October 2001
Preliminary and incomplete.
Please do not cite without permission.

Prices, Preferences or Endowments?: Accounting for Excess Inequality in Brazil

François Bourguignon, Francisco H. G. Ferreira and Phillippe G. Leite[1]

Keywords: Inequality, Distribution, Micro-simulation

JEL Classification Codes: C15, D31, I31, J13, J22

Abstract: This paper develops a micro-econometric approach to investigating differences between income distributions across different countries, based on regarding them as marginals of joint multivariate distributions of household characteristics and entitlements. We thus decompose differences in distributions into labor market price effects, occupational preferences, and human and non-human asset endowments. We apply the method to the differences between the Brazilian income distribution and those of the United States and Mexico, and find that most of Brazil's excess inequality is due to underlying inequalities in the distribution of two key endowments: access to education and to sources of non-labor income.

[1] Bourguignon is with DELTA, Paris, and the World Bank. Ferreira and Leite are at the Department of Economics of the Pontifícia Universidade Católica do Rio de Janeiro. We thank Nora Lustig and Cesar Bouillon at the IDB for making the Mexican data available to us, ready to use. The opinions expressed here are those of the authors and do not necessarily reflect those of the World Bank, its Executive Directors or the countries they represent.

1. Introduction

The distribution of economic welfare in Brazil is one of the world's most unequal. In *Facing Up to Inequality in Latin America*, the IDB (1998) ranked both Brazil's total Gini coefficient (0.60) and its urban-only Gini coefficient (0.57) as the highest in the region.[2] Its ratio of per-capita urban to per-capita rural household incomes (3.0) was also the highest in Latin America. The World Bank's point estimates for Gini coefficients, listed in *Attacking Poverty* (WDR 2001) for as many countries as the Bank dares, include only two higher than Brazil's, worldwide: Sierra Leone's and the Central African Republic's.

This paper asks why that is the case. What is special about Brazil, to make it so much more unequal than other countries? In particular, we investigate the comparative roles of three factors: the distribution of endowments; the structure of returns to these endowments, and evidence on the preferences that govern their use. Does Brazil's income inequality arise predominantly from a larger inequality in the distribution of the ownership of productive factors – mostly human capital, physical capital and land - or from higher relative returns to those factors that are the most unequally distributed? Or does it originate in the unequal use being made of these factors – through personal choices relating to labor supply, education or fertility? Or is it the way in which factors that are individually owned are combined within households – say through marriages?

Obviously, we are not the first people to notice Brazil's inequality... Among economists working with household-level data, Fishlow (1972) and Langoni (1973) would have a better claim to that title. Reis and Barros (1991) and Ramos (1993) also made seminal contributions, establishing the importance of education and its remuneration in explaining - or at least describing - the phenomenon. Henriques (2000) is probably the most comprehensive current reference on the state of research on income distribution and poverty in Brazil. Yet - although many studies have compared aggregate measures of Brazil's inequality with those of other countries, and many cross-country regressions of inequality on other aggregate variables have been run - we are not aware of a previous

[2] At 0.54, the country's rural-only Gini coefficient came second to Paraguay's.

attempt to identify the microeconomic sources of inequality by comparing endowments, their rates of return and household behavior, as observed in household surveys in different countries. This is what we do, using a micro-simulation-based approach, which is described in Section 3 below, and applying it to a comparison of the determinants of the distributions of income in Brazil, Mexico and the United States.[3] The new perspective that such analysis provides for thinking about distributional issues in Brazil in both original and fruitful.

The comparison relies on a simple econometric model of household behavior, in which individual earnings, labor supply, demand for education and fertility choices are treated as endogenous. The model is estimated for each country, on household-level survey data. Then, by combining the various country models – or sub-models – with the data of another country, differences in the distributions across countries are decomposed into three broad types of effects: those related to differences in the structure of returns in the labor market; those originating from differences in the distributions of household or personal characteristics or endowments; and those pertaining to different labor supply or occupational choice behavior. This approach is a cross-country extension of a methodology previously developed to analyze the dynamics of the distribution of income within a single country.[4]

The data source used in Brazil is the 1999 PNAD[5]. The strengths and the weaknesses of this data source are well known. Among the former is the size of the sample (about one hundred thousands households) and its extensive national coverage. Weaknesses include the very imperfect coverage of rural incomes and non-labor incomes. This latter problem is

[3] As will become apparent, space and cost reasons prevent any single research project to apply this methodology to a large sample of countries. One inevitably has to choose a few countries to act as comparators for Brazil and, while other choices are certainly possible, ours was driven by four factors. First, the US economy - and its labor market in particular - have been abundantly researched and are reasonably well understood. Second, the US occupy a position in the inequality ranking of developed countries analogous to Brazil's among emerging economies. Third, as a dominant developed country, and one widely perceived to have reasonably laissez-faire social preferences, the US is often a relevant comparator for Brazilian policy-makers. Fourth, having thus justified the choice of an economy at a level of development so different from Brazil's, we wanted to test the insights against a much more similar economy, with good household data. Mexico, we argue, was the natural choice.

[4] See Bouguignon, Ferreira and Lustig (2001) and for Brazil, Ferreira and Paes de Barros (1999).

[5] The *Pesquisa Nacional por Amostra de Domicilios* (PNAD) is fielded annually (except in Census years) by the Brazilian Statistical Office (IBGE).

common to most household survey data and Brazil is no exception. Yet, a further decomposition of non-labor incomes into various sources will allow us to address that difficulty as well as we can. As for the poor coverage of rural incomes, there is no other remedy than to limit the analysis to urban areas.

The paper is organized as follows. Section 2 summarizes what can be learned from conventional comparisons of income distributions, summary inequality and poverty measures, and decomposition of some of these measures with respect to key population characteristics in the three countries being analyzed. The methodology of the extended decomposition undertaken in this paper is presented in section 3. The results obtained in the case of the Brazil-US comparison are discussed in some detail in section 4. Section 5 discusses the Brazil-Mexico comparison. Section 6 concludes by assessing the effectiveness of this methodology in identifying the determinants of distributional differences across countries, and summarizes our findings regarding what makes Brazil so uniquely unequal.

2. Income Distribution in Brazil, Mexico and the United States.

This section compares the distributions of income in the three most populous countries in the Western Hemisphere. The comparisons are based on an analysis of the original household-level data sets by the authors, with the Pesquisa Nacional por Amostra de Domicílios (PNAD) 1999 being used for Brazil; the Encuesta Nacional de Ingresos y Gastos de Hogares (ENIGH) 1994 for Mexico, and the Current Population Survey (CPS) 2000 for the United States. See Appendix 1 for a brief description of these data sets.

Income, rather than consumption, data is used for two reasons. The first is to ensure consistency, since the decompositions described in the remainder of the paper are based on earnings determination. The second – and sufficient - reason is that no nationally representative household survey with a consumption module has been fielded in Brazil since 1975.[6] As elsewhere, however, there are reasons to suspect that incomes are measured with substantial error. In the case of Brazil, evidence suggests that the problem is particularly severe in rural areas, to the extent that the usefulness of any estimate based on these income data is thrown into doubt.[7] For this reason, we prefer to confine our attention to urban areas only, in Brazil and Mexico.[8] Care is taken to ensure that the distributions used are as comparable as possible, and this requires that we work with data unadjusted for misreporting, imputed rents, or for regional price level differences within countries.[9]

Table 1 below reports some key summary statistics of the income distributions for our three countries. In addition to mean and median incomes, three inequality measures are computed: the Gini Coefficient, the Theil T and L indices – in what follows, the last two are sometimes labeled E(1) and E(0), respectively, as members of the class of generalized entropy inequality measures. Each of these statistics is presented for the distribution of household income per capita, as well as for a distribution of equivalised income, where the Buhmann et. al. ($\theta = 0.5$) equivalence scale is used. [10] All households are weighed by the number of individuals they comprise.

[6] Although this is about to change with the planned launch of a nationally representative Household Budget Survey (POF) in 2001-2002.

[7] For evidence on these weaknesses, see Ferreira, Lanjouw and Neri (2000) and Elbers, Lanjouw, Lanjouw and Leite (2001).

[8] For the US, the CPS does not identify households as belonging to rural or urban areas. All households are therefore included.

[9] These are three common types of adjustment to raw household survey data. The first scales responses in different income categories to bring aggregate values into line with National Account estimates. The second imputes to households that own their homes a value for imputed rent. The third deflates incomes in accordance with a regional price index, so as to compensate for aggregate regional differences in the cost of living. Although the methodological details of each of these procedures are debatable, we think all are reasonable practices in principle. We do not undertake them here simply because, in each case, it could not be done in a comparable way for all five countries.

[10] According to that method, the equivalised income of a household with income y and size N is taken to be y/N^{θ}. This definition coincides with income per capita when $\theta=1$.

Table 1: Descriptive Statistics						
Country	Population (millions, 1999)	GDP per capita (monthly, USD)	Mean equivalised income (monthly, USD)	Gini Coefficient	Theil-T	Theil-L
$\theta = 1.0$						
Brazil	168	526.42	290.34	0.587	0.693	0.646
Mexico	97	643.25	280.90	0.536	0.580	0.511
USA	273	2550.00	1691.64	0.445	0.349	0.391
$\theta = 0.5$						
Brazil	168	526.42	551.08	0.560	0.613	0.572
Mexico	97	643.25	587.91	0.493	0.478	0.423
USA	273	2550.00	2791.78	0.415	0.298	0.344
Notes: Population and GDP per capita figures are from World Bank (2001). The other figures are from calculations by the authors from the household surveys. GDP per capita and mean equivalised income (MEY) are monthly and measured in 1999 US dollars at PPP exchange rates. Mexican survey data is for 1994; Brazilian survey data is for 1999, and US survey data is for 2000. Values of θ are for the economy of scale parameter in the Buhmann et.al. (1988) equivalence scale - $\theta = 1$ corresponds to income per capita. .						

Similarities between Brazil (in 1999) and Mexico (in 1994) are immediately apparent. Across those different years, the two countries had broadly similar levels of GDP per capita. Mexico's was 22% higher than Brazil's , which pales in comparison to the difference between the two countries and the US: 384% higher than Brazil's. Brazil's inequality is ranked highest by all three measures reported, followed by Mexico and the United States. The difference between Brazil's and Mexico's Ginis, at approximately five points, is not too large, while there are a full fourteen points between Brazil and the US. It is interesting to note that the effect of allowing for (a good deal) of scale economies in household consumption differs across both countries and measures. Focusing on the Gini coefficient, the reduction in inequality in Mexico from reducing θ from 1.0 to 0.5 is larger than either in the US or Brazil.

The considerable differences in both mean incomes and inequality across these five countries must translate into different poverty levels as well. Table 2 below presents the three standard FGT[11] poverty measures for each country, based on the distribution of per capita household incomes. The first panel shows poverty rates for the entire countries, whereas the second panel shows them for urban areas only, which is the universe for the analysis carried out in the next sections of the paper. In both cases, we use two alternative poverty thresholds. The first block in each panel employs an absolute poverty line, originally calculated as a strict indigence line for Brazil by Ferreira, Lanjouw and Neri (2000). Translated to 1999 values, it was set at R$74.48, or US$83.69 at PPP exchange rates. Having the lowest mean and the highest inequality of the three countries, Brazil has the most poverty by all three measures, in urban areas and overall. The United States has, by this ungenerous developing country standards, only traces of poverty. As for Mexico, it is striking how much of its poverty is rural: poverty incidence falls from 23% nationally, to less than 7% in urban areas. While being mindful that urban-rural definitions vary across countries, it would seem that poverty has an even more predominantly rural profile in Mexico than in Brazil.

But when one considers welfare across countries at such different levels of development and per capita income as these three countries, a strong argument can be made that a relative poverty concept might be more appropriate. For this reason we also present the same poverty measures, in the same distributions, calculated with respect to a line set at half the median income in each distribution, in the second block of each panel. By these more relative standards, poverty in the US reaches a full quarter of the population, which happens to be quite similar to Brazil's urban incidence. Mexico's P(0) also rises to 15% in urban areas.

These poverty measures and, more directly, the scalar inequality indices presented earlier in Table 1 confirm Brazil as the most unequal of our three countries. They are usefully complemented by an inspection of the Lorenz Curves for each country, as shown in Figure 1. As in the remainder of our analysis, the distributions considered for Brazil and Mexico

[11] Foster, Greer and Thorbecke (1984). In what follows, we use the three common measures of that family of

are urban only. Even though this reduces their inequality, it can be seen that, apart from the usual difficulty with establishing clear rankings at the tails, Brazil is Lorenz dominated by both Mexico and the United States, whereas those two countries, at least with only urban Mexico being considered, can not be Lorenz ranked.

