994 the country experienced extremely high levels of inflation (see Figures 3 and 4), partly as a consequence of high public expenditures, which were used to finance the previous economic development. As a consequence of a well succeeded monetary plan implemented in 1994 hyperinflation was finally controlled. Still, the economy only started to grow again during the 2000's – modestly, though.

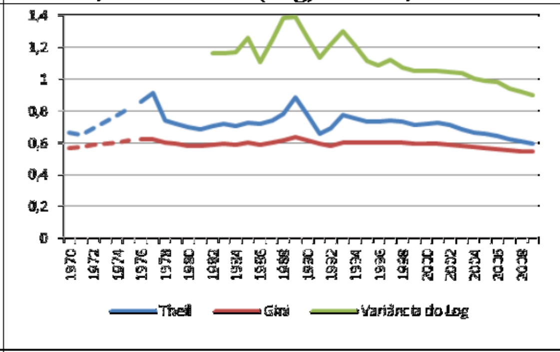
We will consider those economic conjunctures and processes as *period effects*, that is, changes that occur or affect all people living in a specific context at the same time (whatever the age group they belong). According to Yang Yang, a period effect "reflects changes in social, historical and epidemiological conditions that are unique to a time period [and] that affect all living conditions regardless of age or life stage" (2011, p. 18). In other words, it is possible to identify a time component that is independent of (orthogonal to) age and cohort.

igure 1 – GDP *per capita* – Brazil, 1960-2009



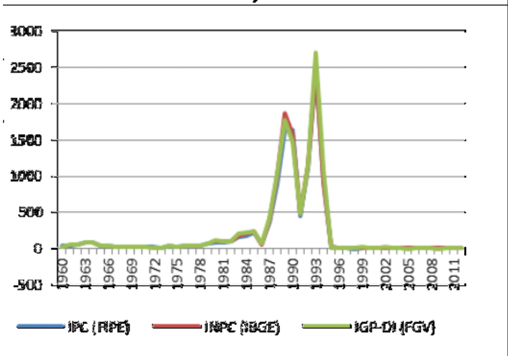
: IPEA Data.

Figure 2 – Labor Income Inequality: Gini, Theil and Var(Log) – Brazil, 1970-2009



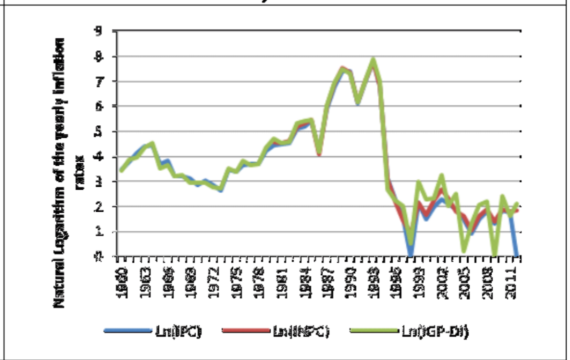
Fontes: 1970: Censo 1970 (Langoni, 2005); 1971-1975: Estimação por interpolação linear; 1976-2009: PNADs, elaboração própria.

igure 3 – Inflation Rates according to three indexes – Brazil, 1960-2012



: IPEA Data.

Figure 4 – Logarithm of the inflation Rates – Brazil, 1960-2012



Fonte: IPEA Data, elaboração própria.

During the whole period of fast economic growth, the educational system was extremely underdeveloped – illiterate rates in Brazil were: 39% in 1960, 25% in 1980 and 13% in 2000. In fact, the most significant changes in the educational composition of the Brazilian population began to happen only from 1980 onward – in spite of the economic downturn. As the majority of the working population were always composed by people aged 25 or more (and these older cohorts were exposed to limited educational opportunities), low educational attainment lasted long as a strong feature of the Brazilian labor marked. In fact, the income returns to higher levels of education (high school or college degree) remained very high.

Only younger cohorts were characterized by lower educational heterogeneity. With more people having middle and higher school degrees, the observed educational wage premium decreased since the mid 1990's and especially during the 2000's. In other words, inequality in the period studied was also related to changes that happened along the composition of birth cohorts represented in each year since 1981, but who attended school and entered the labor market since earlier decades (Ferreira et al, 1998, 2001 e 2006). According to Barros et al (2006b), the educational equalization within the work force and the drop in the returns to human capital investment accounts approximately for 40% of the recent drop in the labor earnings inequality between 2001 and 2005.

Although we do not analyze surveys from periods previous to 1981, we can observe long-term income and inequality level differences looking at generational composition of the work force. Assuming that the educational attainment process typically ends at the age 25 or earlier, different results for people older than that are related earlier structures of educational supply. A cohort effect can be understood exactly as a formative experience (Ryder, 1965): a set of events or process that shape life conditions of individuals born in the same period. These people share not only the moment of birth, but also a continuous exposure to the same historical opportunities and events – so they can be regarded as a group.

In sum, there are many good historical reasons to believe that in addition to a period effect on trends of income inequality, there are significant cohort effects that must be untangled. This is precisely one of the contributions we want to bring to literature with the present paper. Previous research has pointed to the importance of untangling period and cohort effects on income inequality (Menezes-Filho, Fernandes e Picchetti, 2006), but did not analyze in a proper way these two aspects.

Whereas period stand for "instantaneous" or conjectural effects and cohort effects trace back long duration inequalities; age effects represent regularities in the individual life courses. It is expected, for instance, for one person to increase her own earnings during her life, as a consequence of the experience in the labor market, human capital acquisition and other institutional and social factors. The income, however, is expected to decrease as people age. Not only income levels, but also inequalities are linked to the life course.

Another very important issue that we will explore in our analysis is the impact of the increasing entrance of women in the labor market in the period analyzed. While in 1980 only 27% of the workers were female, in 2010 practically half of the workers were females. Therefore, the entrance of women in the labor market is closely related to trends in income inequality (we have to explore more this issue).

In addition to specifying age, period and cohort effects, our analysis in this paper also contributes to the literature on income inequality in Brazil because we use models that distinguish inequality between groups from inequality within groups. More specifically we will analyze trends in inequality between and within each period (survey years, from 1981 to 2011). The same distinction will be made for cohorts (born between 1916 and 1986) and age (people having 25 and 65 years old). “Between effects” represent differences between groups and also overall or average income levels. Inequality within groups are expressed by the variance within each category and can also be understood as the risk to which each group is exposed to.

3 – Modeling Strategy

As we discussed above, income inequality trends are closely related to temporal dimensions: economic cycles, the demographic structure, intergenerational and life course dynamics. The challenge we face is modeling the inequality dynamics accounting for all these factors. Following Zheng, Yang and Land (2011), we adopt an Age-Period-Cohort (APC) regression, which can estimate the determinants of both the income level and inequality.

Age-Period-Cohort (APC) models (Mason et al., 1973; Yang, Fu & Land, 2004 Yang et al., 2008), which are appropriate to disentangle those three kinds of “time effects”, were usually restricted to aggregated data. Yang and Land (2006), however, developed an APC approach specific for individual level data, which do not incur in the “identification problem”[4]. They propose the use of a hierarchical mixed model (HAPC), with period and cohorts effects[5] in the second level (group-level) – and age (being or not regarded as a polynomial including age squared, for instance) in the first one, being considered as an individual-level variable. They present a random effects model in which the intercept is cross-classified by period and cohort (Cross-classified Random Effects Model – CCREM). But Yang and Land (2008) showed that the period and

[4] Methodologically it is very hard to distinguish the three effects because there is an exact linear dependency: period = cohort + age. This problem means that APC models cannot be mathematically identified. A long tradition of studies in sociology, demography and epidemiology has been attempting a solution for this identification problem (Fienberg e Mason, 1985; Mason et al, 1973;O´Brien, R. 2000). Although there is no unequivocal solution, there is a recent methodology that has been showing very useful in many applications (Yang, Fu & Land, 2004).

[5] In repeated cross-section surveys, age, period and cohort are respondent characteristics or variables. Making use of this kind of individual level data, Yang and Land (2006) regarded period and cohorts as groups to which the respondents belong. By grouping periods or cohorts, it is possible to break the identification problem (individuals of several ages can be part of the same period and cohort), allowing finite-valued solutions to a regression model. It is possible to make up various differential groupings of period and cohort. Here we follow the same strategy of those authors, grouping cohorts into 5-year groups.

cohort effects can be treated both as fixed or random effects, since the assumptions of the chosen model are respected.

