United Nations University
WIDER
World Institute for Development Economics Research

Discussion Paper No. 2001/88

The Effects of Expanding Education on the Distribution of Income in Ceará

A Microsimulation

Francisco H.G. Ferreira*
and Phillippe George Leite*

September 2001

Abstract

Does more education really mean less poverty and less inequality? How much less? And what are the transmission mechanisms? This paper presents the results of a microsimulation exercise for the Brazilian State of Ceará, which suggests that broad-based policies aimed at increasing educational attainment would have substantial impacts on poverty reduction, but muted effects on inequality. While these results are highly dependent on assumptions about the behaviour of returns to education for the distribution of earnings, they are much more robust for the distribution of household income per capita. Over half of the poverty reducing effect of more education operate through greater incentives for labour force participation among the poor, and through reductions in fertility. Both of these effects operate largely through decisions made by poor women.

Keywords: education, poverty, inequality

JEL classification: C15, D31, I31, J13, J22

---

*Pontifícia Universidade Católica, Rio de Janeiro

This study is linked to the theme of Growth and Poverty discussed at the WIDER Development Conference, 25-26 May 2001, Helsinki. UNU/WIDER gratefully acknowledges the financial support from the governments of Denmark, Finland and Norway to the 2000-2001 research programme.

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

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

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

ISSN 1609-5774
ISBN 952-455-300-7 (printed publication)
ISBN 952-455-301-5 (internet publication)

1. Introduction

Ever since the introduction of the Human Capital model by Gary Becker and Jacob Mincer, economists have thought of earnings and income distributions as being fundamentally determined by the interaction between educational endowments and their market rates of return. In the specific case of Brazil, the seminal analysis of the country's income distribution by Carlos Langoni (1973) very much confirmed that view, and made education into the principal suspect in the search for culprits for the country's extreme levels of inequality. More recently, Barros et al. (2000) suggest that some 40 percent of overall inequality in the country's personal distribution of income can be ascribed to education.

It should presumably follow that if a government wishes to reduce poverty and inequality in Brazil, the first policy it ought to adopt should be a general expansion of education. Nevertheless, the historical evidence causes one to be less sanguine: in the United States, where 93 percent of the population reports nine or more years of schooling, income inequality has not been falling recently. The literature speaks of a changing structure of returns to education, whereby skill-biased technical progress (and in some contexts, possibly international trade) might be increasing demand for highly-educated workers, and offsetting (or more than offsetting) some of the equalizing results of expanding education. See Tinbergen (1975) for the classic reference, and Katz and Murphy (1992) for evidence on the US.

How might a substantial increase in the stock of education affect the income distribution in Brazil? In this paper, we simulate the impacts of a substantial expansion of education for the northeastern Brazilian state of Ceará. The simulation is carried out at the household level, using the complete Ceará sub-sample of the IBGE's 1999 Pesquisa Nacional por Amostra de Domicílios (PNAD). In addition to simulating the effects on earnings of people having more education to trade in the labour market, under different sets of assumptions about the evolution of returns, we also simulate the likely effects of additional education on labour force participation, occupational choice and fertility behaviour at the household level, and find that these matter a great deal for the overall picture.

As expected, the effects of a substantial educational expansion on poverty incidence are very substantial. Somewhat to our own surprise, however, we find the impact on inequality to be rather muted. Because of the changes in fertility and labour supply, we find that a very large part of the distributional changes arising from greater education depend on the behaviour of women. And location would matter more, rather than less: while we do not simulate the effects on migration, our simulated poverty profile indicates that of the (fewer overall) poor people, more (proportionately) would be in rural areas.

The paper is structured as follows. Section 2 describes the reduced-form model of the income distribution which was estimated. Section 3 describes the specific simulation exercises which were undertaken. Section 4 briefly highlights the main results and suggests some interpretations.

2. The model

In order to understand the impacts of different policies aimed at increasing educational endowments in the population of Ceará, we estimated a simple model of household income determination. The model builds on Ferreira and Paes de Barros (1999), which was in turn heavily influenced by Bourguignon et al. (1998) and Bourguignon et al. (2000).[1] This model—which is estimated on 1999 PNAD data for the state of Ceará—is recursive, and consists of five blocks, as follows:

Block I: household income aggregation

(1) $$Y_h = \sum_{i \in h} w_i L_i^w + \sum_{i \in h} \pi_i L_i^{se} + Y_{0h}$$

This equation simply adds up labour incomes for all household members, across the two sectors into which we assume the labour market is segmented: a wage sector (denoted by the superscript *w*) and a self-employment sector (denoted by the superscript *se*). *L* might have denoted hours, but given the nature of the information on labour supply in the PNAD data, it is actually a 0-1 participation dummy. Hence, w_i denotes the labour earnings of individual *i* in sector *w*, and π_i denotes the profits of individual *i* in the self-employment sector. The final term comprises all reported non-labour incomes accruing to the household.