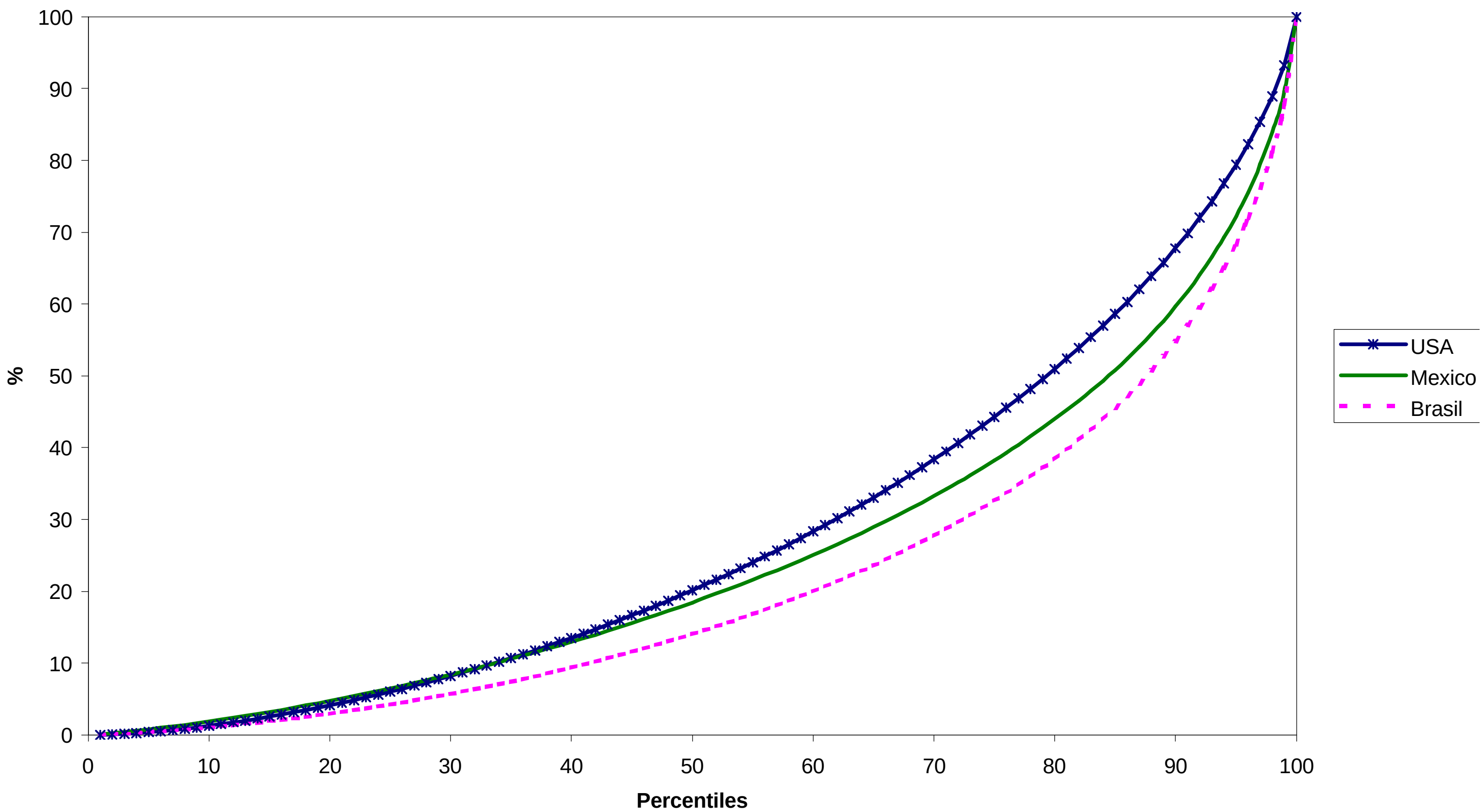
FGT(α) measures for Urban and Rural areas

	P(0)	P(1)	P(2)	***Poverty line*** [1]
Brazil	29,18	12,10	6,74	83,69
Mexico	23,29	8,02	3,84	83,69
USA	1,41	0,75	0,54	83,69
Brazil	30,02	12,22	6,82	84,27
Mexico	17,86	5,59	2,57	70,11
USA	25,02	10,19	5,92	687,70

FGT(α) measures for Urban areas

	P(0)	P(1)	P(2)	***Poverty line*** [1]
Brazil	22,33	8,40	4,37	83,69
Mexico	6,66	1,52	0,51	83,69
Brazil	26,74	10,42	5,55	95,51
Mexico	14,98	3,73	1,39	110,46

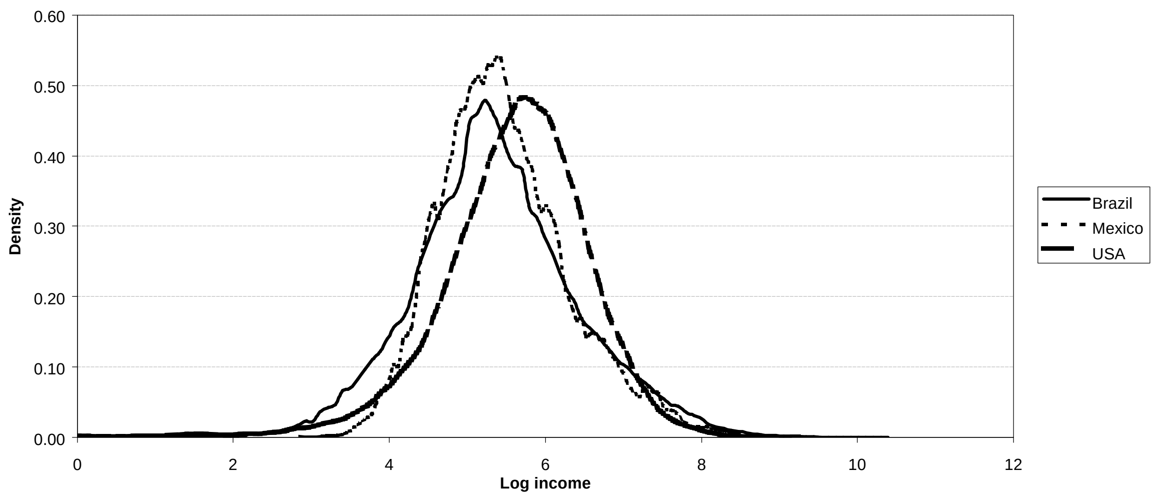
Figure 1: Urban Lorenz Curve For Brazil, Mexico and the US.



poverty indices : P(0), the headcount, P(1), the poverty gap and P(2), the cumulated squared gap.

The Atkinson Theorem (1970) – which establishes the link between normalized second-order stochastic dominance and unambiguous inequality ranking - makes Lorenz Curves very useful diagrammatic tools to compare income distributions. Nevertheless, because they are two levels of integration above a density function, we can do even better in terms of picturing the distribution. Figure 2 below plots kernel estimates of the density functions for the distribution of household per capita income in our three countries.

Figure 2: Income Distributions for Brazil, Mexico and The United States



Sources: PNAD/IBGE 1999, CPS/ADS 2000
Note: Gaussian Kernel Estimates (with optimal window width) of the density functions for the distributions of the logarithms of household per capita incomes. The distribution were scaled so as to have the brazilian mean. Brazil and Mexico are urban areas only. Incomes were converted to US dollar at PPP exchange rates (see Appendix).

Finally, Table 3 reports on standard decompositions of E(0) and E(1) by population subgroups[12], computing the R_B statistic developed by Cowell and Jenkins (1995). This statistic is an indicator of the relative importance of each attribute used to partition the population, in the process of "accounting for" the inequality. The idea is that the larger the share of dispersion which is between groups defined by some attribute - rather than within those groups - the more likely it is that something about the distribution of or returns to that attribute are causally related to the observed inequality.

The attributes to be used include education of the household head (or main earner for the distribution of household incomes); his or her age; his or her race or ethnic group; his or her

[12] See Bourguignon (1979), Cowell (1980) and Shorrocks (1980).

gender; as well as the location of the household (both regional and rural/urban) and its size or type. The results of each decomposition do *not* control for the other attributes, and are not to be interpreted as tests of causality, but can provide useful indications of the nature of inequality in the different countries.

The results are suggestive. In Brazil, education of the head is clearly the most important partitioning characteristic, followed by race and family type. In the US, family type dominates, with education a surprisingly low second, and age of head third. In Mexico, education and urban/rural vie for first place, with family type third. It is clear that education accounts for more inequality in the Brazil (and Mexico) than in the US, although this technique can not tell us whether this is due predominantly to different returns or different endowments of education – i.e. a different distribution of the population across educational levels. The greater role of the urban/rural partition in Mexico is in line with our findings regarding total and urban poverty rates there, and suggest that any comparison we make between the two countries in this paper must clearly be understood to refer only to urban areas, and need not apply overall.

Table 3: Theil Decompositions of Inequality by Population Characteristics

	Brasil			**USA**			**Mexico**		
	RB(0)	RB(1)	RB(2)	RB(0)	RB(1)	RB(2)	RB(0)	RB(1)	RB(2)
Região	0,092	0,076	0,031	0,003	0,004	0,003	0,113	0,103	0,050
Tipo de Família	0,126	0,121	0,060	0,192	0,210	0,155	0,194	0,180	0,092
Zona	0,101	0,073	0,026	-	-	-	0,253	0,194	0,079
Gênero	0,000	0,000	0,000	0,002	0,002	0,002	0,000	0,000	0,000
Raça	0,137	0,119	0,051	0,024	0,024	0,016	-	-	-
Faixa de educação	0,266	0,316	0,213	0,129	0,133	0,093	0,247	0,255	0,150
Faixa etária	0,051	0,047	0,021	0,082	0,091	0,066	0,042	0,037	0,017

Strikingly little of overall US inequality is between different regions of the country, reinforcing the widespread perception of a well-integrated economy. This is in contrast to the two Latin American countries, where some 10% of the Theil-L is accounted for by the regional partition. For a much more detailed analysis of the importance of regional effects in Mexican inequality, see Legovini, Bouillon and Lustig (2000). Finally, it is interesting to note that inequality between households headed by people of different races - which one would expect to be prominent in the US - is five to six times as large in Brazil.

3. The Micro-Simulation-Based Cross-Country Decomposition Methodology

In this section we go beyond summary inequality measures, and seek a deeper understanding of the differences between income distributions in Brazil, Mexico and the US, based on the differences between the determinants of household incomes across these countries. We consider three broad groups of determinants: prices, preferences and endowments. Endowments are personal and household characteristics which can be transacted in markets, generating income. Given the nature of the household data, we have information on four types of characteristics which we treat as endowments. The first three affect remuneration rates in the labor markets, and are as follows: (i) exogenous observed characteristics – such as race, age or gender; (ii) "endogenous" observed characteristics – such as education or fertility); (iii) unobserved characteristics (such as individual ability). The fourth includes claims on physical or financial assets, including entitlements to public or private transfers, which generate non-labor incomes. Ownership of these is generally not observed in our household surveys, and certainly not in a comprehensive manner, and we measure them by the reported income flows that they generate.

Prices refer to the structure of returns to the first three types of endowments, in the labor markets. And preferences refer to the observed relationships between characteristics and patterns of behavior. These relationships are parametrically estimated through discrete choice models for three types of behavior: occupational choice, fertility decisions, and educational decisions.

The main purpose of the decomposition exercise described below is to understand how each of these three groups of determinants explain differences in the income distributions between Brazil and the two other countries. This can be done through changes in aggregate inequality and poverty measures, but also by looking at the differences in the incomes of individual percentiles of the distribution. We explore both the aggregated and disaggregated routes below.

We think of household incomes as a function of a set of personal and family characteristics, and of occupational choice and market return parameters. We write:

$$y_h = G(E, A, R, n_h, r, g; y_0; \alpha, \beta, \lambda) \tag{1}$$

where E is the vector of educational endowments of the household members; A is their age vector, R is the vector of their racial characteristics, n_h is the number of household members, r is the spatial subset where the household is located, g is a vector of gender dummies, y_0 is total household non-labor income; α and β are parameters of a simple earnings function described below, and λ are parameters of an occupational choice function described below.

We impose a fairly conventional structure on the G function. Household incomes are an aggregation of individual earnings y_{hi}, and of additional, unearned income such as transfers or capital income, y_0. Per capita household income for household h is given by:

$$y_h = \frac{1}{n_h}\left[\sum_{i=1}^{n_h}\sum_{j=1}^{J} I_{hi}^{j} y_{hi}^{j} + y_0\right] \tag{2}$$

where I_{hi}^{j} is an indicator variable that takes the value 1 if individual i in household h participates in earning activity j, and 0 otherwise. y_{hi}^{j} is assumed to be log-linear in α_j and $\boldsymbol{\beta}_\mathbf{j}$, and the individual earnings equation is estimated separately for males and females, as follows:

$$\log y_{hi}^{j} = \alpha_j + \mathbf{x}_{hi}\beta_j + \varepsilon_i \tag{3}$$

where **x** = {1, education dummies, age, age squared, age * education, and intercept dummies for region, race}. In the absence of specific information on experience, the education and age variables are the standard Becker - Mincer human capital terms. The intercept dummies allow for a simple level effect of possible spatial, segmentation of the labor market, as well as for the possibility of racial discrimination. Earnings activity are defined by the sector of activity and the informal status. To simplify, it is also assumed that earning functions across activities differ only through the intercepts, so that the sets of coefficients $\boldsymbol{\beta_j}$ are the same across activities ($\boldsymbol{\beta_j} = \boldsymbol{\beta}$**)**. These earnings equations constitute the first level of the model, and generate two sets of parameters: the α_js and βs are the "price effect" parameters – interpreted as the returns to the attributes in $\mathbf{x_{hi}}$ - and the variance of the residual terms, s^2_{em}, s^2_{ef}, are the conduits for the effects of unobserved characteristics.

The second level models the determination of participation behavior and occupational choice, thus determining I^j_{hi} for each individual i; that is his/her participation as well as the nature (both sectoral and formal/informal) of the earning activity he or she engages in. This is done through a multinomial logit of the form:

(4)
$$\Pr\{j = s\} = P^s(Z_{hi}, l) = \frac{e^{Z_{hi} l_s}}{e^{Z_{hi} l_s} + \sum_{j \neq s} e^{Z_{hi} l_j}}$$

where $P^s(\ )$ is the probability of choosing earning activity s among: inactivity, formal employment in industry, informal employment in industry, formal employment in services and informal employment in services. Once again, separate but identically specified models are estimated for males and females. Z = {age, age squared, education dummies, age * education, race, and region for the individual in question; average endowments of age and education among adults in his or her household; numbers of adults and children in the household; whether the individual is the head or not; and if not whether the head is active}. As it is well known, the multi-logit model may be interpreted as a utility-maximizing discrete choice model where the utility associated with choice j is given by :

(5)
$$U^j_{hi} = Z_{hi}.l_j + h^j_{hi}$$

where the last term on the RHS, that stands for unobserved choice determinants of individual hi, is assumed to be distributed according to a double exponential law in the population.

This level of the methodology generates the occupational choice parameters, λ, and (random) estimates of the residual terms h_{hi}^{s} .[13] Equation (5) can be interpreted as a utility-maximizing labor supply model (by sector) in reduced form, since **x** is contained in Z, as are averages of the determinants of the wage-earning power in the household. However, we prefer not to rely on such an explicit formulation of supply-driven occupational choice behavior, with which this model might (or might not) be consistent. Instead, we prefer to see (4) and (5) as a purely statistical description of the correlations between various personal and household endowments, on the one hand, and occupational choices on the other; estimated under a maintained assumption about the (multilogit) functional form. Behavior behind this functional form may be that of individuals choosing freely among various occupations, or that of firms selecting in a discriminatory way their employees, or any combination of these two extreme assumptions.