In order to model inequality, Western and Bloome (2009) presented what they called Variance Function Regression Models (VFR). According to them, regression-based studies of inequality take into account only the "between-group differences" (expressed by the coefficients). They argue that the within-group inequality is commonly disregarded as being unexplained variation and, therefore, not being treated as an object of sociological interest. Therefore, they intended to show that the residual heterogeneity has been overlooked and may be important in several substantive ways. A group can be structurally more insecure or unequal than another, for instance. Western and Bloome (2009) present a two-folded strategy of regression modeling that includes covariates for both the mean and variance of a dependent variable. First, they estimate a linear regression. In this model, if categorical variables are groups, their coefficients represent the difference between the categories' means. Second, they save the unstandardized residuals of the first model, raise them to the square and run a Gamma regression model on them[6]. In this second model, the residual squared represents the variance, and is considered the target of analysis. In other words, this two-step modeling does not assume constant variance as in a context of linear regression analysis. The variance function model is intrinsically heteroscedastic, once the variance can be treated as being depended on covariates. Once each model is estimated separately, the standard errors of both regressions are incorrect: the estimates of the Gamma regression do not take into account the uncertainty in the linear regression coefficients, and the last are inefficient once they ignore the heteroscedasticity. Western and Bloome (2009) point out that maximum likelihood estimates can be obtained by iterating these two steps – leading to unbiased standard errors and efficient estimates[7].

Zheng, Yang and Land (2011) proposed to integrate the HAPC to the Variance Function Regression, allowing for studying the effects of those three different temporal

---

[6] If the errors are normally distributed, the squared residuals will approach a gamma distribution. The gamma regression is a kind of generalized linear model (GLM) tailored for just-positive and right skewed variables. The use of a common OLS can lead to negative predicted values (which cannot exist for the squared residuals) and highly underestimated standard errors (making hypothesis tests biased).

[7] This two-step iterative process is executed through a weighted least squares estimation. The predicted values of the gamma regression (estimated variance: $\sigma 2$) are saved into a variable. Then the first step is executed again, using $1/\hat{\sigma}^2$ as weight (in our case, $\frac{1}{\hat{\sigma}^2} \times [original\ analitic\ survey\ weights]$, once the survey sample design already needed weighting). This procedure is repeated until the maximum likelihood statistic achieve convergence (details are given by Western and Bloome, 2009).

dimensions on inequalities. The goal of the analysis is to simultaneously decompose the between-group (variations in the conditional mean) and within-group inequality (variations in the conditional variance) into age, period, and cohort components. In their original study, the authors applied their model for analyzing self-reported health inequality.

Our modeling has some differences in relation to Zheng, Yang and Land (2011). We estimate the models using the two-step iterative process. So we got maximum likelihood estimates (they use a Restricted Maximum Likelihood Estimator). Western and Bloome (2009) showed that in big samples, different estimation methods tend to converge. In our case, as indicated below, we have an extremely large sample size.

Summing up, our modeling strategy is based on the use of a Hierarchical Age-Period-Cohort Model integrated to a Variance Function Regression/heteroscedastic regression (HAPC-VRF/HR), with cross-classified random effects (CCREM). In this preliminary work, APC effects will be treated as changing just the intercepts (random intercept model), that is, we will use a simple model with no interactions with the other covariates. This means that all covariates effects are treated as being constant over time[8]. Our two regression models (for the mean and the variance) are given by the following hierarchical equations below:

Mean Model

Level-1 / Within-Cell Model: individual level

$$Y_i = \beta_{0jk} + \sum_{p=1}^{P} \beta_p X_{pijk} + \beta_a Age_{ijk} + \beta_{a2}(Age)^2{}_{ijk} + \varepsilon_{ijk},\ \varepsilon_{ijk} \sim N(0, \sigma_i^2) \tag{1.1}$$

Level-2 / Between-Cell (random intercept) model: period and cohort level

$$\begin{gathered} \beta_{0jk} = \gamma_0 + u_{0j} + v_{0k}, \\ u_{0j} \sim N(0, \tau_u), \\ v_{0k} \sim N(0, \tau_v) \end{gathered} \tag{1.2}$$

Variance Model

Level-1 / Within-Cell Model: individual level

$$\ln(\sigma_i^2) = \lambda_{0jk} + \sum_{p=1}^{P} \lambda_p X_{pijk} + \lambda_a Age_{ijk} + \lambda_{a2}(Age)^2{}_{ijk} \tag{2.1}$$

Level-2 / Between-Cell (random intercept) model: period and cohort level

$$\begin{gathered} \lambda_{0jk} = \pi_0 + \omega_{0j} + \varphi_{0k}, \\ \omega_{0j} \sim N(0, \psi_\omega), \\ \varphi_{0k} \sim N(0, \psi_\varphi) \end{gathered} \tag{2.2}$$

---

[8] Surely this brings limitation to our analysis. But we are making tests and model sophistication including interactive terms, different estimators (Restricted Maximum Lihelihood), and another modeling strategies.

For $i = 1, 2, \ldots, n_{jk}$ individuals within cohort *j* and period *k*;
β_p= level-1 fixed effects for the mean regression for each *p* covariate
β_{0jk}=level-1 random intercept in the mean regression
γ_0= level-2 intercept in the mean regression
u_{0j}= level-2 cohort random effect in the mean regression (assumed as normally distributed)
v_{0k}= level-2 period random effect in the mean regression (assumed as normally distributed)
ε_{ijk}= heterokedastic conditional normally distributed individual-level residuals

λ_p= level 1 fixed effects for the variance regression for each *p* covariate
λ_{0jk}=level-1 random intercept in the variance regression
π_0= level-2 intercept in the variance regression
ω_{0j}= level-2 cohort random effect in the variance regression (assumed as normally distributed)
φ_{0k}= level-2 period random effect in the variance regression (assumed as normally distributed)

Both hieraquical linear equations can be presented as combined mixed-effects models:

Mean and variance combined models

$$Y_i = \gamma_0 + \sum_{p=1}^{P} \beta_p X_{pijk} + \beta_a Age_{ijk} + \beta_{a2}(Age)^2{}_{ijk} + u_{0j} + v_{0k} + \varepsilon_{ijk} \tag{3.1}$$

$$\ln(\sigma_i^2) = \pi_0 + \sum_{p=1}^{P} \lambda_p X_{pijk} + \lambda_a Age_{ijk} + \lambda_{a2}(Age)^2{}_{ijk} + \omega_{0j} + \varphi_{0k} \tag{3.2}$$

The two expressions above have the same independent variables. Y_i stands for our dependent variable in the linear regression of the first step; σ_i^2 is the squared residuals which will be used in the second step, the Gamma Regression[9] in order to estimate the group-dependent heterocedastic variance. X_p represent a set of individual-level covariates. Age and age squared are continuous and also individual-level variables. The random effects of period and cohort cross-classify and change the intercept. In the following section, we explain the dependent and independent variables that will be used in these equations.

4 – Data and variables

In this paper we use the main Brazilian data source for labor and income dynamics: Pesquisa Nacional por Amostragem de Domicilios (PNAD)[10]. We have

[9] It is importante to notice that the natural logarithm in the Gamma regression is part of the GLM link function. One does not need to log the squared residuals before using it as a input in the second one.

[10] The PNAD is a national household survey by Instituto Brasileiro de Geografia e Estatística (IBGE), a governmental official bureau of Statistics. It was first collected in 1967 – quarterly, but not yet in a national scale. From 1971 on, it runs yearly. The sample scope and size became more standardized since 1981. Therefore we use the PNAD series from 1981 to 2011 (which is the last year available yet). The PNAD was not collected in Census years (1970, 1980, 1991, 2000 and 2010), in 1974 and 1975 (when there occurred another big sample survey by IBGE) and in 1994 (due to political and administrative concerns). As a consequence those of these years included in our analytical time interval will be missing

selected a sub-sample for each year including just individuals (men and women) from 25 to 65 years old, with positive working hours and income at the moment they were surveyed. This procedure gave us a total sample size of 2,806,358 cases (considering all the years, from 1981 to 2011). Our dependent variable is the logarithm of monthly earnings, which was corrected for inflation using the standard deflation procedure adopted in Brazil (Courseiul & Fogel, 2002) which brings the currencies to the latest year survey values. We also use a series of relevant independent variables besides age, period and cohort.

The Brazilian labor market, as many others in Latin America, include a formal and an informal sector (reference category in our model). Workers in the formal sector are mainly those having unemployment and pension benefits, while those in the informal sector do not count with these benefits. In addition to the formal versus informal division, we included the occupation position of individuals in our analysis: employees (reference category), employers and self-employed. We also control for race using a dichotomous division. As usual in the Brazilian social sciences literature we use a dummy variable which divides race groups into white (whites and Asians) and non-white (blacks, mulattos and indigenous people – reference category). Whites have, in general, privilege in relation to non-whites in terms of educational attainment and labor market outcomes. Once there is high and a long term established regional inequalities among, we also included regions dummy variables (Southeast is the reference category). As they are fixed effects, they control for all the unobserved heterogeneity between regions. We also include a continuous variable for weekly working hours.