Block II: earnings equation

(2) $$Log w_i = X_i^P \beta^w + \varepsilon_i^w$$

(3) $$Log \pi_i = X_i^P \beta^{se} + \varepsilon_i^{se}$$

Equations (2) and (3) are standard Mincerian earnings equations, estimated separately for the two labour market sectors. Both formal ('com carteira') and informal ('sem carteira') workers were treated as wage sector workers. Own account ('conta própria') workers were treated as self-employed. Employers were grouped alongside wage workers. Workers were assigned to the sectors of their principal occupation. The vector *X*, as is customary, contained characteristics both of the worker and of the job. In this case, *X* included years of schooling (year dummies), age, age squared, age*schooling, gender dummy, race (white, non-white), spatial (RM Fortaleza, other urban, rural) and sector (agriculture, services, industry). The results are reported in Table 1.

Block III: occupational choice

(4) $$P_i^s = \frac{e^{Z_i \gamma_s}}{e^{Z_i \gamma_s} + \sum_{j \neq s} e^{Z_i \gamma_j}} \quad \text{where } s, j = (0, w, se)$$

[1] See also Almeida dos Reis and Paes de Barros (1991) and Juhn, Murphy and Pierce (1993).

This block models the choice of occupation (into wage employment, self-employment or inactivity) by means of a discrete choice model—specifically, a multinomial logit—which estimates the probability of choice of each occupation as a function of a set of family and personal variables, namely: age, age squared, education, age*education, gender, race, spatial location, family composition, average age in the family (excluding the individual), average education in the family (excluding the individual), dummy if head of household, dummy if the head is inactive, dummy if spouse. Note that this occupational choice model is written in reduced form, as it does not include the wage rate (or earnings) of the individual (or of its family members) as explanatory variables. Instead, his or her productive characteristics (and the averages for the household) are included to proxy for earning potential. This approach is adopted to maintain the econometrics of joint estimation (with Block II) tractable. See Bourguignon et al. (1998) for a discussion.

Block IV: demographic choices

(5) $$ML\ (n_c \mid a, e, r, s, n_a)$$

This block uses a similar model to (4), which we now write in short form—ML stands for multinomial logit. This estimates the probability of choosing a certain number of children (0, 1, 2, 3, 4, 5+), as a function of the woman's age, education, race, spatial location, and the number of adults in the household.

Block V: educational choice

(6) $$ML\ (e \mid a, r, g, s)$$

This block applies the same multinomial logit estimation approach to the choice of level of education (into categories of years of schooling) of individuals, as a function of their age, race, gender and spatial characteristics.

Note that we interpret equations (2)-(6) merely as descriptions of the statistical associations present in the data, under some maintained assumptions about the form of the relevant joint multivariate distributions.

Table 1
The earnings equations for Ceará, 1999

Mincerian Equation **Estimated Coefficients**

		Earnings				Self-employed			
		R^2	coef	std	p-value	R^2	coef	std	p-value
		0,60				0,49			
Intercept			3,74	0,14	0,00		4,09	0,36	0,00
Education									
	0		-0,64	0,13	0,00		-2,07	0,35	0,00
	1 to 3		-0,65	0,13	0,00		-1,81	0,35	0,00
	4		-0,59	0,13	0,00		-1,61	0,36	0,00
	5		-0,61	0,15	0,00		-1,70	0,41	0,00
	6		-0,38	0,17	0,03		-1,23	0,45	0,01
	7		-0,59	0,15	0,00		-1,15	0,45	0,01
	8		-0,31	0,14	0,03		-1,04	0,40	0,01
	9 to 12		-0,44	0,12	0,00		-0,92	0,37	0,01
	13 or more		-	-	-		-	-	-
Age			0,09	0,01	0,00		0,09	0,01	0,00
Age^2			0,00	0,00	0,00		0,00	0,00	0,00
Age * education									
	0		-0,03	0,00	0,00		**-0,01**	**0,01**	**0,44**
	1 to 3		-0,03	0,00	0,00		**-0,01**	**0,01**	**0,25**
	4		-0,03	0,00	0,00		**-0,01**	**0,01**	**0,29**
	5		-0,03	0,00	0,00		**0,00**	**0,01**	**0,90**
	6		-0,03	0,01	0,00		**-0,01**	**0,01**	**0,31**
	7		-0,03	0,00	0,00		**-0,02**	**0,01**	**0,17**
	8		-0,03	0,00	0,00		**-0,01**	**0,01**	**0,22**
	9 to 12		-0,02	0,00	0,00		**0,00**	**0,01**	**0,64**
	13 or more		-	-	-		-	-	-
Race - White			0,11	0,02	0,00		0,14	0,04	0,00
Gender - Male			0,46	0,02	0,00		0,94	0,04	0,00
Metropolitan area			0,45	0,03	0,00		0,30	0,06	0,00
Urban non metropolitan			0,12	0,03	0,00		0,11	0,05	0,05
Sector of activity									
	Agriculture		-0,17	0,04	0,00		-0,67	0,06	0,00
	Industry		0,07	0,02	0,00		-0,10	0,05	0,04