But the income distribution function $F(y_h)$ is not fully explained by α, β, λ, σ. In fact, we can think of its density function as obtained from the joint distribution function of household characteristics through integrating along the household income generation function G( ):

$$(6) \qquad f(y) = \int_{G(.)=y} j\left(E, A, R, n_h, r; y_0\right) \partial E \partial A \partial R \partial N \partial r \partial y_0$$

Equations (2), (3) and (4) generate estimates of the price and participation parameters α, β, λ, σ and define the function G( ). To complete our understanding of the determinants of f(y), we must now turn to the distribution of the endowments (observed personal characteristics and claims on non-labor incomes) of household members, i.e. the function φ( ) above. In particular, we focus on their numbers and age structure, their education levels, race and gender. Among these, we treat the number of adults (n_{ah}) and their age structure, as well as race and gender, as fully exogenous. We do allow, however, for the fact that the individual educational levels are chosen (at least in part) by individuals and that this choice depends on their other personal characteristics. Educational endowments are thus not distributed independently of other personal and household characteristics. The same is true for the number of children in the household, $n_{ch} = n_h - n_{ah}$.

We must therefore estimate the conditional distribution functions of education and fertility on the exogenous endowments: $H(E \mid A, R, r, g, n_{ah})$ and $J(n_{ch} \mid E, A, R, r, n_{ah})$. Again, these are estimated through multinomial logit functions analogous to that in equation (4). Unlike those for participation, these are estimated jointly for men and women. The educational choice multilogit $H(E \mid .)$ has as choice categories 1-4; 5-6; 7-8; 9-12; and 13 and more years of schooling, with 0 as the omitted category, and the independent variables are the individual's age, race, gender, cohort and region of residence. Estimation of this model generates the educational endowment parameters, $\boldsymbol{\gamma}$**.** The demographic multilogit $J(n_{ch} \mid .)$ has as choice categories the number of children in the household: 1, 2, 3, 4 and 5 and more, with 0 as the omitted category. The independent variables are the number of adults in the household; as well as the race, region, education and age of the spouse of the head.[14] Estimation of this model generates the demographic endowment parameters, $\boldsymbol{\psi}$. These two "endowment demand" models constitute level 3 of the model.

The fourth and final level of the model consists of simulating the effect of the truly exogenous endowment effects – age (A), race (R) and adult household size (n_{ah}). Since these variables do not depend on other exogenous variables in the model, this estimation is carried out simply by re-calibrating the population by the weights corresponding to the joint distribution of these attributes in the target country.

After each of these reduced-form models have been estimated on two countries (Brazil and each of the comparator nations), the simulation strategy is as follows: Step 1 imports the αs and βs from one country, separately and jointly, into the earnings equations of the other country, and recomputes inequality and poverty measures, as well as the differences in log incomes along the entire distribution arising from the exercise. Since this step simulates the

---

[13] For details on how the latter may be determined, see Bourguignon, Ferreira and Lustig (2001).
[14] Or of the head if there is no spouse.

impact of adopting in one country the returns (or "price") structure of the other; we call it the *price effect*. Step 2 does the same thing but adds the importing of the residual variances in the earnings equations. It combines the price and unobserved characteristic effects.

Step 3 imports the level 2 participation parameters (λ), and only them. This is modifying the utility of the various occupations as given by (5), so that the choice of some individuals will be modified, depending on their observed (Z_{hi}) and unobserved (h_{hi}^{s}) characteristics. The simulated distribution and the measures related to it are constructed by giving those people whose occupational choice has changed the incomes which model (3) predicts they would earn. Residuals for equation (3) are drawn randomly from a normal distribution in with mean zero and the empirically observed variances. Step 4 combines steps 3 and 2: once individuals have been given their new incomes in accordance with their simulated participation behavior, their incomes are calculated using the imported αs ,βs and σs. This step combines price, unobserved characteristics and participation effects.

Step 5 imports the γs from the educational multilogit, and changes education endowments in the earnings equation to recalculate incomes. Step 6 combines steps 5 and 4 in a manner analogous to step 4. Step 7 imports the ψs from the demographic multilogit and recomputes per capita incomes by changing household sizes in the distribution according to the predictions of this model. Step 8 combines steps 7 and 6 in a manner analogous to step 4.

In step 9, the final exogenous endowment effects are estimated, by importing the age and racial make-up of the population in one country into the other. This is done by partitioning the two populations by the numbers of adults in the household. To remain manageable, the partition is in three groups: households with a single adult; households with two adults; and households with more than two adults. Each of these groups is then further partitioned by the race (whites and non-whites) and age category (six groups) of each adult.[15] The number of household in each of these subgroups can be denoted $M_{a,r}^{n,C}$, where *a* stands for the age

[15] In the case of households with more than two adults, this is done for two adults only: the head and a randomly drawn other adult. In this manner, the group of single adult households is partitioned into 12 sub-groups, and the other two groups into 144 sub-groups each.

category of the group, *r* for the race of the group, *n* for the number of adults in the household, and *C* for the country. If we are importing the structure from country A (population of households P^A) to country B (population of households P^B), we then simply re-scale the household weights in the sample for country B by the factor:

(7) $$f_{a,r}^{n} = \left(\frac{M_{a,r}^{n,A}}{M_{a,r}^{n,B}} \right) \frac{P^B}{P^A}$$

The income distribution simulated by this step is the original country B distribution, re-weighted by the factors calculated in equation (7).

Step 10 consists of the final simulation for labor incomes: it imports to country B all price (α, β) and participation parameters (λ) from country A, as well as the unobserved endowment parameters (σ); the observed endogenous endowment parameters (γ, ψ), and the observed exogenous endowment structure (**ö**). The income distribution simulated by this step is the distribution simulated in Step 8, re-weighted by the factors calculated in equation (7).

The final step is to simulate the impact of adopting the comparator country's distribution of unearned income, y_0. This is done in Step 11. Changing the distribution of unearned income in one country into the distribution in another would be easy through a standard rank-preserving functional switch. More important than the marginal distribution of y_0 however, is its conditional distribution on other household characteristics. We choose here to condition on household labor earnings, rather than on a larger number of characteristics.

In practical terms, we rank the two distributions by per capita household earned income $y_e = y_h - \frac{y_0}{n_h}$. If $p = F_B(y_e)$ is the rank of household with income y_e in country B, then Step 11 replaces y_{op}^{B} with the unearned income of the household with the same rank (by earned income) in country A, after normalizing by mean unearned incomes: $y_{op}^{A} \frac{m_B(y_0)}{m_A(y_0)}$.

This operation was carried out cumulatively for various types of unearned income, namely retirement, pension (mostly survivor), capital and property income, private and public transfers, and other incomes.

Once this is done, the income distribution simulated by this step is merely the distribution of y_h simulated as above, once again ranked by y_h. In the tables that follow, this step is represented by the letter î. Step 12 combines this treatment of unearned incomes with the complete simulation of labor incomes from Step 10. The simulation exercise is then complete.

4. The Brazil-US Comparison.

The decomposition described in the previous section was conducted for differences in distributions between Brazil in 1999 and the United States in 2000. Table 4 - in the Appendix - presents the results for importing the parameters from the US into Brazil, in terms of means and inequality measures for the individual earnings distributions, separately for men and women. Table 5 displays analogous results for household per capita incomes, and includes also three poverty measures. Figures 4 to 8 present the full picture, by plotting differences in log incomes between the distributions simulated in various steps and the original distribution, for each percentile of the new distribution.[16]

Looking first at individual earnings, the observed differences between the Gini coefficients in Brazil and the US are nine points for men, and ten for women. Brazil's gender-specific earnings distributions have a Gini of 0.5, whereas those of the US are around 0.4. Broadly speaking, price effects (identified by simulating Brazilian earnings with the US á and â parameters) account for half of this difference. As we shall see, this is a much greater share than that which will hold for the distribution of household incomes per capita. Among the different price effects, the coefficient on the interaction of age and education stands out as making the largest difference.

[16] Analogous figures for differences in log incomes by percentiles ranked by the original distribution – which show the re-rankings induced by each simulation - are available from the authors on request.

Differences in participation behavior are unimportant in isolation. Importing the US participation parameters only contributes to reducing Brazilian earnings inequality when combined with importing US prices, as may be seen by comparing the rows α,β (vii) and the row λ,α,β. In other words additional participants or changes in the sectoral allocation of earners are spread along the whole distribution with the Brazilian earning structure, whereas they tend to concentrate at the middle of the distribution with the US earning structure.

Educational and fertility choices are more important effects. The former raises educational endowments and hence both increases and upgrades the sectoral profile of labor supply. The latter leads to increased participation rates by women. This effect accounts for nearly all of the remaining four to five Gini points. As one would expect, demographic effects are particularly important for the female distribution, where, in combination with the effect of education, it reduces the Brazilian Gini by a full five points even before any changes are made to prices. Reweighing the purely exogenous endowments - including race - has no effect.

Table 5, which reports on the simulations for the distribution of household incomes per capita, can be read in an analogous way. The first two lines present inequality and poverty measures for the actual distributions of household per capita income by individuals in Brazil (in 1999) and the US (in 2000). In terms of the Gini coefficient, the gap we are trying to "explain" is substantial: it is twelve and a half points higher in Brazil than in the US. The difference is still much bigger when entropy inequality measures E( ) are used.

The first block of simulations suggests that differences in the structure of returns to observed personal characteristics in the labor market can account for some five of these thirteen points.[17] When one disaggregates by individual âs, it turns out that returns to education, conditionally on experience – as for individual earnings- play the crucial role.

[17] The relative importance of each effect varies across the four inequality measures presented, but the orders of magnitude are broadly the same, and the main story could be told from any of them. All are presented in Table 5, but we use the Gini for the discussion in the text.

Table 5 : Simulated Poverty and Inequality for Brazil in 1999, Using 2000 USA coefficients.

		Mean p/c Income	Inequality				Poverty: Z =median/2 per month		
			Gini	E(0)	E(1)	E(2)	P(0)	P(1)	P(2)
1	Brasil	294,8	0,569	0,597	0,644	1,395	26,23	10,10	5,36
2	USA	294,8	0,445	0,391	0,349	0,485	25,02	10,19	5,92
3	α e β	294,9	0,516	0,486	0,515	1,049	20,32	7,53	3,92
4	α , β e σ^2	294,9	0,530	0,517	0,545	1,119	21,92	8,39	4,46
5	λ	277,9	0,579	0,632	0,653	1,313			
6	λ, α e β	255,4	0,535	0,536	0,542	1,022	28,06	11,58	6,46
7	λ, α , β e σ^2	255,5	0,548	0,565	0,572	1,093	29,59	12,50	7,06
8	γ	454,0	0,505	0,489	0,460	0,719			
8a	γ, α e β	283,9	0,480	0,425	0,425	0,732	18,81	7,12	3,75
8b	γ, α , β e σ^2	283,9	0,494	0,453	0,452	0,786	20,33	7,84	4,18
9	λ e γ	469,0	0,511	0,514	0,467	0,711			
10	λ, γ, α e β	274,2	0,490	0,450	0,445	0,780	21,15	8,36	4,54
11	λ, γ, α , β e σ^2	274,2	0,505	0,480	0,474	0,837	22,73	9,19	5,07
12	ψ	295,2	0,576	0,613	0,663	1,449			
13	ψ e γ	464,6	0,505	0,493	0,454	0,686			
14	ψ, γ, α e β	287,1	0,486	0,437	0,434	0,746	19,31	7,31	3,85
15	ψ, γ, α , β e σ^2	287,1	0,499	0,464	0,459	0,794	20,85	8,09	4,35
16	ψ, λ e γ	507,2	0,500	0,492	0,441	0,641			
17	ψ, λ, γ, α e β	299,2	0,481	0,433	0,423	0,709	18,14	7,00	3,75
18	ψ, λ, γ, α , β e σ^2	299,2	0,495	0,462	0,448	0,755	19,59	7,77	4,24
19									
20	ϕ	404,7	0,585	0,637	0,683	1,496			
21	ϕ, ψ, λ, γ, α , β e σ^2	387,7	0,511	0,490	0,489	0,874	14,35	5,43	2,88
	10000 partições								
22	**ψ, λ, γ, α , β e σ^2 ; ξ^1**	**258,1**	**0,488**	**0,470**	**0,428**	**0,684**	**22,53**	**9,57**	**5,60**
23	**ψ, λ, γ, α , β e σ^2 ; ξ^2**	**249,9**	**0,494**	**0,511**	**0,436**	**0,677**	**25,26**	**11,72**	**7,45**
24	**ψ, λ, γ, α , β e σ^2 ; ξ^3**	**277,7**	**0,447**	**0,382**	**0,349**	**0,488**	**16,61**	**6,41**	**3,59**
25	**ψ, λ, γ, α , β e σ^2 ; ξ^4**	**282,6**	**0,449**	**0,409**	**0,352**	**0,485**	**16,43**	**6,71**	**4,02**
26	**ψ, λ, γ, α , β e σ^2 ; ξ^5**	**284,7**	**0,448**	**0,407**	**0,351**	**0,484**	**16,14**	**6,57**	**3,93**
27	**ϕ ; ψ, λ, γ, α , β e σ^2 ; ξ^1**	**322,2**	**0,502**	**0,499**	**0,462**	**0,779**	**19,00**	**8,00**	**4,71**
28	**ϕ ; ψ, λ, γ, α , β e σ^2 ; ξ^2**	**303,2**	**0,510**	**0,554**	**0,472**	**0,762**	**21,88**	**10,47**	**6,85**
29	**ϕ ; ψ, λ, γ, α , β e σ^2 ; ξ^3**	**344,8**	**0,450**	**0,388**	**0,353**	**0,497**	**12,20**	**4,74**	**2,69**
30	**ϕ ; ψ, λ, γ, α , β e σ^2 ; ξ^4**	**352,1**	**0,450**	**0,411**	**0,353**	**0,490**	**12,09**	**5,04**	**3,08**
31	**ϕ ; ψ, λ, γ, α , β e σ^2 ; ξ^5**	**355,2**	**0,450**	**0,410**	**0,353**	**0,489**	**11,88**	**4,93**	**3,01**