It is widely known that education is one of the most important determinants of income level. The literature on income inequality indicates that at least 40% of the recent drop in earnings inequality since 2001 is due to declining returns to education (cf. MENEZES-FILHO, ANO). Once education and the log income have a linear relation, we used a continuous variable for years of study. Because mean educational achievement varies strongly through generations, our education variable is cohort centered. So it measured how many more or less years of schooling a individual has than the average for his cohort. We intend to further explore the interaction between education and period and cohort effects (including a random slope term for years of

in our analysis. The PNAD surveys are very large (the number of cases is around 400.000 individuals per year – or 150.000 households) and quite reliable for demographic, labor market and educational information.

schooling) when expanding the present paper. But at the moment we are only controlling for this effect and not explicitly modeling if the effect changes across the temporal dimensions.

We introduced age in the model as a continuous grand-mean centered variable in linear and squared format. Period is defined by the survey years (ranging from 1981 to 2011 – excluding 1991, 1994, 2000 and 2010, for which there is no data). As described above we already know that earnings inequality increase from 1981 until 1989, fluctuates from 1989 to 1994, declines from 1995 to 2000, and declines even more from 2001 to 2011.

Finally, cohort is defined by five-year birth cohorts in the following way: (1) 1916-1920, (2) 1921-1925, (3) 1926-1930, (4) 1931-1935, (5) 1936-1940, (6) 1941-1945, (7) 1946-1950, (8) 1951-1955, (9) 1956-1960, (10) 1961-1965, (11) 1966-1970, (12) 1971-1975, (13) 1976-1980, and (14) 1981-1985 and (15) 1986. We also analyze data for men and women in two different ways. First, we use gender as an independent variable in a pooled model and, then we estimate separated models for men and women (assuming full interaction between all the convariantes and gender). This second strategy allows us to understand and compare the trend of income inequality taking into account the disparity of income between men and women. The period analyzed is also one of growing participation of women in the labor market. The entrance of women changes the competition between men and women in the labor market and probably explains part of the general trend.

5 – Results: Variation in Mean Earnings and Earnings Inequality by Age, Period and Cohort, 1981 to 2011.

Figures 5 to 10 display the observed values (not controlled for the explanatory covariates described above) for the mean and the variance of the logarithmic income. These results do not disentangle age, period and cohort– so they remain highly correlated and it is not possible to distinguish each component specifically. Nevertheless, they will be taken as a baseline in order to assess the explanatory power of our models.

Age, period and cohort effects – Observed values

Mean (R$)

Figure 5 - Age

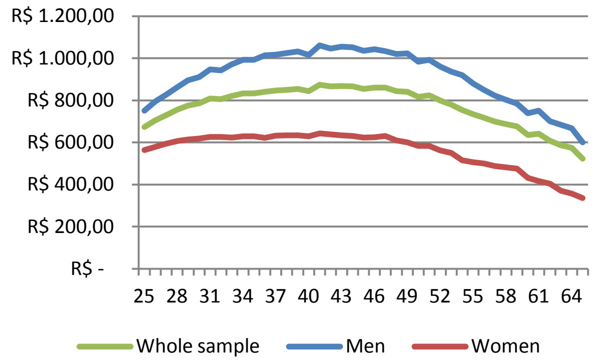


Figure 6 - Period

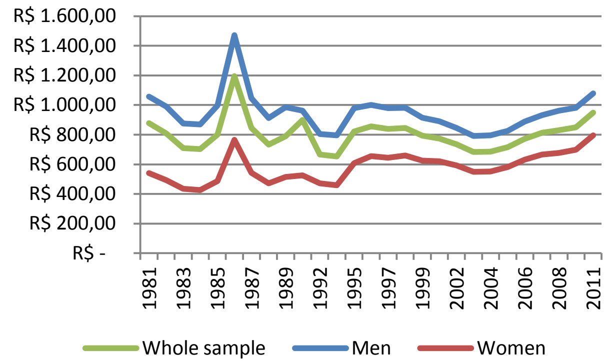
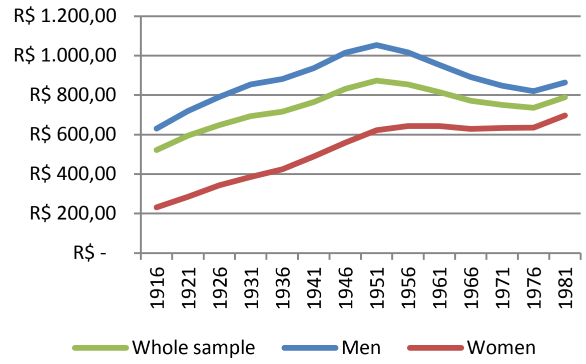


Figure 7 - Cohort



Variance (ln[income])

Figure 8 - Age

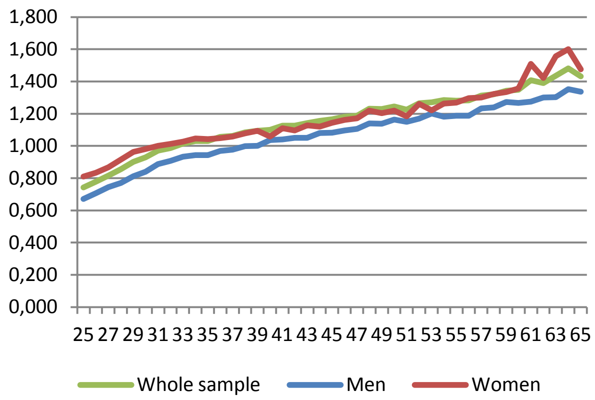


Figure 9 - Period

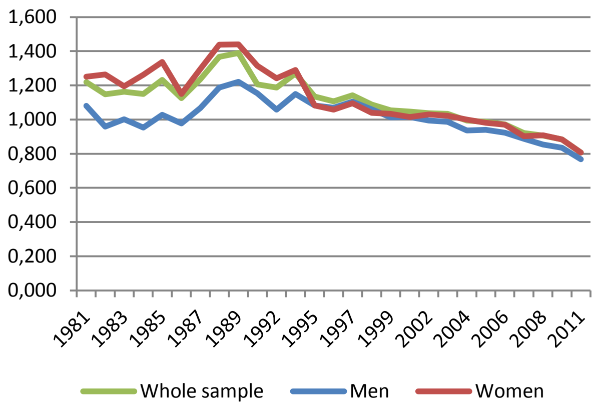


Figure 10 - Cohort

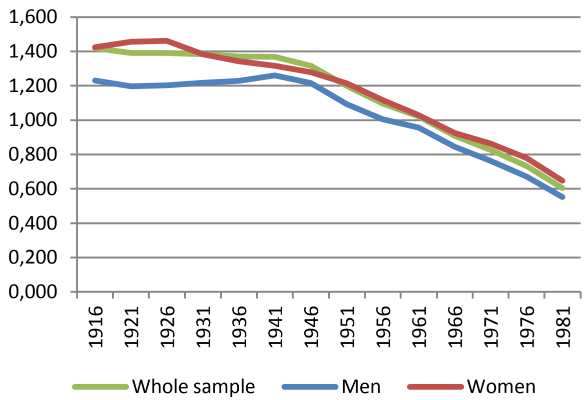


Table 1 presents parameter estimates and model fit for the HAPC-VFR/HR Models of Log monthly earnings for the whole sample. The linear regression column presents results for the first stage regression which estimates the mean income for each group (between/across groups effect). The Gamma Regression columns present results for the second-stage, which estimates the log income variance – understood here as a measure of income inequality within groups.

The individual fixed effects of the linear regression are consistent with the widely known behavior of the selected explanatory variables towards income. Each additional year of education leads to a 10.7% increase in the expected income level. Working more hours weekly also raises income (each additional hour is associated with 1% point increase). Being in the formal labor market is associated with approximate 40% increase in the expected income. Self-employed earn 5% more than employees and employers earn around 70% more. Regarding regions, in the Southeast (reference category) we find the biggest expected income – all the dummy variables indicating the other regions have negative effect. The lowest salaries are found in Northeast (40% lower than in Southeast). Regarding gender, males have a income 44,8% bigger than females.

The fixed effects for age, period and cohort are presented as a graph. For age, we calculated the fitted values keeping all the other covariates at their mean, changing just the values for age and age squared. The same was done for period and cohort random effects. We applied the exponential in order to change the results measure (log income) to the actual currency (R$).