Fonte: PNAD 1999

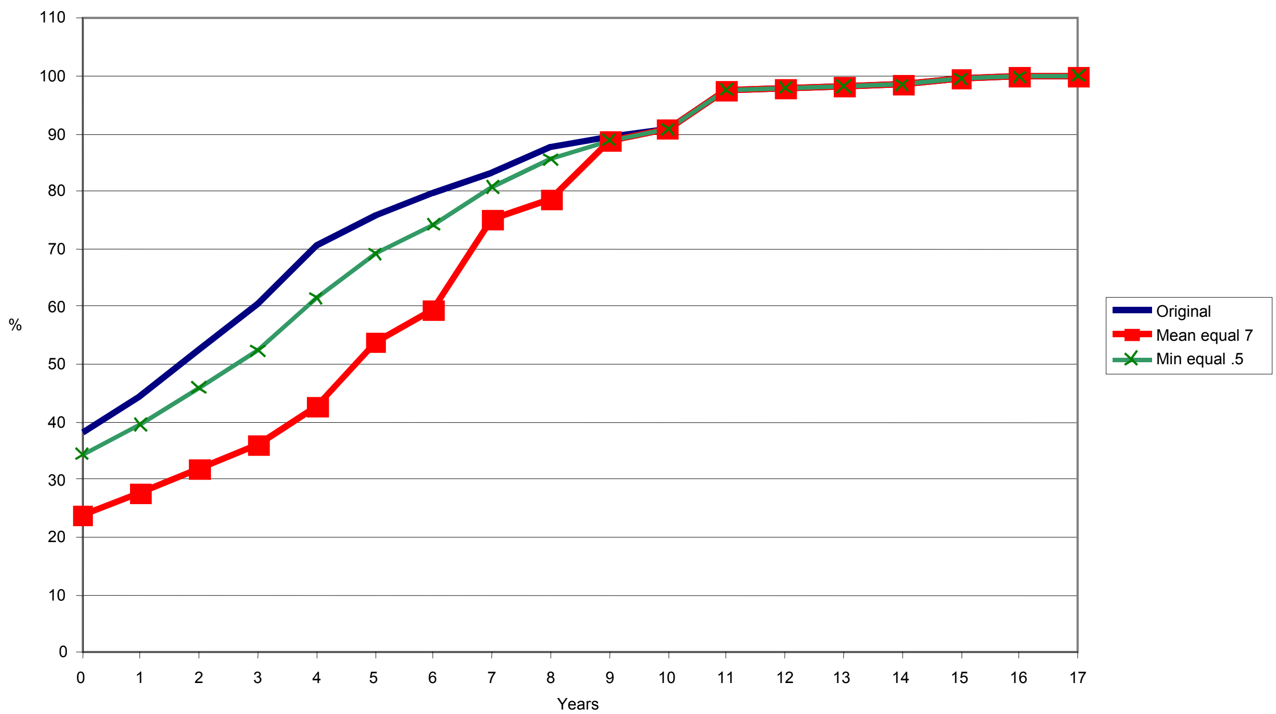
Source: authors' calculations from the PNAD 1999 dataset.[2]

[2] Pesquisa Nacional por Amostragem de Domicílios (PNAD), microdata.

3. The simulations

Educational expansions are not, of course, all alike. One would expect to obtain very different distributional results from two policies, one of which aimed to triple the number of university graduates in the state, and another which aimed to halve the number of illiterate people. How exactly the histogram of the distribution of years of schooling changes matters almost as much as how the overall mean evolves. In addition—and as alluded to above—the same expansion in education will have different impacts depending on how demand for skills changes in the labour market. To allow for both of these concerns to the extent possible, six simulations were undertaken, corresponding to two different 'policy choices', with different aims in terms of the distribution of education; and to three sets of assumptions about returns in the labour market.