Source: PNAD 1999 e CPS March 2000

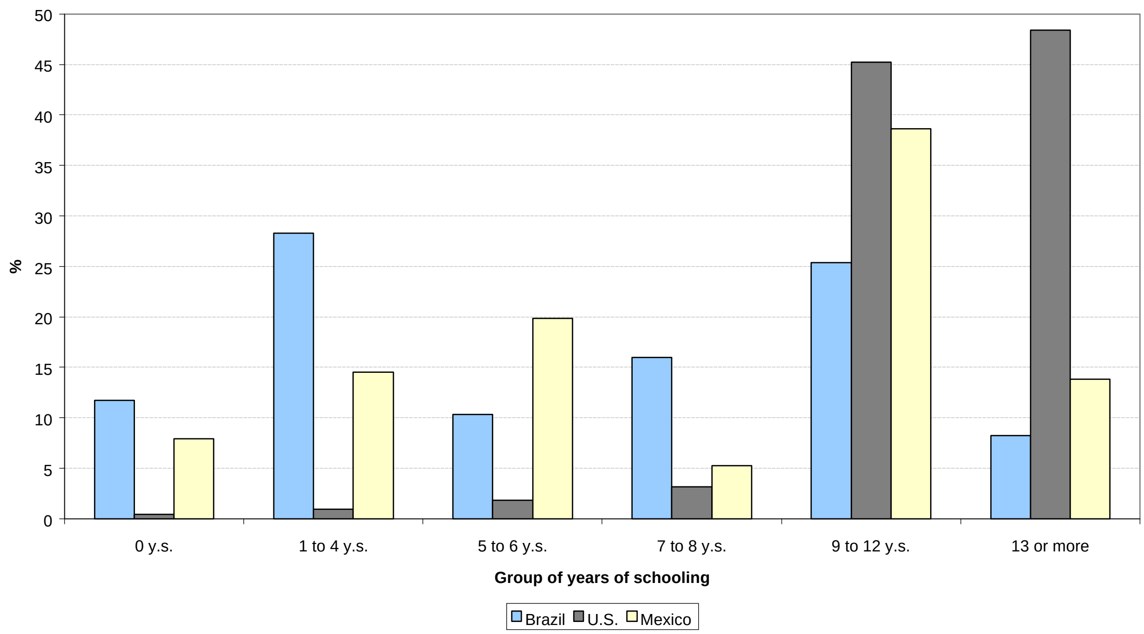
Note: ξ^1 – Private transfers | ξ^2 - Pensions + ξ^1 | ξ^3 - Retirement + ξ^2 | ξ^4 - Capital income + ξ^3 | ξ^5 - Other incomes + ξ^4

Overall, it can thus be said that difference in returns to schooling and experience together explain approximately 40 per cent of the difference of inequality between Brazil and the US. The order of magnitude is practically the same with E(1) and E(2) but it is higher with E(0), suggesting that the problem is not only that returns to schooling are relatively higher at the top of the schooling scale but also that they are relatively lower at the bottom. This is confirmed by the fact that importing US prices lowers poverty in Brazil, even though (relative) poverty is initially comparable in the two countries.

Importing the US variance of residuals goes in the opposite direction, contributing to an increase of almost 1.5 Gini points in Brazilian inequality. Two candidate explanations suggest themselves: either there is greater heterogeneity amongst US workers along unobserved dimensions (such as ability) than among their Brazilian counterparts, or the US labor market is more efficient at observing and pricing these characteristics. This is an interesting question, which deserves further investigation. In the absence of additional information on, say, the variance of IQ test results or other measures of innate ability, orthogonal to education, we are inclined to favor the second interpretation. It may be that the lower labor market turnover and longer tenures that characterize the US labor market translate into a lessened degree of asymmetric information between workers and managers in that country, with a more accurate remuneration of endowments which are unobserved to researchers. We thus consider the ó^{2} effect as a price effect, which dampens the overall contribution of price effects to some 3.5 to 4 points of the Gini.

The next block shows that importing US occupational choice "preferences" - or behavioral parameters - by themselves, has almost no impact on Brazilian inequality, but lowers average incomes and raises poverty. This is a consequence of the great differences in the distribution of education across the two countries, as revealed by Figure 3 below. Since education is negatively correlated with inactivity, and positively with employment in industry and with formality in the US, when we simulate participation behavior with US parameters but Brazilian levels of education, we withdraw a non-negligible number of people from the labor force, and 'downgrade' many others. This negative effect on lower incomes intensifies when one also imports the US structure of returns (i.e. simulates ë, á, â, ó²). Figure 5 shows the impoverishing effect of imposing US occupational choice behavior, combined with its price effect, on Brazil's original distribution of endowments.

Figure 3: Distribution of education across the countries



Sources: PNAD/IBGE 1999, CPS/ADS 2000, ENIGH 1994

Turning to endowments, and beginning with the 'endogenous' endowment effects, we see further support for the aforementioned role of education in determining occupational choice. When US educational choice parameters are simulated by themselves, this raises education levels in Brazil substantially, thus significantly increasing incomes and reducing poverty. Education endowments increase more for the poor (as expected by the upper-bounded nature of the education distribution), and inequality also falls dramatically. The ã simulation alone takes six points of the Gini off the Brazilian coefficient and, crucially, takes the impoverishing effect away from the occupational choice simulation. The latter result suggest that the most important difference in the distribution of educational endowments between Brazil and the US might actually be in the lack of minimum compulsory level in Brazil – see figure 3. This also explains why importing US labor-supply and occupational choice behavioral patterns is poverty-augmenting in Brazil only when Brazil keeps its own educational distribution. When it adopts US educational behavior too, participation parameters have a muted effect on both the distribution's mean and dispersion. This can also be seen in Figure 6, for the whole distribution.

Moving on to demographic behavior, we observe a similar role for education. As with occupational choice, importing ø alone hardly changes inequality – it would even increase it slightly. However, fertility is negatively correlated with educational attainment, particularly of women. If the change in fertility were taking place in the Brazilian population with US levels of schooling and participation behavior, inequality would drop by 1 percentage point of the Gini coefficient and poverty would fall. This seems to mean that fertility behavior differs between the two countries mostly for lowest educated households.

When educational and demographic (i.e. 'endogenous' endowment) effects are combined with occupational choice and price effects (the line for ø, ë, ã á, â, *ó*^{2}), we see an overall reduction of seven points in the Gini. Most of this (four to five points) seems to be directly or indirectly associated with adopting the US endowments of education, either directly or indirectly, through knock-on effects on participation and fertility. The remaining (and non-negligible) three points or so are due to the price effects.[18] This still leaves, however, some additional five Gini points - a rather substantial amount - in the difference in inequality between the two countries unexplained. Figure 7 illustrates the results of the combined simulations for the entire distribution: while the simulated line has moved much closer to the actual (log income percentile) differences, it is not yet a very good fit.

Of the various candidate factors we are considering, two remain: the exogenous observed personal endowments (basically race, age and adult family composition); and non-labor incomes. The two final blocks of simulations show that it is the latter, rather than the former, that accounts for the remaining inequality differences. While reweighing the households in accordance with the procedure outlined in Steps 9 and 10 actually has an increasing effect on Brazilian inequality (thus weakening the explanatory power of the

---

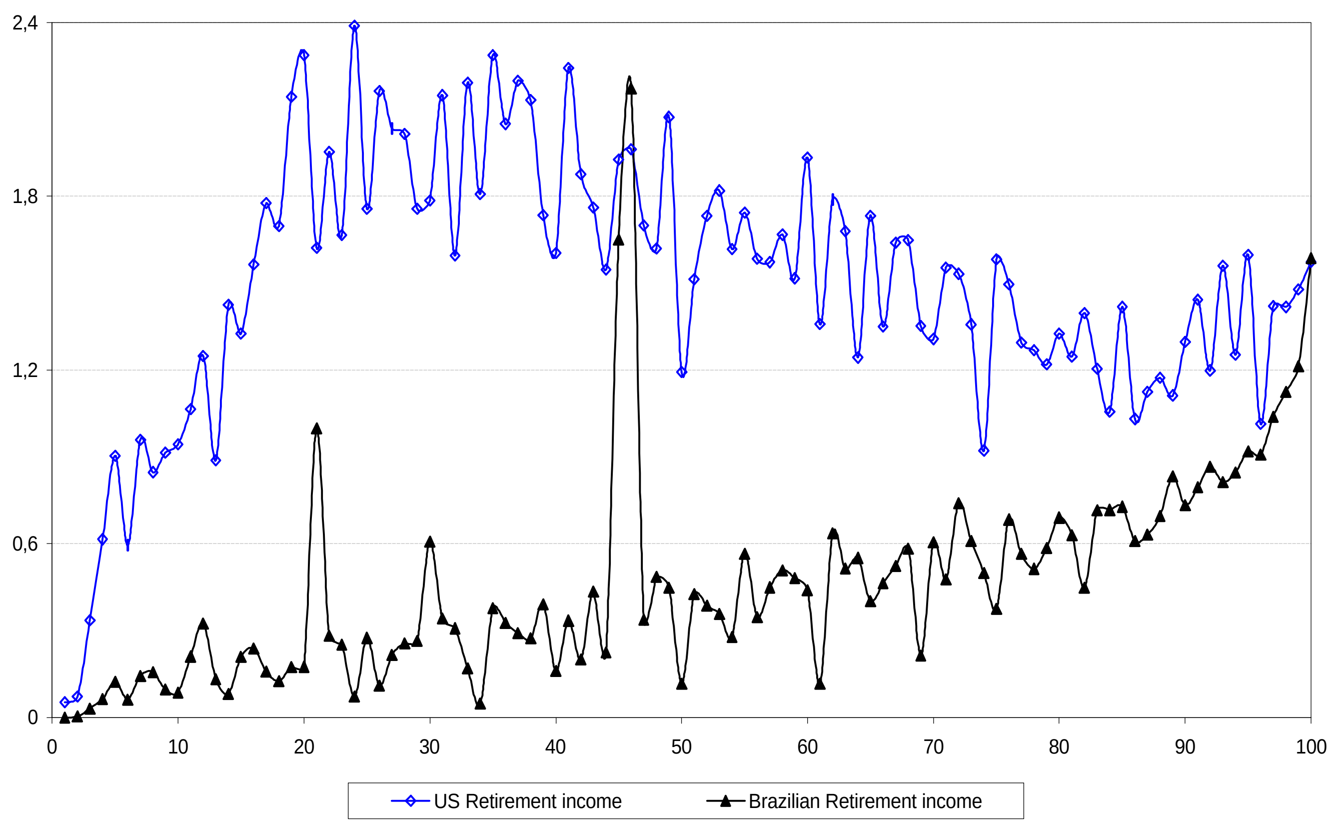
[18] This allocation of the various effects is made difficult by the fact that their size depends on the other effects already being accounted for. The figures mentioned here are obtained as averages over the various possible configurations appearing in table 5.

overall simulation by about one and half Gini points), Steps 11 and 12, whereby the US distribution of non-labor incomes is superimposed on the Brazilian distribution of earned income, has surprisingly large explanatory power. As may be seen from the bottom row of the two blocks of the bottom panel of table 5, it actually moves the simulated Gini coefficient for Brazil to within two tenths of a Gini point of the true US Gini. Figure 8, which shows the curve for simulated income differences for Step 12 nearly on top of the actual differences curve, graphically illustrates the success of the simulated decomposition.

The detailed results which appear in the bottom of table 5 show very clearly that the contribution of unearned income to the difference between the distribution in Brazil and in the US comes essentially from two components of unearned income. Private transfers are responsible for a drop in the Gini coefficient equal to 0.7 percentage points, certainly not a negligible effect. But most of the effect of unearned income is in effect due to retirement income. Even after equalizing its mean and its marginal distribution across the two countries, it is the case that retirement income is strongly unequalizing in Brazil, whereas it would be (mildly) equalizing in the US. This appears quite clearly in figure 9 which shows mean unearned incomes, by sources, for each centile of the distribution of household income. If it were not for some outliers in the middle of the distribution, retirement income clearly concentrates among the riches households in Brazil, whereas it is the largest in the deciles just below the median in the US. The explanation of that difference is simple. Retirement income in Brazil concentrates among retirees of the formal sector who tend to be better off than the rest of the population. In the US, on the contrary, retirees are more evenly distributed in the population. When summing up all income sources, they tend to be around the median of the distribution. Hence the switch from Brazilian to US retirement income is very strongly equalizing, reflecting first of all the universality of retirement in the US and the privilege that it may represent in Brazil.

Overall, the bottom line seems to be that differences in income inequality between Brazil and the United States are predominantly due to differences in the underlying distributions of endowments in the two countries, including among endowments the right to retirement income. Of the almost thirteen Gini points difference, almost ten can be ascribed to endowment effects. Among these, the data suggest almost equally important roles to inequalities in the Brazilian distribution of human capital (as proxied by years of schooling), and other claims on resources, measured by flows of non-labor income. But when the composition of unearned income is taken into account, the distinction between these two types of 'endowments' is no longer so clear. If one were to consider retirement income as a part of labor income in the formal sector, then very much of the difference in the way that that income is distributed in Brazil and in the US could be attributed to schooling as well. More highly educated workers in Brazil are more often employed during part of their lifetime in the formal sector and collect high retirement income. Less educated workers, on the contrary, collect almost nothing because they did not access the formal sector. If it were possible to endogenize retirement income, it is thus probable that education would be the main explanatory variable, which would still reinforce its dominant role in explaining the huge distributional difference between Brazil and the US.

Figure 9: Incidence of Retirement Pensions in Brazil and the US



The remaining three points of the Gini are due to price effects and, in particular, steeper returns to education in Brazil than in the US. Combined to the more unequal distribution of educational endowments themselves, this confirms the importance of education (prices and quantities) in driving Brazilian inequality, as previewed by the Theil decompositions reported in Section 2. While human capital remains firmly at the center-stage, our results suggest that it is joined there by the distribution of non-labor incomes and, in particular, of post-retirement incomes.

5. *The Brazil - Mexico Comparison*

The differences between the distributions of household income per capita of Brazil and Mexico are much smaller than those between either country and the US. The two Latin American countries are at roughly the same level of development, and both are high inequality countries in international terms. Nevertheless, urban Brazil is much poorer than urban Mexico, and more unequal by any of the four measures reported in Table 7 below. Its Lorenz curve, in Figure 1, lies everywhere below Mexico's.