Figure 11 points that age effects are curvilinear (negative quadratic). The monthly income increases linearly until the age of 52 and decreases thereafter. Figure 12 represent the expected period effects on the mean income. It follows the same general trends than the observed univariate mean income across the years. It shows fluctuation during the 1980s. After a great drop caused by the 1981-1983 crisis, there was an increasing trend till 1986. Thereon it followed a steady and constant decline until 1993 (the year in which there were the highest inflation rates in Brazil's recent history). In the middle of 1994, inflation was controlled (as we discussed above). The consequences of this fact are quite identifiable in the year of 1995. The mean income increased and stayed relatively high until 1998, when there was a major devaluation of the Brazilian currency. In that year, there was the Russian financial crisis (affecting the entire world) and the Brazilian macroeconomic monetary policy changed from fixed exchange rate to floating exchange rate. So after a brief interval of stability, the mean income declined from 1999 to 2003. It is important to mention that during these same years unemployment and labor market informality achieved the highest level ever registered. When the economic growth regained its pace from 2004 onwards, the average income level increased steadily accompanying this movement.

Age, period and cohort effects – Fitted Values

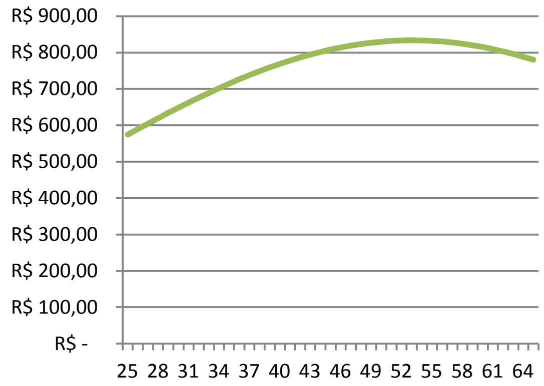
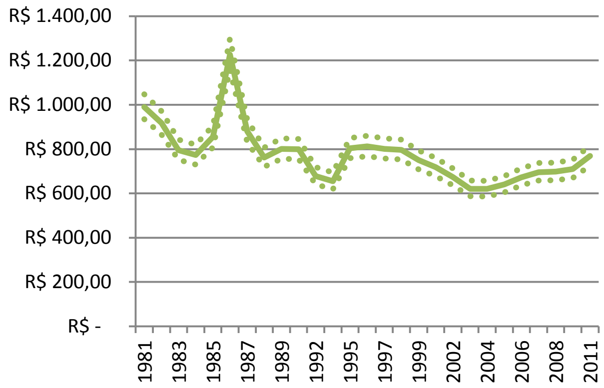
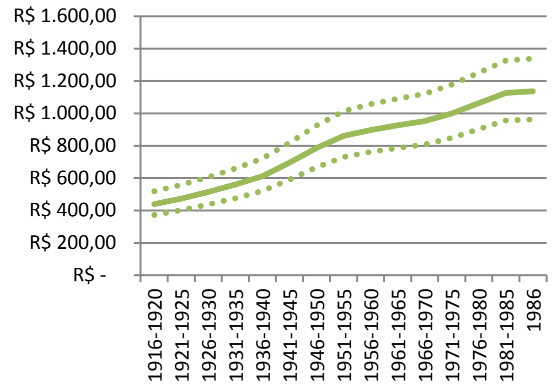
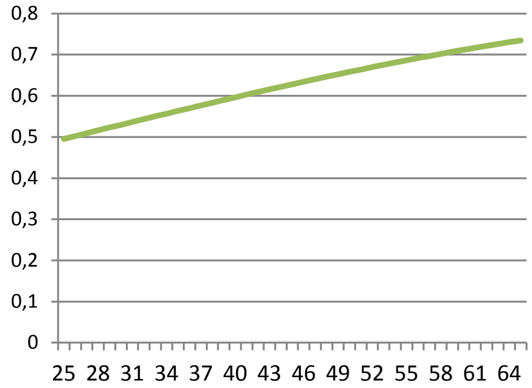
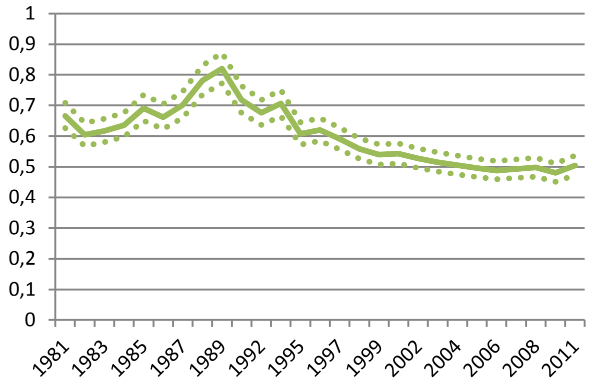
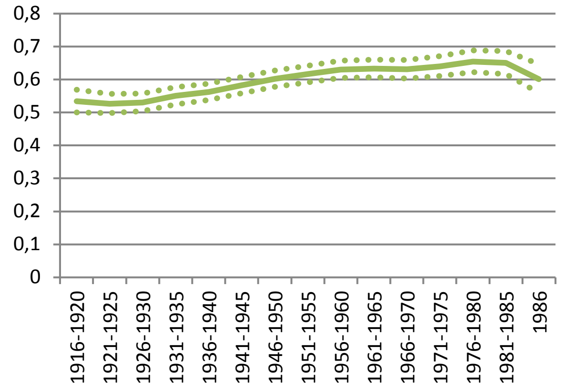
<table>
<tr><th></th><th>Figure 11 - Age</th><th>Figure 12 - Period</th><th>Figure 13 - Cohort</th></tr>
<tr><td>Mean (R$)</td><td>
</td><td>
</td><td>
</td></tr>
<tr><th></th><th>Figure 14 - Age</th><th>Figure 15 - Period</th><th>Figure 16 - Cohort</th></tr>
<tr><td>Variance (ln[income])</td><td>
</td><td>
</td><td>
</td></tr>
</table>

Figure 13 presents the trend on mean earnings across birth cohorts. From the oldest cohort to the one born between 1951 and 1955 mean log earnings increased steadily. This is very clearly related to the history of economic development in Brazil. Those born between 1916 and 1955 were around their twenties during all the industrialization and development process which took place in Brazil between years 1930 and 1975. Once the mid 1970's economic growth were the highest ever registered, it seems very plausible that those born in 1951-1955 were exactly the ones who most benefited of it – they were at the right period with the right age to reach for the opportunities. After this point the economy entered a path of deceleration which could preannounce the 1980's crisis and stagnation. It seems the cohort trend is quite related to all the economic opportunities which were being created since the early stages of industrialization and that came to a halt at the end of the 1970's. There was no significant increase in the expected income for those born between 1955 and 1975 – these people are exactly the ones who experienced their entrance in the labor market in the midst of a highly unfavorable scenario. The acceleration of cohort effects after 1975 shows that the recent period of growth is starting to bring positive consequences for the young cohorts.

It is a very interesting fact that cohort effects trend is monotonically positive. It means that the expected income for younger generations is always bigger than for the older ones. Economic crises and stagnation didn't lead to a decline in the income level. This result is probably associated to the fact that the economic development process meant a continuous monetarization, that is, it established monetary earnings as the dominant form of economic return – instead of goods, for instance. New cohorts have more propensities to be exposed to urban labor markets where monetary salaries are main or exclusive reward for working.

The column for the Gamma Regression in Table 1 shows how the covariates affect log earnings inequality. Estimated within-group log earnings disparities increases as one achieves more years of education, but decreases as he works more hours per week. In general, those with lower educational levels have smaller earnings that are more similar within the same category of educational attainment – the same for the quantity of hours spent working weekly. Men are more unequal than women – we will approach this issue with more details bellow.

Concerning the labor market variables (formality and position), one way of seeing the inequality within each category is by taking them also a measure of uncertainty – once the log variance is a measure of variability or dispersion. From this perspective, being in a formal job reduces the inequality or uncertainty, besides increasing the expected salary. But being a employer or self-employed increases uncertainty – which is very consistent with the known information about these categories: there are several kinds of employers and self-employed, and just few of them work in very profitable areas or at/for big firms.

Let's now consider the log earnings disparities across age, period and cohort (or within-age, within-period, and within-cohort earnings inequality). The three temporal dimensions have statistically significant effects on log earnings inequality. After controlling for socioeconomic, educational and demographic characteristics, inequality increases constantly with age, as can be seen in Figure 14. The low effect of the squared term leads to this almost linear trend. It means there are cumulative inequalities throughout the aging process: there are more income similarities among younger people similar than among older people.

Figure 15 shows that within period inequality follows a pattern which is similar to the descriptive univariate trend of inequality across the years. It increases until 1989 and decreases steadly and constantly (with some bumps) thereafter. It is very relevant to notice that the estimated period declining inequality trend is much more accentuated when we control for individual level (including age) and cohort effects.

This finding is also particularly interesting because it indicates that a great part of the inequality drop is concentrated between 1989-1992 and 1993-1995. The recent period of inequality fall (from 2001 onwards) in fact appears to be a deceleration of the falling trend. This curve behavior is very different from the univariate descriptive graph of the var log (see Figure 9). It does not mean that inequality has not fallen recently. Period here represent conjuncture effects which are uncorrelated to the covariates included in the model – and also disentangled from age and cohort. This result means that the recent fall of income inequality in Brazil was not driven by those conjuncture effects. It is more related to compositional changes in the labor market.

In fact, Figure 16 indicate that within cohort inequality increases until the cohort born between 1951 and 1955, and does not change for younger cohorts. The fact that inequality increases for older cohorts and do not change for younger cohorts shows that part of the declining inequality trend since the 1990´s is due to cohort replacement. That is, cohorts with increasing inequality were gradually substituted by cohorts with stable inequality in the surveys we have studied. Or, in other words, the trend of increasing inequality across cohorts stopped for cohorts born after 1955, that are precisely the cohorts that compose the majority of surveys from the 1990’s and 2000´s when inequality was decreasing.

Gender-specific effects – Fitted Values

Mean (R$)

Figure 17 - Age

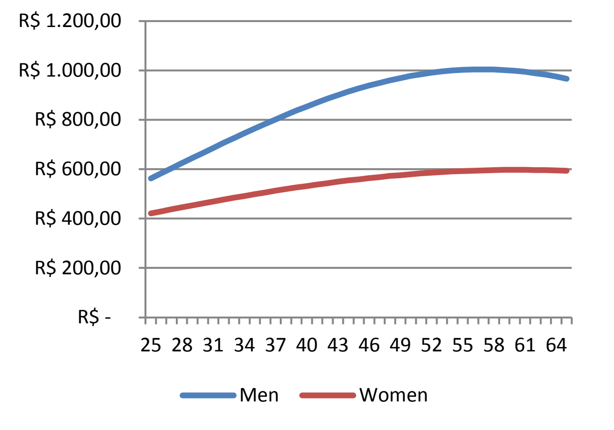


Figure 18 - Period

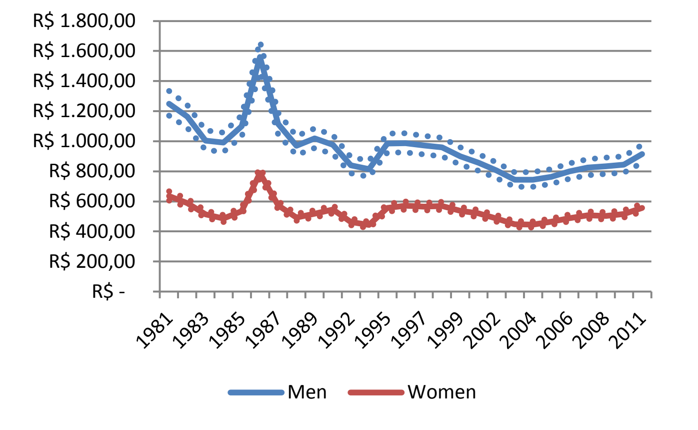
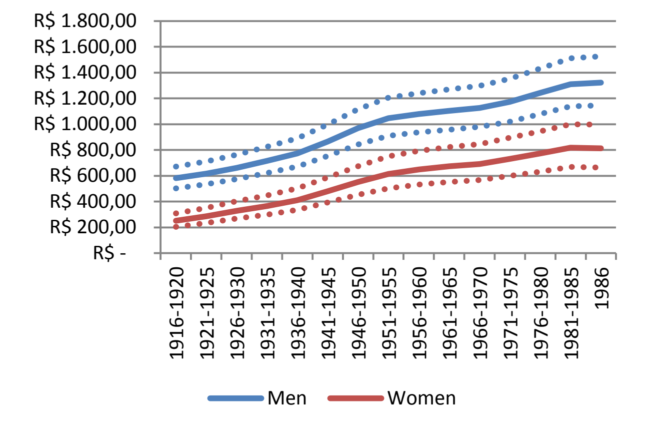


Figure 19 - Cohort



Variance (ln[income])

Figure 20 - Age

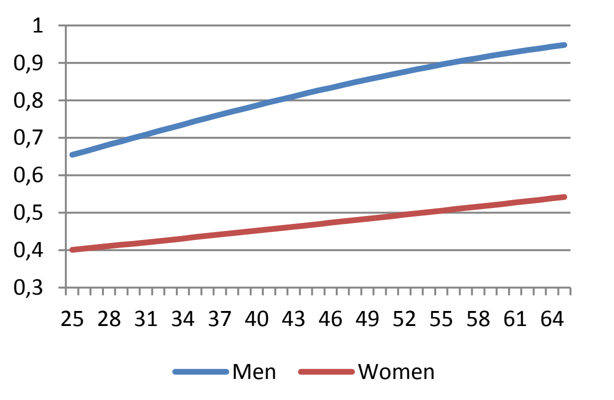


Figure 21 - Period

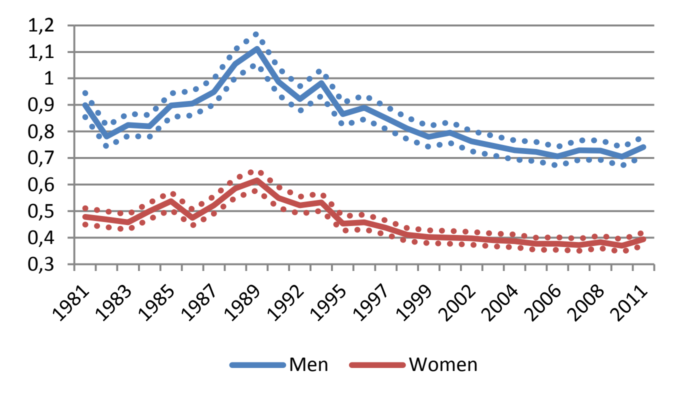
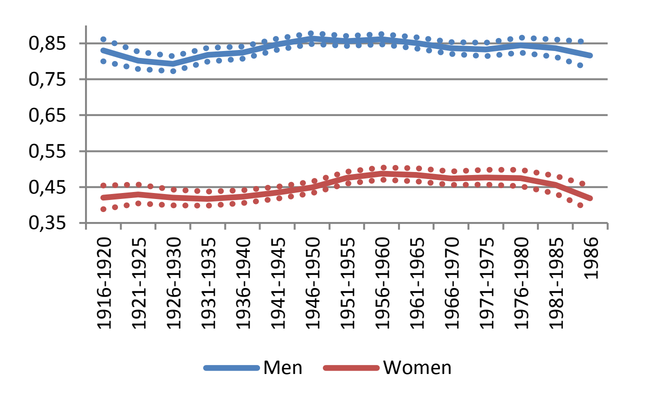


Figure 22 - Cohort



7 – Conclusions and prospective

(To be completed)

Table 1 - Estimates HAPC-CCREM VFR/HR Models of Log Earnings, PNAD, 1981 to 2011.