In terms of the Gini coefficient, Brazil's excess inequality amounts to some seven points. Price effects account for 1.2 of these, with the variance of the residuals making no contribution at all to differences between Mexico and Brazil. Participation and occupational choice behavior also account for about a Gini point, but its interaction with the price effects is more-than-additive. The combined impact of all price and participation effects is of more than three points of the Gini.

Education alone also accounts for some three Gini points, but its interaction with occupational choice and price effects is less-than-additive. Joint simulation of Mexican ë, ã, á, â and ó^{2} account for some four and a half of the seven-point difference. Interacting demographic effects takes away another Gini point from Brazil's measure, again only once Mexican educational behavior has been imported too.

Reweighing the Brazilian population so that its make-up in terms of exogenous characteristics - such as age, race and household type - is the same as Mexico's has a small inequality-reducing effect – the Gini coefficient falls by 0.7 percentage point. This effect is slightly bigger when these new exogenous endowments are interacted with Mexican (endogenous endowments of) education and fertility, as well as its price and occupational choice effects. They also help subtract a Gini point.

Table 7: Simulated Poverty and Inequality for Brazil in 1999, Using 1994 Mexico coefficients.

	Mean p/c Income	Inequality				Poverty Z =median/2 per month		
		Gini	E(0)	E(1)	E(2)	P(0)	P(1)	P(2)
Brasil	294,8	0,569	0,597	0,644	1,395	26,2	10,1	5,4
Mexico	294,8	0,498	0,420	0,495	1,028	15,0	3,7	1,4
α e β								
α , β e σ^2								
λ	289,5	0,557	0,567	0,608	1,229	25,5	9,6	5,0
λ, α , β e σ^2	281,3	0,535	0,518	0,552	1,079	23,6	8,7	4,5
γ	375,3	0,537	0,544	0,532	0,908	18,0	6,9	3,6
λ e γ	399,2	0,535	0,540	0,525	0,889	16,5	6,1	3,2
λ, γ, α e β	285,1	0,522	0,500	0,513	0,950	22,9	8,6	4,4
λ, γ, α , β e σ^2	285,1	0,524	0,502	0,516	0,957	23,1	8,6	4,5
ψ	275,5	0,579	0,619	0,671	1,496	29,9	11,9	6,4
ψ e γ	348,0	0,537	0,550	0,529	0,891	20,5	8,1	4,4
ψ, λ e γ	389,7	0,532	0,538	0,514	0,844	17,4	6,6	3,5
ψ, λ, γ, α e β	282,6	0,514	0,490	0,493	0,887	22,9	8,8	4,7
ψ, λ, γ, α , β e σ^2	282,6	0,515	0,491	0,494	0,888	22,9	8,8	4,7
ϕ	284,5	0,562	0,579	0,625	1,330	26,5	10,2	5,4
ϕ, ψ, λ, γ, α , β e σ^2	269,2	0,506	0,471	0,473	0,834	23,4	8,9	4,7
Brasil ; ξ_0	291,9	0,529	0,488	0,554	1,216	20,6	6,3	2,8
ψ ; $\xi 0$	272,7	0,535	0,498	0,571	1,294	23,7	7,5	3,3
ψ, λ, γ, α , β e σ^2 ; $\xi 0$	279,9	0,447	0,348	0,356	0,539	14,7	4,4	1,9
ϕ ; Brasil ; ξ_0	283,7	0,522	0,475	0,535	1,138	20,9	6,5	2,8
ϕ ; ψ ; $\xi 0$	265,7	0,527	0,484	0,551	1,205	23,9	7,5	3,4
ϕ ; ψ, λ, γ, α , β e σ^2 ; $\xi 0$	268,6	0,437	0,331	0,337	0,496	14,9	4,4	1,9

Fonte: PNAD 1999

Altogether, the preceding effects account for almost all the difference observed between Brazil and Mexico, in terms of the Gini coefficient. This is not true, however, of the other inequality measures or of poverty, as shown in table 7. In particular, it can be seen that very little of the excessive relative poverty in Brazil is explained by the decomposition methodology, when it is limited to price, occupational choice and socio-demographic effects, a feature that also appears quite clearly in figure 12. As in the comparison with US, it may thus be expected that what is left explained actually corresponds to the factors behind unearned income.

The impact of importing the Mexican distribution of non-labor incomes into Brazil is powerfully equalizing. This is particularly clear on Figure 13. By itself, it subtracts four points from the Brazilian Gini, and six points from the headcount index.[19] Tellingly, it almost halves the distribution-sensitive poverty measure FGT(2). At the same time, it may also be seen that, when combined with all the preceding changes, importing the structure of Mexican unearned incomes overshoots the observed difference between the two countries – see also figure 13. This suggests that some interaction between the first set of factors and unearned income, or some components of unearned income, are not properly taken into account. One reason for this may be the conditioning of the distribution of unearned income that relies exclusively on labor income and not directly on socio-demographic characteristics. More investigation is necessary here. In any case, however, the results obtained so far suggest that the Brazilian urban poor are at a disadvantage in terms of access to non-human assets and to public or private transfers when compared not only to their US counterparts - which might not be so surprising - but also when compared to the Mexican urban poor. This is an issue of clear relevance for the design of poverty-reduction policy in Brazil. Identifying more precisely the reasons of the difference with Mexico deserves further investigation. [20]

[19] Note that the Brazil - Mexico simulations appear, on the whole, to be less additively separable than the Brazil - US simulations. The sum of individual effects in Table 7 is, on the whole, further away from the corresponding combined effects than in Table 5. From a theoretical viewpoint there is, of course, no reason to expect additive separability. But there would certainly be gains from understanding the economics of the interaction between the different effects that do not add up in the Brazil - Mexico simulations.

6. Conclusions

This paper proposed a micro-econometric approach to investigating the nature of the differences between income distributions across countries. The approach regards the observed distribution of household incomes as a marginal of the joint distribution of a number of household attributes. We then impose a standard Mincerian structure on the determination of earnings in the labor market, and estimate these pricing parameters econometrically on household level data. We then impose a logistic structure on the distribution of occupational choices, conditional on other household attributes, as well as on the conditional distributions of education and fertility choices, so as to enable an estimation of parameters of these conditional distributions.

We then follow micro-simulation techniques that build on earlier work by Oaxaca (1973), Blinder (1973), and more recently by Juhn, Murphy and Pierce (1993), to decompose differences across income distributions into effects due to three broad sources: differences in the returns or pricing structure prevailing in the countries' labor markets; differences in the parameters of the occupational choice conditional distribution (which we associate with the underlying preferences that govern decisions in that realm); and differences in the endowments of age, race, gender, education, fertility and non-labor assets, broadly defined.

We applied this approach to the question of what makes the Brazilian distribution of income so unequal. In particular, we considered the determinants of the differences between it and the distributions of two other large American nations: Mexico and the United States. We found that differences in preferences account for little in both cases. Prices were not insubstantial in explaining difference between the US and Brazil, with this being due largely to steeper returns to education in Brazil. But the most important source of Brazil's uniquely large income inequality is the underlying inequality in the distribution of its human and non-human endowments. In particular, the main causes of Brazil's inequality

[20] This could not be done at this stage of the project because of the absence of any disaggregated information on Mexican unearned incomes, in the database being used.

- and indeed of its urban poverty - seem to be poor access to education and claims on assets and transfers that potentially generate non-labor incomes.

The importance of these non-labor incomes was one of our chief findings. Income distribution in Brazil would be much improved if only the distribution of this income component was more similar to those of the US or Mexico - themselves hardly paragons of the Welfare State. If this is due to public transfers, which needs to be investigated further, it is possible that our findings would vindicate those who have argued for a speedier public approach to the reduction in inequality than that which would be available from educational policies alone.

Another key finding was that the impact of educational expansion has important gender implications, since growth in women's schooling reduces the number of children per household and increases female labor force participation.

References.

Atkinson, Anthony B. (1970): "On the Measurement of Inequality", *Journal of Economic Theory*, **2**, pp.244-263.

Blinder, Alan S. (1973): "Wage Discrimination: Reduced Form and Structural Estimates", *Journal of Human Resources*, **8**, pp.436-455.

Bourguignon, François (1979): "Decomposable Income Inequality Measures", *Econometrica*, **47**, pp.901-20.

Bourguignon, François, Francisco H.G. Ferreira and Nora Lustig (1998): "The Microeconomics of Income Distribution Dynamics in East Asia and Latin America", World Bank DECRA mimeo.

Buhmann, B., L. Rainwater, G. Schmaus e T. Smeeding (1988): "Equivalence Scales, Well-being, Inequality and Poverty: Sensitivity Estimates Across Ten Countries using the Luxembourg Income Study database", *Review of Income and Wealth,* **34,** pp.115-42.

Cowell, Frank A. (1980): "On the Structure of Additive Inequality Measures", *Review of Economic Studies*, **47**, pp.521-31.

Cowell, Frank A. and Stephen P. Jenkins (1995): "How much inequality can we explain? A methodology and an application to the USA", *Economic Journal*, **105,** pp.421-430**.**

Elbers, Chris, Jean O. Lanjouw, Peter Lanjouw and Phillippe G. Leite (2001): "Poverty and Inequality in Brazil: New Estimates from Combined PPV-PNAD Data", World Bank, DECRG Mimeo.

Ferreira, Francisco H.G., Peter Lanjouw and Marcelo Neri (2000): "A New Poverty Profile for Brazil using PPV, PNAD and census data", PUC-Rio, Department of Economics, TD#418.

Ferreira, Francisco H.G. and Ricardo Paes de Barros (1999): "The Slippery Slope: Explaining the Increase in Extreme Poverty in Urban Brazil, 1976-1996", *Brazilian Review of Econometrics*, **19** (2), pp.211-296.

Fishlow, Albert (1972): “Brazilian Size Distribution of Income”, *American Economic Association: Papers and Proceedings 1972*, pp.391-402.

Foster, J., J. Greer, and E. Thorbecke, 1984, “A class of decomposable poverty measures.” *Econometrica,* **52**, pp.761-65.

Henriques, Ricardo (2000): *Desigualdade e Pobreza no Brasil*, (Rio de Janeiro: IPEA)

IDB (1998): *Facing Up to Inequality in Latin America*, (Washington, DC: Inter-American Development Bank).

Juhn, Chinhui, Kevin Murphy and Brooks Pierce (1993): “Wage Inequality and the Rise in Returns to Skill”, *Journal of Political Economy*, **101** (3), pp. 410-442.

Langoni, Carlos G. (1973): *Distribuição da Renda e Desenvolvimento Econômico do Brasil*, (Rio de Janeiro: Expressão e Cultura).

Legovini, Arianna, César Bouillon and Nora Lustig (2001): "Can Education Explain Income Inequality Changes in Mexico", IADB Poverty and Inequality Unit, mimeo.

Oaxaca, Ronald (1973): "Male-Female Wage Differentials in Urban Labor Markets", *International Economic Review*, **14**, pp.673-709.

Ramos, Lauro (1993): *A Distribuição de Rendimentos no Brasil: 1976/85* (Rio de Janeiro: IPEA).

Reis, José G. Almeida and Ricardo Paes de Barros (1991): “Wage Inequality and the Distribution of Education: A Study of the Evolution of Regional Differences in Inequality in Metropolitan Brazil”, *Journal of Development Economics*, **36**, pp. 117-143.

Shorrocks, Anthony F. (1980): "The Class of Additively Decomposable Inequality Measures", *Econometrica*, **48**, pp.613-25.

World Bank (2000): *Attacking Poverty: World Development Report, 2000/2001*, (New York: Oxford University Press).

Table 4: Simulated Poverty and Inequality for Brazilian earnings in 1999, Using 2000 USA coefficients.