	Linear Regression (Mean)					Gamma Regression (Variance)			
	Estimate	Std,Error	tvalue	Pr(>|t|)		Estimate	Std,Error	tvalue	Pr(>|t|)
Level-1 Individual fixed effects					**Level-1 Individual fixed effects**				
Intercept	5,762	0,088	65,29	<.0001	**Intercept**	-0,583	0,037	-15,85	<.0001
Age	0,013	0,000	45,34	<.0001	**Age**	0,011	0,001	18,1	<.0001
Age^2	-0,0005	0,0000	-107,79	<.0001	**Age^2**	-0,0001	0,00001	-9,06	<.0001
Years of education (cohort centered)	0,107	0,000	1077,1	<.0001	**Years of education (cohort centered)**	0,049	0,000	185,52	<.0001
Hours	0,010	0,000	301,08	<.0001	**Hours**	-0,007	0,000	-74,01	<.0001
Gender (0=Female)	0,448	0,001	523,62	<.0001	**Gender (0=Female)**	0,072	0,002	29,96	<.0001
Region (North)	-0,094	0,002	-51,12	<.0001	**Region (North)**	0,039	0,004	9,73	<.0001
Region (Northeast)	-0,398	0,001	-378,88	<.0001	**Region (Northeast)**	0,165	0,003	58,52	<.0001
Region (South)	-0,054	0,001	-49,96	<.0001	**Region (South)**	-0,055	0,003	-17,29	<.0001
Region (Center-West)	-0,014	0,002	-8,83	<.0001	**Region (Center-West)**	0,065	0,004	17,58	<.0001
Formality indicator (0=informal)	0,399	0,001	427,27	<.0001	**Formality indicator (0=informal)**	-0,288	0,005	113,02	<.0001
Position: Employer	0,716	0,002	313,85	<.0001	**Position: Employer**	0,573	0,003	191,45	<.0001
Position: Self-Employed	0,052	0,001	45,46	<.0001	**Position: Self-Employed**	0,524	0,003	-113,22	<.0001
					Level-2 Random effects				
Period					**Period**				
1981	0,256	0,029	8,75	<.0001	**1981**	0,114	0,032	3,58	0,0003
1982	0,181	0,029	6,17	<.0001	**1982**	0,016	0,032	0,52	0,6054
1983	0,038	0,029	1,29	0,1962	**1983**	0,037	0,031	1,19	0,2329
1984	0,010	0,029	0,35	0,7274	**1984**	0,068	0,031	2,16	0,0304
1985	0,110	0,029	3,78	0,0002	**1985**	0,151	0,031	4,84	<.0001
1986	0,470	0,029	16,16	<.0001	**1986**	0,108	0,031	3,45	0,0006
1987	0,139	0,029	4,78	<.0001	**1987**	0,165	0,031	5,29	<.0001
1988	-0,005	0,029	-0,18	0,8593	**1988**	0,274	0,031	8,83	<.0001
1989	0,045	0,029	1,55	0,1213	**1989**	0,322	0,031	10,36	<.0001
1990	0,043	0,029	1,47	0,1415	**1990**	0,189	0,031	6,05	<.0001
1992	-0,121	0,029	-4,18	<.0001	**1992**	0,129	0,031	4,19	<.0001
1993	-0,156	0,029	-5,38	<.0001	**1993**	0,172	0,031	5,6	<.0001
1995	0,048	0,029	1,66	0,0974	**1995**	0,022	0,031	0,7	0,4821
1996	0,058	0,029	2,02	0,0438	**1996**	0,042	0,031	1,38	0,1664
1997	0,045	0,029	1,55	0,1213	**1997**	-0,006	0,031	-0,21	0,8352
1998	0,039	0,029	1,35	0,1756	**1998**	-0,062	0,031	-2,03	0,0423
1999	-0,021	0,029	-0,73	0,4647	**1999**	-0,098	0,031	-3,18	0,0015
2001	-0,065	0,029	-2,23	0,0259	**2001**	-0,092	0,031	-2,98	0,0029
2002	-0,128	0,029	-4,41	<.0001	**2002**	-0,121	0,031	-3,92	<.0001

2003	-0,210	0,029	-7,22	<.0001	**2003**	-0,144	0,031	-4,67	<.0001
2004	-0,210	0,029	-7,23	<.0001	**2004**	-0,163	0,031	-5,27	<.0001
2005	-0,178	0,029	-6,14	<.0001	**2005**	-0,182	0,031	-5,88	<.0001
2006	-0,129	0,029	-4,44	<.0001	**2006**	-0,198	0,031	-6,36	<.0001
2007	-0,096	0,029	-3,29	0,001	**2007**	-0,188	0,031	-6	<.0001
2008	-0,092	0,029	-3,17	0,0015	**2008**	-0,178	0,031	-5,67	<.0001
2009	-0,076	0,029	-2,61	0,009	**2009**	-0,213	0,031	-6,78	<.0001
2011	0,004	0,029	0,15	0,8816	**2011**	-0,165	0,032	-5,2	<.0001
Cohort					**Cohort**				
1916-1920	-0,556	0,084	-6,59	<.0001	**1916-1920**	-0,108	0,033	-3,3	0,001
1921-1925	-0,482	0,084	-5,74	<.0001	**1921-1925**	-0,121	0,028	-4,26	<.0001
1926-1930	-0,399	0,084	-4,76	<.0001	**1926-1930**	-0,115	0,026	-4,41	<.0001
1931-1935	-0,314	0,084	-3,76	0,0002	**1931-1935**	-0,078	0,024	-3,2	0,0014
1936-1940	-0,226	0,083	-2,7	0,0069	**1936-1940**	-0,056	0,023	-2,46	0,0138
1941-1945	-0,100	0,083	-1,19	0,2326	**1941-1945**	-0,020	0,022	-0,93	0,3506
1946-1950	0,025	0,083	0,3	0,7645	**1946-1950**	0,012	0,021	0,56	0,5777
1951-1955	0,116	0,083	1,39	0,1657	**1951-1955**	0,036	0,021	1,72	0,0862
1956-1960	0,158	0,083	1,89	0,0583	**1956-1960**	0,059	0,021	2,79	0,0052
1961-1965	0,189	0,083	2,27	0,0232	**1961-1965**	0,063	0,022	2,89	0,0039
1966-1970	0,217	0,083	2,6	0,0094	**1966-1970**	0,059	0,023	2,61	0,0092
1971-1975	0,266	0,084	3,19	0,0014	**1971-1975**	0,074	0,024	3,07	0,0021
1976-1980	0,327	0,084	3,91	<.0001	**1976-1980**	0,096	0,026	3,73	0,0002
1981-1985	0,385	0,084	4,59	<.0001	**1981-1985**	0,090	0,028	3,25	0,0012
1986	0,393	0,084	4,68	<.0001	**1986**	0,011	0,037	0,3	0,764

Covariance Parameter Estimates			**Covariance Parameter Estimates**		
	Estimate	**Sd Error**		**Estimate**	**Sd Error**
Period	0,0225	0,0063	**Period**	0,0244	0,0071
Cohort	0,1042	0,0395	**Cohort**	0,0062	0,0038
Residual	1,0030	0,0008	**Residual**	3,2626	0,0028

-2 Res Log Likelihood	**AIC**	**N**	**-2 Res Log Pseudo-Likelihood**	**Generalized Chi-Square**	**Gener. Chi-Square / DF**	**N**
6.328.373,0	6.328.379,0	2.806.358	11.283.089,0	9.156.030,0	3,26	2.806.358

Table 2 - Estimates HAPC-CCREM VFR/HR Models of Log Earnings, Men, PNAD, 1981 to 2011.

	Linear Regression (Mean)				Gamma Regression (Variance)			
	Estimate	Std,Error	tvalue	Pr(>|t|)	Estimate	Std,Error	tvalue	Pr(>|t|)
Level-1 Individual fixed effects					**Level-1 Individual fixed effects**			

Intercept	6,341	0,079	80,46	<.0001
Age	0,012	0,000	35,03	<.0001
Age^2	-0,001	0,000	-106	<.0001
Years of education (cohort centered)	0,105	0,000	815,09	<.0001
Hours	0,007	0,000	150,29	<.0001
Region (North)	-0,085	0,002	-36,14	<.0001
Region (Northeast)	-0,390	0,001	-292,92	<.0001
Region (South)	-0,057	0,001	-41,02	<.0001
Region (Center-West)	0,000	0,002	0,1	0,9192
Formality indicator (0=informal)	0,394	0,001	328,53	<.0001
Position: Employer	0,729	0,003	282,02	<.0001
Position: Self-Employed	0,104	0,001	76,12	<.0001
Level-2 Random effects				
Period				
1981	0,286	0,033	<.0001	6,6263
1982	0,214	0,033	<.0001	6,5544
1983	0,067	0,033	0,0421	6,40787
1984	0,054	0,033	0,1057	6,3941
1985	0,157	0,033	<.0001	6,4979
1986	0,508	0,033	<.0001	6,8483
1987	0,171	0,033	<.0001	6,512
1988	0,032	0,033	0,3339	6,37242
1989	0,081	0,033	0,0143	6,42127
1990	0,038	0,033	0,2504	6,37842
1992	-0,112	0,033	0,0006	6,2285
1993	-0,142	0,033	<.0001	6,1984
1995	0,048	0,033	0,1427	6,38868
1996	0,049	0,033	0,134	6,38975
1997	0,035	0,033	0,2913	6,37521
1998	0,021	0,033	0,5296	6,36122
1999	-0,043	0,033	0,1933	6,29793
2001	-0,091	0,033	0,0057	6,24989
2002	-0,154	0,033	<.0001	6,1863
2003	-0,234	0,033	<.0001	6,1068
2004	-0,233	0,033	<.0001	6,1081
2005	-0,207	0,033	<.0001	6,1332