		MEN						WOMEN					
		Mean p/c Income	Inequality					Mean p/c Income	Inequality				
			Gini	E(0)	E(1)	E(2)	V(log)		Gini	E(0)	E(1)	E(2)	V(log)
Brasil		636,3	0,517	0,467	0,510	0,902	0,837	411,1	0,507	0,450	0,488	0,838	0,819
USA		636,3	0,427	0,355	0,325	0,441	0,820	411,1	0,409	0,336	0,288	0,362	0,814
α e β													
	i. Intercepto	636,3	0,517	0,467	0,510	0,902	0,837	411,1	0,507	0,450	0,488	0,838	0,819
	ii. Educação	636,3	0,513	0,479	0,485	0,783	0,948	411,1	0,479	0,401	0,423	0,674	0,761
	iii. Experiência	636,3	0,575	0,609	0,644	1,244	1,120	411,1	0,535	0,506	0,549	0,986	0,914
	iv. Raça	636,3	0,515	0,463	0,507	0,893	0,830	411,1	0,497	0,430	0,467	0,791	0,783
	v. Iteração idade/educação	636,3	0,439	0,332	0,344	0,504	0,642	411,1	0,461	0,374	0,386	0,586	0,731
	vi. Sector de atividade	636,3	0,513	0,457	0,502	0,884	0,817	411,1	0,508	0,451	0,489	0,839	0,823
	vii. Formalidade	636,3	0,517	0,476	0,509	0,900	0,887	411,1	0,517	0,484	0,506	0,876	0,929
	viii. All betas	636,3	0,460	0,379	0,376	0,545	0,767	411,1	0,453	0,371	0,368	0,544	0,761
α , β e σ^2													
	i. Intercepto	636,3	0,540	0,516	0,562	1,039	0,927	411,1	0,545	0,533	0,578	1,084	0,971
	ii. Educação	636,3	0,536	0,528	0,536	0,910	1,038	411,1	0,519	0,483	0,510	0,888	0,913
	iii. Experiência	636,3	0,594	0,659	0,697	1,415	1,210	411,1	0,570	0,590	0,640	1,260	1,066
	iv. Raça	636,3	0,538	0,512	0,559	1,030	0,920	411,1	0,535	0,512	0,556	1,028	0,935
	v. Iteração idade/educação	636,3	0,465	0,379	0,392	0,600	0,733	411,1	0,503	0,454	0,470	0,779	0,883
	vi. Sector de atividade	636,3	0,536	0,506	0,554	1,020	0,907	411,1	0,545	0,534	0,578	1,085	0,975
	vii. Formalidade	636,3	0,538	0,523	0,557	1,028	0,977	411,1	0,551	0,561	0,589	1,116	1,080
	viii. All betas	636,3	0,484	0,424	0,421	0,638	0,857	411,1	0,492	0,446	0,446	0,720	0,913
λ		722,9	0,502	0,434	0,475	0,803	0,772	465,4	0,503	0,439	0,471	0,781	0,800
λ, α e β		636,3	0,442	0,336	0,345	0,492	0,649	411,1	0,432	0,321	0,332	0,479	0,624
λ, α , β e σ^2		636,3	0,468	0,382	0,392	0,584	0,735	411,1	0,476	0,400	0,415	0,651	0,773
γ		1210,0	0,477	0,408	0,400	0,572	0,825	705,9	0,468	0,391	0,384	0,545	0,789
λ e γ		1306,8	0,464	0,382	0,375	0,526	0,769	809,2	0,456	0,369	0,363	0,506	0,742
λ, γ, α e β		636,3	0,428	0,322	0,315	0,421	0,654	411,1	0,415	0,300	0,297	0,396	0,608
λ, γ, α , β e σ^2		636,3	0,455	0,367	0,361	0,505	0,741	411,1	0,460	0,378	0,376	0,547	0,761
ψ e γ		1235,3	0,469	0,397	0,381	0,529	0,818	732,2	0,457	0,373	0,361	0,500	0,762
ψ, γ, α e β		636,4	0,441	0,346	0,333	0,447	0,717	411,1	0,431	0,328	0,319	0,425	0,674
ψ, γ, α , β e σ^2		636,4	0,465	0,391	0,378	0,532	0,808	411,1	0,474	0,405	0,395	0,573	0,828
ψ, λ e γ		1281,8	0,463	0,385	0,369	0,506	0,796	797,2	0,449	0,361	0,348	0,477	0,743
ψ, λ, γ, α e β		636,3	0,430	0,328	0,315	0,413	0,681	411,1	0,412	0,297	0,289	0,378	0,611
ψ, λ, γ, α , β e σ^2		636,3	0,455	0,373	0,359	0,496	0,772	411,1	0,457	0,374	0,365	0,523	0,764
ϕ		818,7	0,528	0,492	0,518	0,865	0,907	508,7	0,524	0,485	0,510	0,834	0,896
ϕ, ψ, λ, γ, α , β e σ^2		704,3	0,448	0,362	0,349	0,484	0,751	435,3	0,454	0,369	0,362	0,520	0,752

Fonte: PNAD 1999 e CPS March 2000

Table 6: Simulated Poverty and Inequality for Brazilian earnings in 1999, Using 1994 México coefficients.

		MEN						WOMEN					
		Mean p/c	Inequality					Mean p/c	Inequality				
		Income	Gini	E(0)	E(1)	E(2)	V(log)	Income	Gini	E(0)	E(1)	E(2)	V(log)
Brasil		636,2	0,517	0,467	0,511	0,906	0,837	410,3	0,507	0,449	0,486	0,831	0,818
México		636,3	0,498	0,432	0,492	0,925	0,765	411,1	0,466	0,416	0,387	0,565	0,944
α e β													
	i. Intercepto	636,2	0,517	0,467	0,511	0,906	0,837	410,3	0,507	0,449	0,486	0,831	0,818
	ii. Educação	636,2	0,500	0,435	0,470	0,804	0,795	410,3	0,459	0,368	0,384	0,585	0,709
	iii. Experiência	636,2	0,516	0,463	0,509	0,904	0,827	410,3	0,516	0,466	0,508	0,891	0,840
	iv. Iteração idade/educação	636,2	0,504	0,445	0,467	0,756	0,831	410,3	0,511	0,457	0,495	0,848	0,833
	v. Sector de atividade	636,2	0,519	0,471	0,514	0,911	0,847	410,3	0,513	0,469	0,497	0,847	0,886
	vi. Formalidade	636,2	0,539	0,509	0,563	1,052	0,890	410,3	0,520	0,470	0,520	0,934	0,831
	vii. All betas	636,2	0,500	0,431	0,469	0,794	0,776	410,3	0,490	0,421	0,449	0,745	0,793
α , β e σ^2													
	i. Intercepto	636,2	0,511	0,453	0,497	0,869	0,812	410,3	0,532	0,504	0,546	0,989	0,921
	ii. Educação	636,2	0,493	0,421	0,456	0,769	0,769	410,3	0,488	0,423	0,442	0,713	0,812
	iii. Experiência	636,2	0,509	0,449	0,494	0,867	0,802	410,3	0,541	0,521	0,568	1,057	0,942
	iv. Iteração idade/educação	636,2	0,497	0,431	0,453	0,723	0,806	410,3	0,536	0,512	0,554	1,008	0,935
	v. Sector de atividade	636,2	0,512	0,457	0,499	0,873	0,822	410,3	0,538	0,524	0,556	1,006	0,988
	vi. Formalidade	636,2	0,533	0,494	0,547	1,009	0,864	410,3	0,546	0,528	0,584	1,115	0,933
	vii. All betas	636,2	0,493	0,417	0,454	0,758	0,751	410,3	0,518	0,479	0,512	0,903	0,895
λ	15	657,0	0,508	0,449	0,491	0,854	0,805	439,2	0,519	0,477	0,506	0,857	0,882
λ, α e β	16	636,2	0,478	0,392	0,421	0,675	0,718	410,3	0,481	0,399	0,425	0,673	0,738
λ, α , β e σ^2	17	636,2	0,471	0,378	0,406	0,643	0,692	410,3	0,510	0,456	0,486	0,814	0,842
γ	18	912,5	0,523	0,486	0,499	0,803	0,916	615,4	0,514	0,479	0,471	0,703	0,950
λ e γ	19	926,6	0,525	0,493	0,501	0,794	0,940	736,0	0,516	0,495	0,467	0,673	1,022
λ, γ, α e β	20	636,2	0,493	0,426	0,439	0,686	0,809	410,3	0,487	0,421	0,421	0,623	0,830
λ, γ, α , β e σ^2	21	636,2	0,486	0,412	0,425	0,654	0,784	410,3	0,514	0,477	0,480	0,759	0,934
ψ e γ	22	922,4	0,517	0,479	0,484	0,771	0,924	628,2	0,504	0,465	0,444	0,637	0,949
ψ, γ, α e β	23	636,2	0,495	0,435	0,443	0,701	0,842	410,3	0,474	0,402	0,391	0,553	0,813
ψ, γ, α , β e σ^2	24	636,2	0,489	0,422	0,430	0,669	0,817	410,3	0,499	0,453	0,441	0,661	0,916
ψ, λ e γ	26	909,3	0,522	0,491	0,494	0,787	0,950	721,9	0,505	0,479	0,441	0,605	1,015
ψ, λ, γ, α e β	27	636,2	0,483	0,414	0,416	0,629	0,811	410,3	0,469	0,398	0,380	0,520	0,823
ψ, λ, γ, α , β e σ^2	28	636,2	0,477	0,401	0,404	0,600	0,786	410,3	0,494	0,449	0,429	0,624	0,926
ϕ	29	621,3	0,511	0,455	0,500	0,887	0,814	401,3	0,500	0,437	0,474	0,809	0,798
ϕ, ψ, λ, γ, α , β e σ^2	30	615,8	0,476	0,398	0,403	0,602	0,777	400,0	0,495	0,448	0,431	0,630	0,921

Fonte: PNAD 1999

Figure 4: Brazil-US Differences, Actual and Simulated, Steps 1 and 2

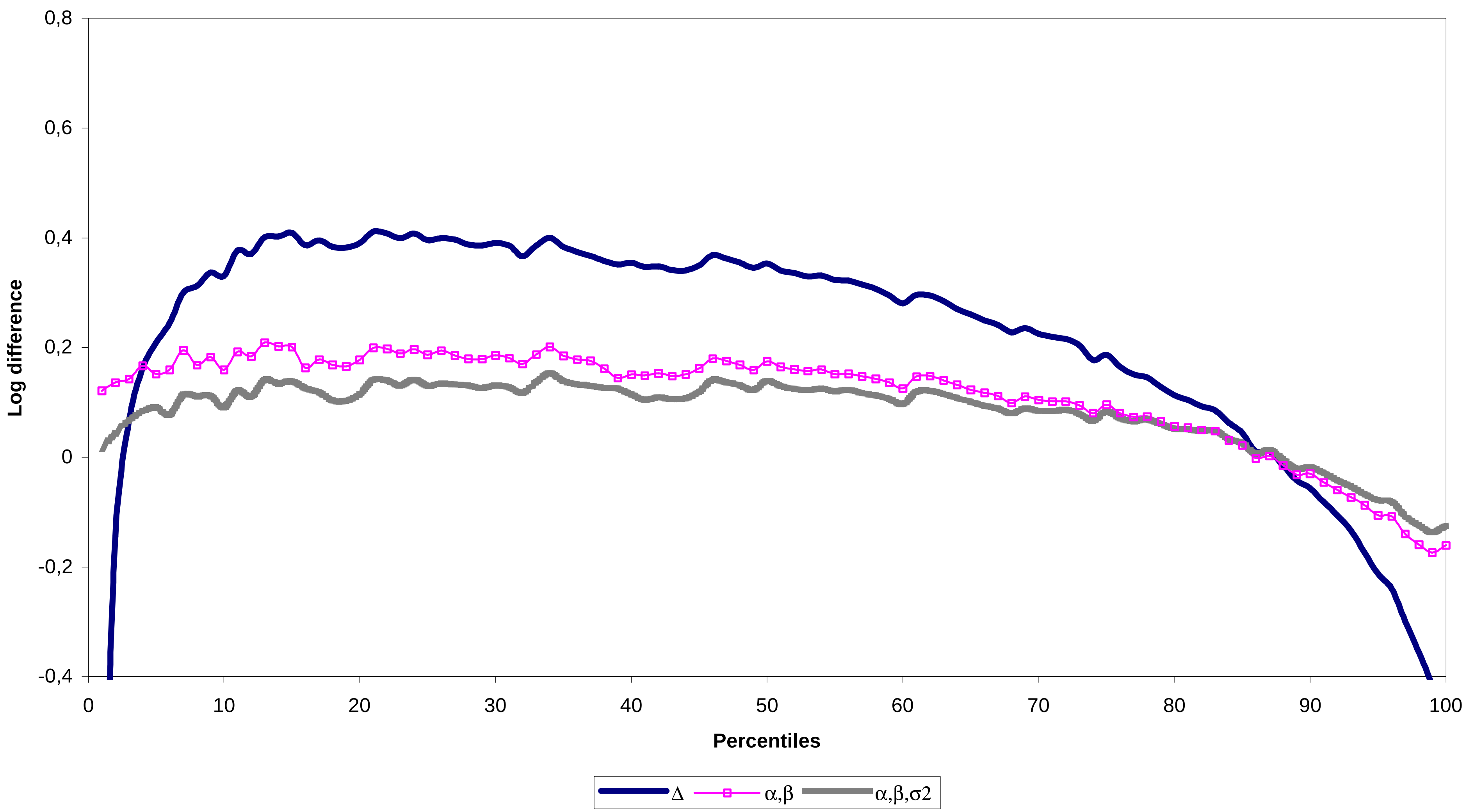


Figure 5: Brazil - US Differences, Actual and Simulated, Step 4

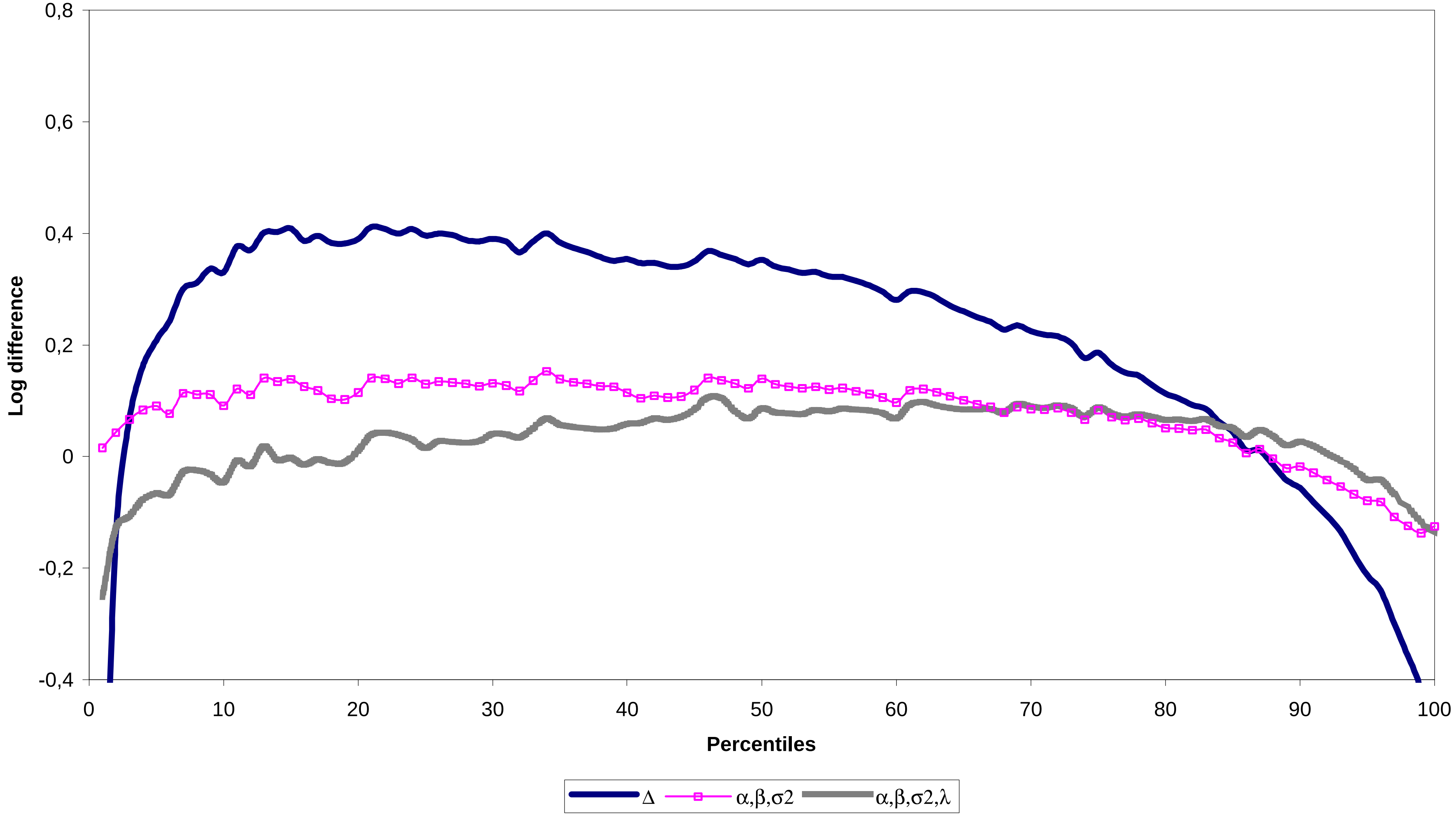


Figure 6: Brazil - US Differences, Actual and Simulated, Step 6.

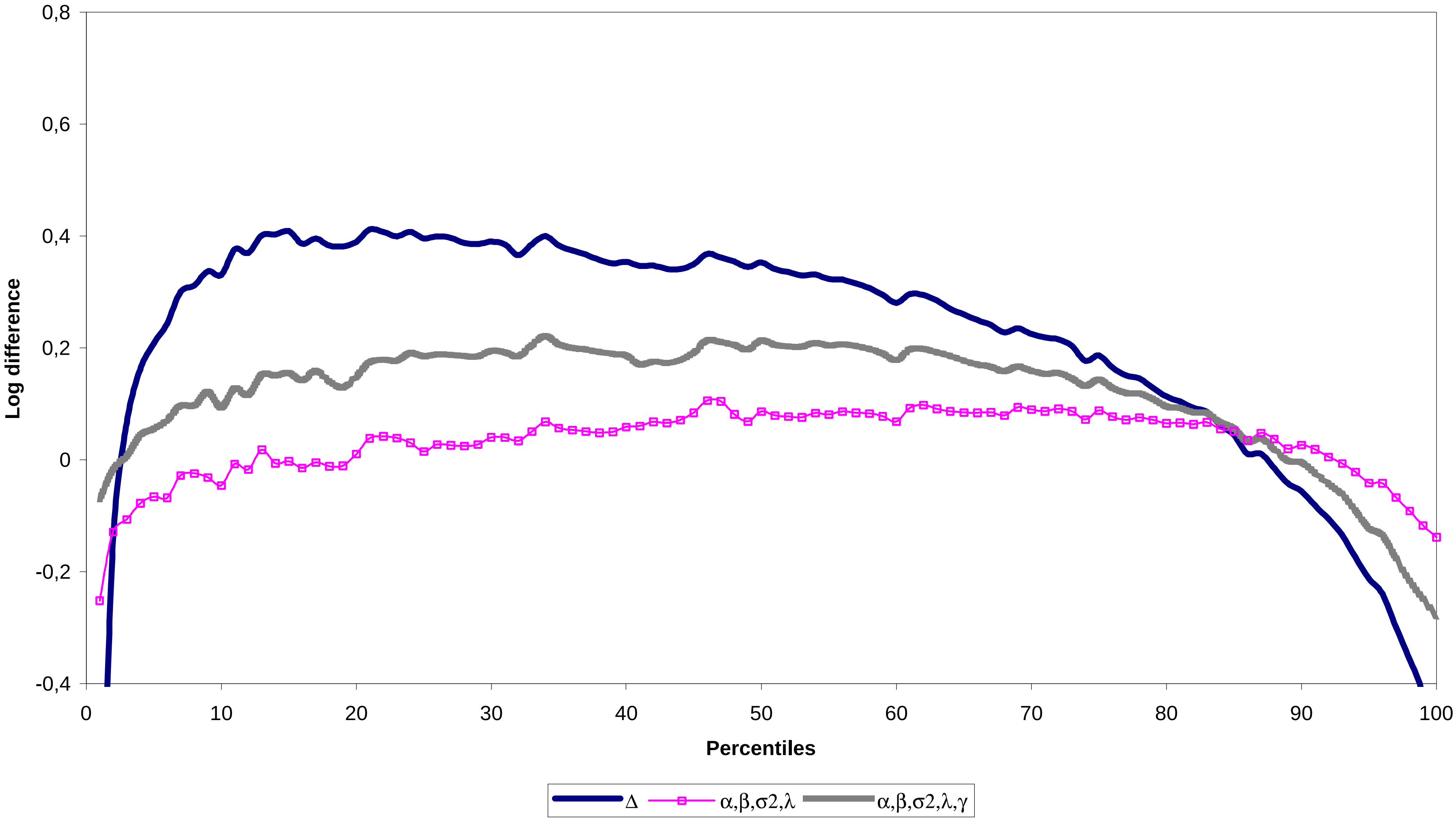


Figure 7: Brazil - US Differences, Actual and Simulated, Step 8

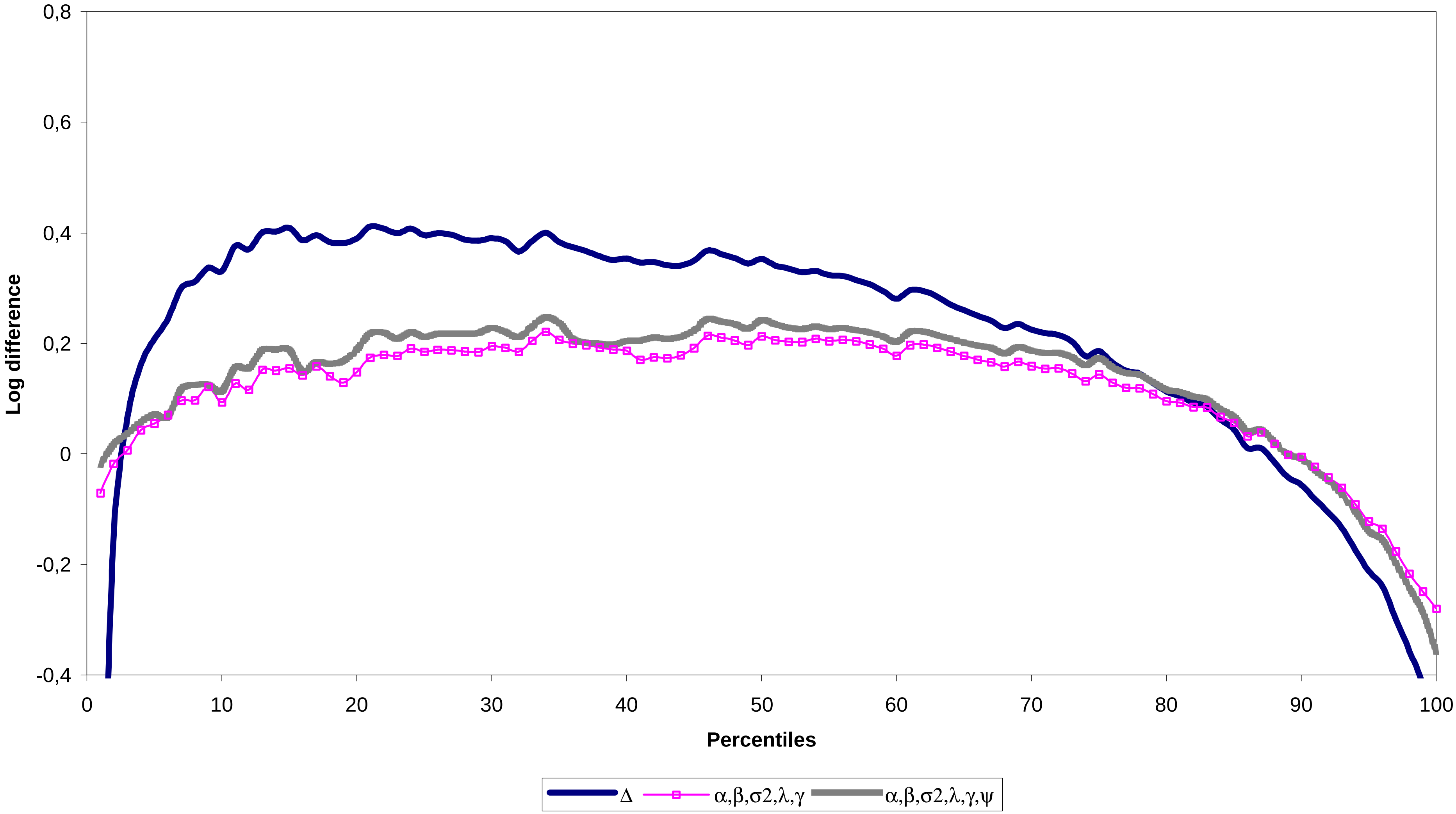


Figure 8: Brazil - US Differences, Actual and Simulated, Step 12

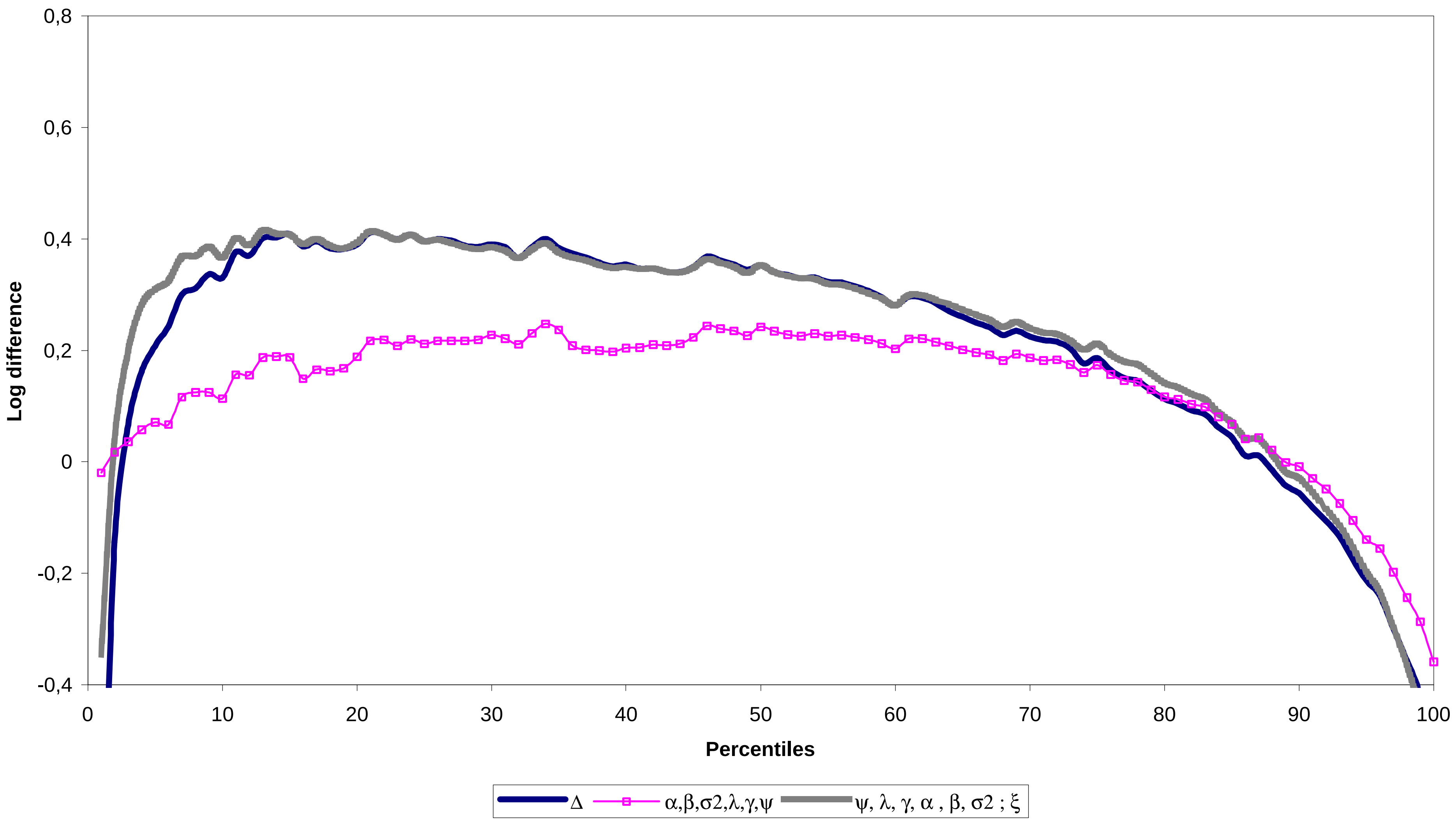


Figure 9: Brazil - Mexico Differences, Actual and Simulated, Steps 1 and 2

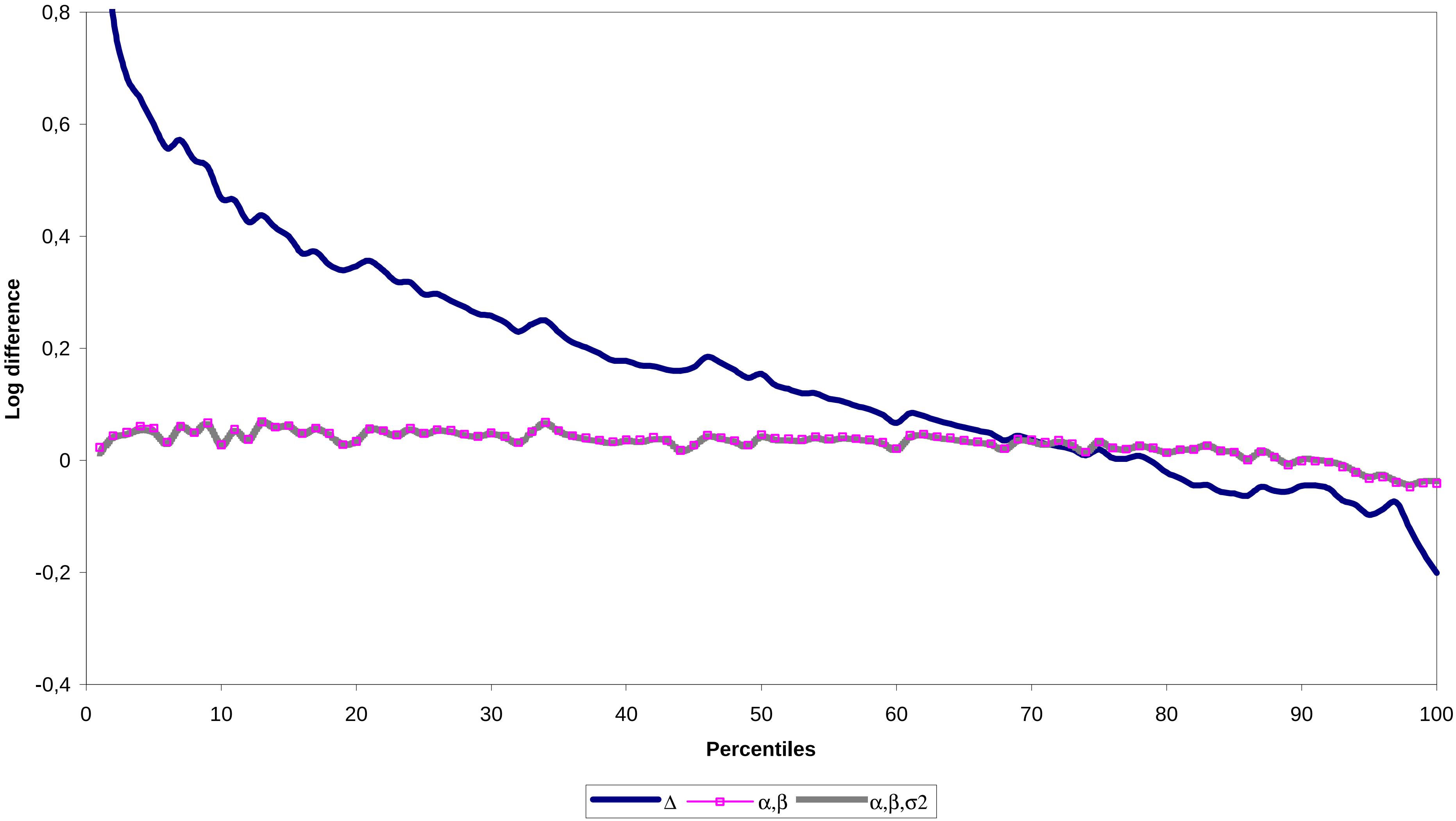