Intercept	-0,566	0,026	-21,35	<.0001
Age	0,009	0,000	24,01	<.0001
Age^2	0,000	0,000	-8,34	<.0001
Years of education (cohort centered)	0,050	0,000	145,04	<.0001
Hours	-0,005	0,000	-42,4	<.0001
Region (North)	0,056	0,005	10,74	<.0001
Region (Northeast)	0,150	0,004	40,67	<.0001
Region (South)	-0,031	0,004	-7,53	<.0001
Region (Center-West)	0,035	0,005	7,37	<.0001
Formality indicator (0=informal)	-0,279	0,006	88,97	<.0001
Position: Employer	0,530	0,003	134,82	<.0001
Position: Self-Employed	0,470	0,003	-84,41	<.0001
Level-2 Random effects				
Period				
1981	0,075	0,026	2,91	0,0036
1982	-0,066	0,026	-2,58	0,0098
1983	-0,012	0,026	-0,45	0,6493
1984	-0,017	0,026	-0,67	0,5002
1985	0,074	0,025	2,92	0,0035
1986	0,082	0,026	3,15	0,0016
1987	0,129	0,026	4,98	<.0001
1988	0,235	0,026	9,09	<.0001
1989	0,287	0,026	11,05	<.0001
1990	0,169	0,026	6,5	<.0001
1992	0,101	0,026	3,94	<.0001
1993	0,164	0,026	6,39	<.0001
1995	0,037	0,026	1,44	0,1507
1996	0,064	0,026	2,51	0,0122
1997	0,020	0,026	0,8	0,4245
1998	-0,028	0,026	-1,1	0,2692
1999	-0,066	0,026	-2,6	0,0093
2001	-0,048	0,025	-1,87	0,0614
2002	-0,090	0,025	-3,52	0,0004
2003	-0,112	0,026	-4,39	<.0001
2004	-0,134	0,025	-5,26	<.0001
2005	-0,142	0,026	-5,58	<.0001

2006	-0,161	0,033	<.0001	6,1792
2007	-0,130	0,033	<.0001	6,2107
2008	-0,119	0,033	0,0003	6,222
2009	-0,108	0,033	0,0011	6,2327
2011	-0,026	0,033	0,434	6,31464
Cohort				
1916-1920	-0,481	0,073	<.0001	5,8598
1921-1925	-0,422	0,073	<.0001	5,9189
1926-1930	-0,350	0,072	<.0001	5,991
1931-1935	-0,271	0,072	0,0002	6,0696
1936-1940	-0,194	0,072	0,007	6,1469
1941-1945	-0,082	0,072	0,2545	6,25883
1946-1950	0,032	0,072	0,6594	6,3722
1951-1955	0,108	0,072	0,1316	6,4487
1956-1960	0,138	0,072	0,0536	6,4789
1961-1965	0,160	0,072	0,0257	6,5007
1966-1970	0,182	0,072	0,0112	6,5228
1971-1975	0,223	0,072	0,002	6,5635
1976-1980	0,280	0,072	0,0001	6,6205
1981-1985	0,333	0,072	<.0001	6,6739
1986	0,342	0,073	<.0001	6,6826

Covariance Parameter Estimates

	Estimate	**Sd Error**
Period	0,02885	0,008047
Cohort	0,07699	0,02892
Residual	1,0059	0,001068

-2 Res Log Likelihood	**AIC**	**N**
3.999.383,0	3.999.389,0	1.775.136.

2006	-0,166	0,026	-6,52	<.0001
2007	-0,135	0,026	-5,27	<.0001
2008	-0,135	0,026	-5,27	<.0001
2009	-0,167	0,026	-6,52	<.0001
2011	-0,117	0,026	-4,53	<.0001
Cohort				
1916-1920	-0,005	0,019	-0,24	0,8097
1921-1925	-0,039	0,015	-2,51	0,012
1926-1930	-0,050	0,013	-3,74	0,0002
1931-1935	-0,019	0,012	-1,61	0,1074
1936-1940	-0,012	0,010	-1,13	0,2595
1941-1945	0,016	0,009	1,68	0,0932
1946-1950	0,034	0,009	3,95	<.0001
1951-1955	0,027	0,008	3,16	0,0016
1956-1960	0,032	0,009	3,74	0,0002
1961-1965	0,020	0,009	2,18	0,0289
1966-1970	0,003	0,010	0,33	0,7398
1971-1975	-0,001	0,011	-0,11	0,9114
1976-1980	0,012	0,013	0,99	0,3245
1981-1985	0,003	0,015	0,18	0,8579
1986	-0,022	0,023	-0,96	0,3384

Covariance Parameter Estimates

	Estimate	**Sd Error**
Period	0,01605	0,004481
Cohort	0,000762	0,00044
Residual	3,4759	0,00369

-2 Res Log Pseudo-Likelihood	**Generalized Chi-Square**	**Gener. Chi-Square / DF**	**N**
7.249.472,0	6.170.079,0	3,48	1.775.136.

Table 3 - Estimates HAPC-CCREM VFR/HR Models of Log Earnings, Women, PNAD, 1981 to 2011.

Linear Regression (Mean)					**Gamma Regression (Variance)**			
	Estimate	Std,Error	tvalue	Pr(>|t|)	Estimate	Std,Error	tvalue	Pr(>|t|)
Level-1 Individual fixed effects					**Level-1 Individual fixed effects**			

Intercept	5,651	0,105	53,94	<.0001
Age	0,012	0,000	26,94	<.0001
Age^2	0,000	0,000	-37,42	<.0001
Years of education (cohort centered)	0,108	0,000	701,27	<.0001
Hours	0,014	0,000	268,23	<.0001
Region (North)	-0,108	0,003	-37,07	<.0001
Region (Northeast)	-0,424	0,002	-248,45	<.0001
Region (South)	-0,053	0,002	-31,3	<.0001
Region (Center-West)	-0,032	0,003	-12,41	<.0001
Formality indicator (0=informal)	0,397	0,001	267,27	<.0001
Position: Employer	0,732	0,005	149,72	<.0001
Position: Self-Employed	-0,049	0,002	-23,85	<.0001
Level-2 Random effects				
Period				
1981	0,188	0,025	7,56	<.0001
1982	0,113	0,025	4,57	<.0001
1983	-0,022	0,025	-0,91	0,3641
1984	-0,080	0,025	-3,27	0,0011
1985	0,016	0,024	0,66	0,5107
1986	0,404	0,024	16,66	<.0001
1987	0,085	0,024	3,53	0,0004
1988	-0,065	0,024	-2,7	0,0069
1989	-0,018	0,024	-0,73	0,4641
1990	0,036	0,025	1,42	0,156
1992	-0,128	0,024	-5,35	<.0001
1993	-0,172	0,024	-7,21	<.0001
1995	0,056	0,024	2,35	0,0188
1996	0,083	0,024	3,5	0,0005
1997	0,070	0,024	2,95	0,0032
1998	0,078	0,024	3,28	0,001
1999	0,021	0,024	0,87	0,3825
2001	-0,014	0,024	-0,58	0,5644
2002	-0,078	0,024	-3,25	0,0011
2003	-0,163	0,024	-6,83	<.0001
2004	-0,166	0,024	-6,93	<.0001
2005	-0,126	0,024	-5,26	<.0001
2006	-0,073	0,024	-3,04	0,0023
2007	-0,036	0,024	-1,5	0,1324

Intercept	-0,630	0,035	-18,13	<.0001
Age	0,008	0,001	11,25	<.0001
Age^2	0,000	0,000	-1,3	0,1947
Years of education (cohort centered)	0,049	0,000	119,28	<.0001
Hours	-0,005	0,000	-41,84	<.0001
Region (North)	0,020	0,006	3,12	0,0018
Region (Northeast)	0,209	0,004	47,3	<.0001
Region (South)	-0,081	0,005	-16,02	<.0001
Region (Center-West)	0,124	0,006	20,79	<.0001
Formality indicator (0=informal)	-0,321	0,010	62,21	<.0001
Position: Employer	0,649	0,005	129,45	<.0001
Position: Self-Employed	0,590	0,004	-79,2	<.0001
Level-2 Random effects				
Period				
1981	0,065	0,032	2,01	0,0445
1982	0,044	0,032	1,36	0,1745
1983	0,021	0,032	0,66	0,5107
1984	0,107	0,032	3,38	0,0007
1985	0,180	0,032	5,7	<.0001
1986	0,057	0,032	1,78	0,0748
1987	0,152	0,032	4,77	<.0001
1988	0,267	0,032	8,42	<.0001
1989	0,316	0,032	9,96	<.0001
1990	0,203	0,036	5,57	<.0001
1992	0,151	0,031	4,84	<.0001
1993	0,172	0,031	5,52	<.0001
1995	0,009	0,031	0,28	0,7794
1996	0,021	0,031	0,67	0,5047
1997	-0,025	0,031	-0,81	0,4165
1998	-0,089	0,031	-2,88	0,004
1999	-0,109	0,031	-3,52	0,0004
2001	-0,115	0,031	-3,72	0,0002
2002	-0,122	0,031	-3,96	<.0001
2003	-0,138	0,031	-4,47	<.0001
2004	-0,148	0,031	-4,76	<.0001
2005	-0,175	0,031	-5,64	<.0001
2006	-0,174	0,031	-5,58	<.0001
2007	-0,187	0,031	-5,95	<.0001

2008	-0,044	0,024	-1,82	0,0682	**2008**	-0,158	0,032	-5,02	<.0001
2009	-0,021	0,024	-0,85	0,3946	**2009**	-0,194	0,032	-6,12	<.0001
2011	0,057	0,025	2,3	0,0214	**2011**	-0,131	0,032	-4,07	<.0001
Cohort					**Cohort**				
1916-1920	-0,744	0,106	-7,04	<.0001	**1916-1920**	-0,066	0,040	-1,66	0,0961
1921-1925	-0,607	0,104	-5,87	<.0001	**1921-1925**	-0,043	0,031	-1,39	0,1646
1926-1930	-0,469	0,103	-4,56	<.0001	**1926-1930**	-0,065	0,027	-2,45	0,0144
1931-1935	-0,364	0,103	-3,55	0,0004	**1931-1935**	-0,072	0,024	-3,06	0,0022
1936-1940	-0,246	0,102	-2,4	0,0162	**1936-1940**	-0,060	0,021	-2,81	0,0049
1941-1945	-0,091	0,102	-0,9	0,3708	**1941-1945**	-0,032	0,019	-1,67	0,0947
1946-1950	0,048	0,102	0,47	0,6411	**1946-1950**	0,000	0,018	-0,01	0,9929
1951-1955	0,155	0,102	1,52	0,1283	**1951-1955**	0,060	0,018	3,38	0,0007
1956-1960	0,209	0,102	2,05	0,0407	**1956-1960**	0,082	0,018	4,62	<.0001
1961-1965	0,247	0,102	2,42	0,0155	**1961-1965**	0,075	0,019	3,99	<.0001
1966-1970	0,273	0,102	2,67	0,0076	**1966-1970**	0,056	0,020	2,75	0,006
1971-1975	0,330	0,102	3,22	0,0013	**1971-1975**	0,061	0,022	2,76	0,0058
1976-1980	0,385	0,103	3,75	0,0002	**1976-1980**	0,057	0,024	2,34	0,0192
1981-1985	0,440	0,103	4,28	<.0001	**1981-1985**	0,016	0,027	0,6	0,5472
1986	0,435	0,104	4,2	<.0001	**1986**	-0,069	0,040	-1,72	0,0858

Covariance Parameter Estimates			**Covariance Parameter Estimates**		
	Estimate	**Sd Error**		**Estimate**	**Sd Error**
Period	0,01494	0,004168	**Period**	0,02349	0,006944
Cohort	0,1561	0,05939	**Cohort**	0,00408	0,002502
Residual	0,9993	0,001392	**Residual**	3,0447	0,00424

-2 Res Log Likelihood	**AIC**	**N**	**-2 Res Log Pseudo-Likelihood**	**Generalized Chi-Square**	**Gener. Chi-Square / DF**	**N**
2.297.850,0	2.297.856,0	1.031.222	4.074.978,0	3.139.756,0	3,04	1.031.222

References

BARROS, Ricardo Paes de; CARVALHO, Mirela de; FRANCO, Samuel; MENDONÇA, Rosane. “A queda recente da Desigualdade de Renda no Brasil”. In: BARROS, R. P.; FOGUEL, M. N.; ULYSSEA, G. (orgs) Desigualdade de Renda no Brasil: uma análise da queda recente (vol 1). Brasília: IPEA, 2006a.

BARROS, Ricardo Paes de; FRANCO, Samuel; MENDONÇA, Rosane. “A Recente Queda na Desigualdade de Renda e o Acelerado Progresso Educacional Brasileiro da Última Década”. In: BARROS, R. P.; FOGUEL, M. N.; ULYSSEA, G. (orgs) Desigualdade de Renda no Brasil: uma análise da queda recente (vol 2). Brasília: IPEA, 2006b.

BARROS, Ricardo Paes de; CARVALHO, Mirela de; FRANCO, Samuel; MENDONÇA, Rosane. “Determinantes Imediatos da Queda da Desigualdade de Renda Brasileira”. In: BARROS, R. P.; FOGUEL, M. N.; ULYSSEA, G. (orgs) Desigualdade de Renda no Brasil: uma análise da queda recente (vol 1). Brasília: IPEA, 2006c.

CHAIMOWICZ, Flávio. “A saúde dos idosos brasileiros às vésperas do século XXI: problemas, projeções e alternativas”. Rev. Saúde Pública, 31 (2): 184-200, 1997.

CORSEUIL, Carlos Henrique; FOGUEL, Miguel N. “Uma sugestão de deflatores para rendas obtidas a partir de algumas pesquisas domiciliares do IBGE”. Rio de Janeiro: Ipea, 2002. (Texto para Discussão, n. 897).

FERREIRA, Francisco H. G.; BARROS, Ricardo Paes. “Climbing a moving moutain: explaining the decline in income inequality in Brazil from 1976 to 1996”. First Workshop of the LACEA/IDB/World Bank Inequality and Poverty Network. Buenos Aires, 1998.

FERREIRA, Francisco H. G.; LITCHFIELD, Julie. (2001), “Education or inflation? The Micro and Macroeconomics of the Brazilian Income Distribution during 1981-1995”. Cuadernos de Economía, vol 38, no 114, p. 209-238, 2001.

FERREIRA, Francisco H. G.; LEITE, Philippe G.; LITCHFIELD, Julie A.; ULYSSEA, Gabriel. “Ascensão e queda da desigualdade de renda no Brasil”. Econômica, v.8, no 1, p. 147-169, 2006.

Fienberg, S. E.; Mason , W. M. “Specification and Implementation of Age, Period, and Cohort Models.” in Cohort Analysis in Social Research, edited by W. M. Mason and S. E. Fienberg. New York: Springer-Verlag, 1985.

FISHLOW, A. Brazilian size distribution of income. American Economic Review, v. 62, n. 2, 1972.

HOFFMANN, R.; DUARTE, J. C. A distribuição da renda no Brasil. Revista de Administração de Empresas, v. 12, n. 2, 1972, p. 46-66.

LANGONI, C. Distribuição de renda e desenvolvimento econômico no Brasil. Rio de Janeiro: Expressão e Cultura, 1973.

Mason, K. O., W. H. Mason, H. H. Winsborough, and K. Poole. 1973. “Some Methodological Issues in Cohort Analysis of Archival Data.” American Sociological Review 38:242–58.

Menezes-Filho, Naércio; Fernandes, Reynaldo; Picchetti, Paulo. “Educação e Queda Recente da Desigualdade no Brasil”. IN: BARROS, R. P.; FOGUEL, M. N; ULYSSEA, G. (Orgs). *Desigualdade de renda no Brasil: uma análise da queda recente*. Brasília: IPEA, 2006.

O’BRIEN, R. M. 2000. “Age Period Cohort Characteristic Models.” Social Science Research 29:123–39.

RYDER, Norman. “The cohort as a concept in the study of social change”. American Sociological Review 30 (6), 1965: 843-861.

SOARES, S. Distribuição de renda no Brasil de 1976 a 2004 com ênfase no período de 2001 a 2004. Brasília: IPEA, 2006 (Texto para discussão, n. 1166).

WESTERN, BRUCE e DEIRDRE BLOOME. Variance function regressions for studying inequality. Sociological Methodology, v.39, p.293-326. 2009.

YANG, Yang. "Aging, Cohorts, and Methods". In B. Binstock and L.K. George (eds). *The Handbook of Aging and the Social Sciences* (7th ed). Academic Press, 2011.

Yang, Y., W. J. Fu, and K. C. Land. 2004. “A Methodological Comparison of Age-Period-Cohort Models: Intrinsic Estimator and Conventional Generalized Linear Models.” Sociological Methodology 34:75–110.

YANG, Yang; SCHULHOFER-WOHL, Sam; FU, Wenjiang J.; LAND, Kenneth C. “The Intrinsic Estimator for Age-Period-Cohort Analysis: What It Is and How to Use It”. American Journal of Sociology, Volume 113, Number 6, 2008: 1697–1736

YANG, Yang; LAND, Kenneth C. “A mixed models approach to the age-period-cohort analysis of repeated cross-section surveys, with an application to data on trends in verbal test scores”. Sociological Methodology. Volume 36, Issue 1, pages 75–97, December 2006

YANG, Yang; LAND, Kenneth C. "Age–Period–Cohort Analysis of Repeated Cross-Section Surveys: Fixed or Random Effects?". Sociological Methods & Research, 36 (3), February 2008 pp. 297-326.

ZHENG, Hui; YANG, Yang; LAND, Kenneth. Variance Function Regression in Hierarchical Age-Period-Cohort Models: Applications to the Study of Self-Reported Health. American Sociological Review, v.76, p.955-983. 2011.