Figure 10: Brazil - Mexico Differences, Actual and Simulated, Step 4

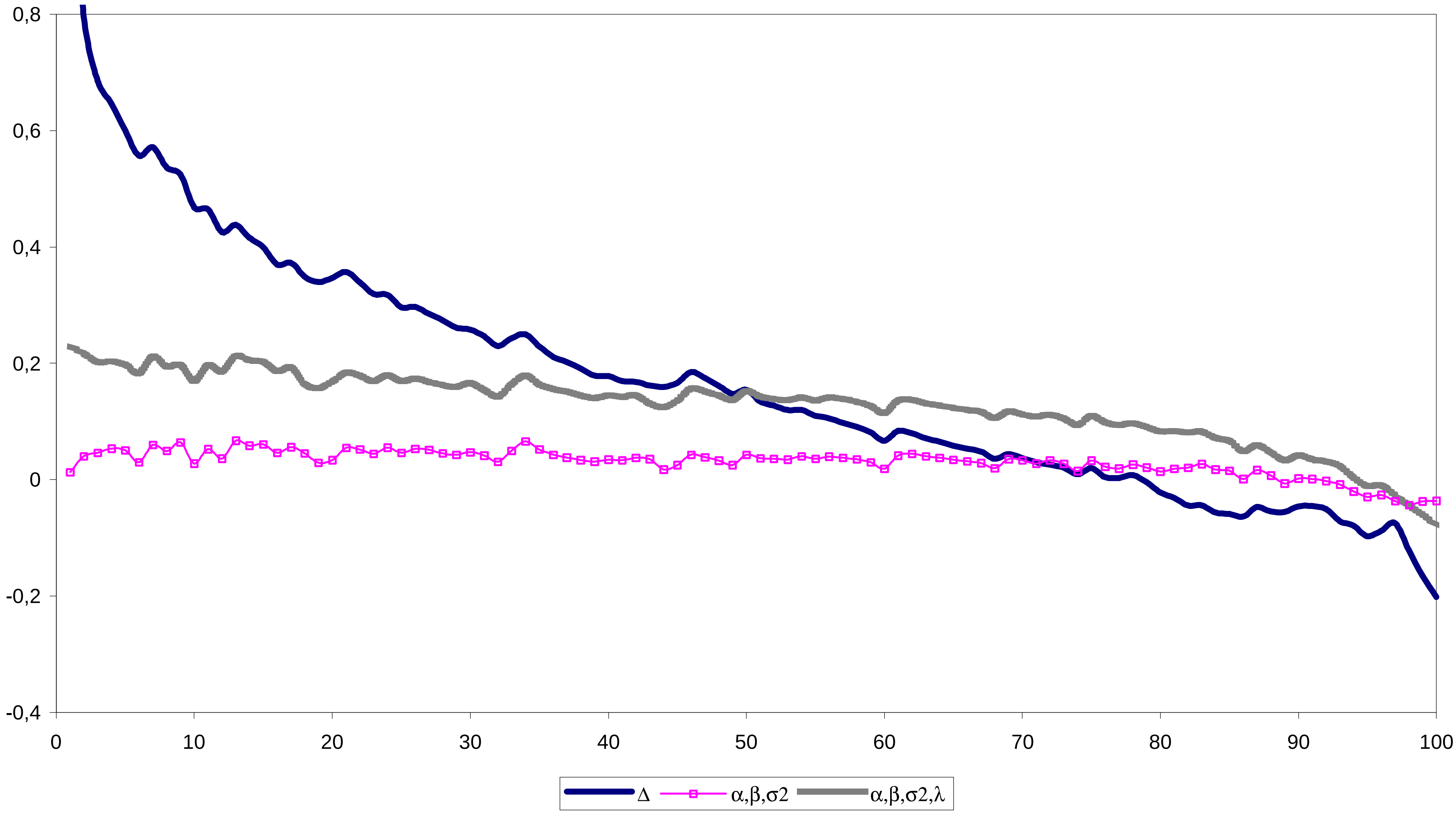


Figure 11: Brazil - Mexico Differences, Actual and Simulated, Step 6

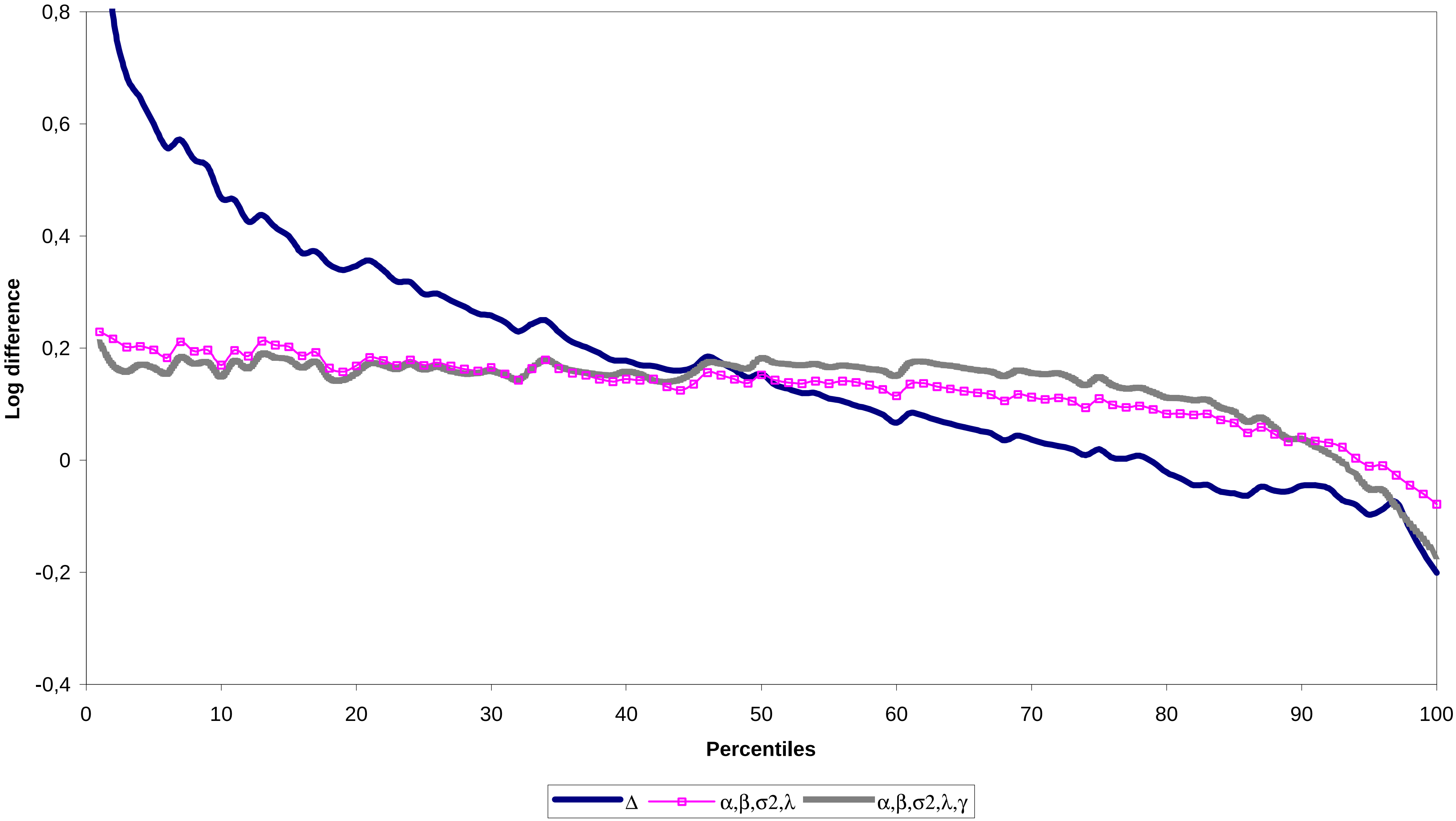


Figure 12: Brazil - Mexico Differences, Actual and Simulated, Step 8

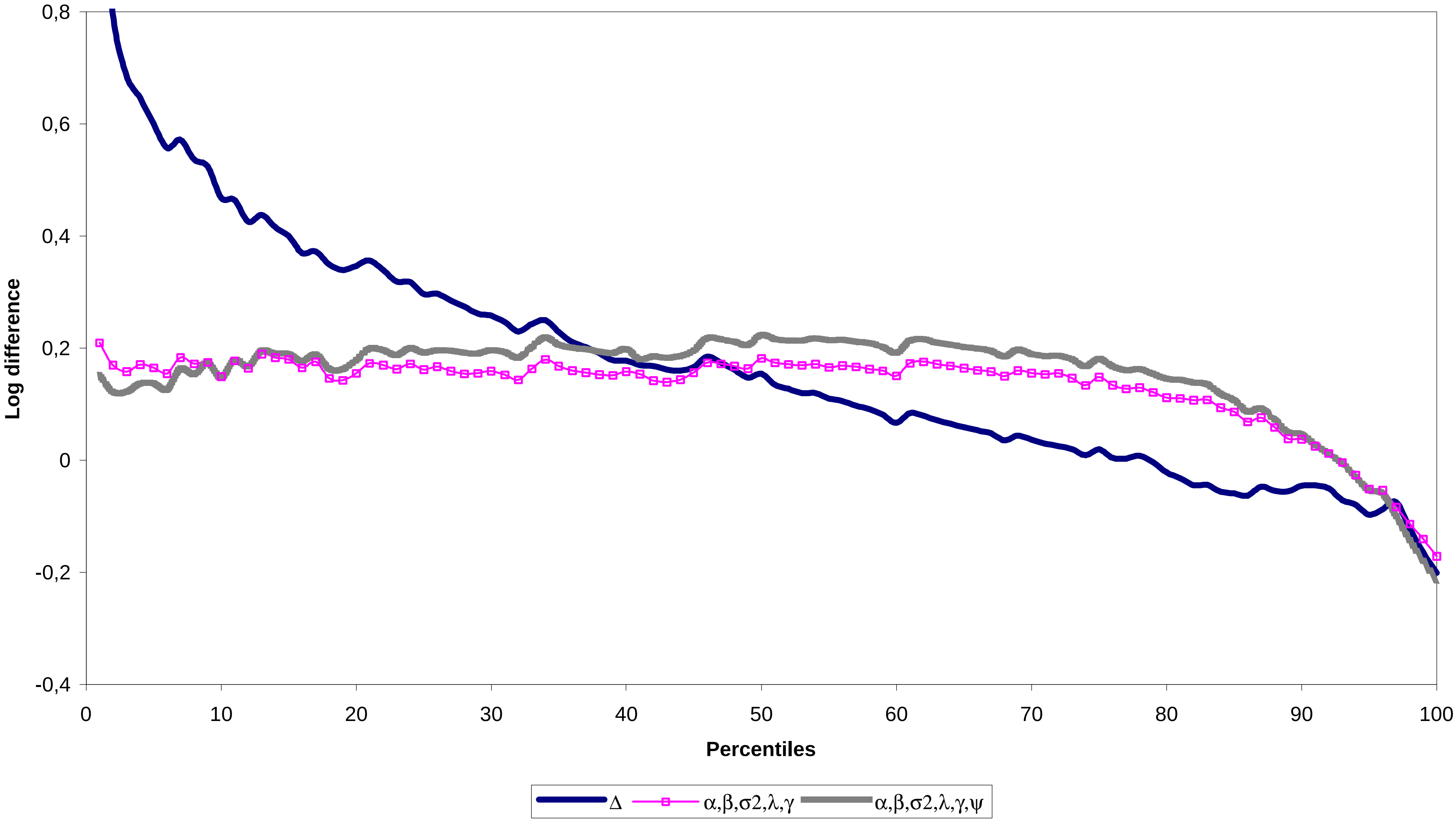


Figure 13: Brazil - Mexico Differences, Actual and Simulated, Step 12