Figure 1
CDRs of years of schooling in Ceará: actual and simulated



The first 'policy' was one of indiscriminate expansion. We simulate this as a rise in the mean of the distribution of years of schooling, from 4.5 (the observed level in 1999), to seven years (which the government of Ceará regards as a feasible target for a ten- or twenty-year plan). Of course, one might raise the mean of a distribution in very different ways. Since we observe how educational attainment is distributed jointly with age, gender, race and spatial location in the state, through our estimation of equation (6) above, we simulate the expansion in a manner consistent with that pattern. Specifically, for each individual, a rise in the constant term in (6) generates a simulated level of schooling e*, different from the actual level of e observed in 1999. In the simulation, we replace e by max (e, e*).[3] By shifting the distribution in this manner, without altering the other (gamma) coefficients in (6), we preserve the observed conditionality of educational choices on other characteristics.

The second ‘policy’ we investigate is a focused effort to reduce illiteracy. We change the distribution of education by moving *fifty percent* of those with less than four years of schooling to five years (exactly), by selecting those with the highest probability of moving rom amongst all possible candidates. As in (a), the constant term in (6) is raised, and max (e, e*) is chosen, until the mass of individuals with $e < 5$ is reduced by fifty percent, falling from an observed 46.1 percent in 1999 to a simulated 23.9 percent. The original cumulative distribution function of years of schooling in Ceará in 1999, as well as the two simulated distributions are shown in Figure 1.

The results of each of these two educational ‘policies’ are simulated under three alternative returns scenarios, namely:

1. β_{99} : Keep all βs values as estimated for the 1999 regressions.
2. β convex: with respect to category 13+ (omitted), lower β for 0-4 years of schooling by 20 percent; for 5-8 by 15 percent; and for 9-12 by 10 percent.
3. β concave: with respect to category 13+ (omitted), raise β for 0-4 years of schooling by 30 percent, for 5-8 by 20 percent and 9-12 by 10 percent.

So the six simulations are given by the following schematic 2x3 matrix:

Simulation	β (1999)	β concave	β convex
First Policy			
Second Policy			

In all six exercises, we also included a change in the constant term of equations (2) and (3), to simulate a returns-neutral growth rate of 2.5 percent p.a. on average, over eleven years, so that the simulations might be given the concrete interpretation of changes between the years of 1999 and 2010, in Ceará.

[3] Max (e, e*) is used, instead of e*, because the ML sometimes generates falls in educational attainment, mainly for those with very high schooling. This is largely due to the fact that some key determinants of educational choices, such as parental income or wealth, are omitted from (6). We do not resolve this problem, and explicitly warn the reader against interpreting (6) causally. For the purposes of the simulation, however, preventing declines in individual levels of education alleviates the main undesired consequence of the incomplete specification of (6).

Table 2
Distribution of individual earnings

Simulated Poverty and Inequality

	β_{99}						
	Mean p/c Income	Inequality					Pop
		Gini	E(0)	E(1)	E(2)	V(log)	
Ceara	413,7	0,542	0,532	0,627	1,532	0,930	1.673.243
1 - mean of education equal 7							
α , β e σ^2	469,3	0,523	0,489	0,560	1,250	0,876	1.673.243
γ, α , β e σ^2	450,4	0,525	0,492	0,568	1,287	0,872	1.854.771
ψ, γ, α , β e σ^2	449,3	0,523	0,489	0,561	1,260	0,872	1.852.698
2 - w=0.50 | education less than 4 | age between 15 - 40							
α , β e σ^2	428,8	0,536	0,519	0,609	1,455	0,909	1.673.243
γ, α , β e σ^2	427,4	0,537	0,520	0,608	1,438	0,912	1.746.710
ψ, γ, α , β e σ^2	427,0	0,536	0,517	0,608	1,452	0,902	1.744.842

	$\beta_{concave}$						
	Mean p/c Income	Inequality					Pop
		Gini	E(0)	E(1)	E(2)	V(log)	
Ceara	480,3	0,521	0,480	0,556	1,229	0,842	1.673.243
1 - mean of education equal 7							
α , β e σ^2	520,5	0,511	0,462	0,523	1,094	0,828	1.673.243
γ, α , β e σ^2	498,9	0,513	0,464	0,530	1,126	0,823	1.854.771
ψ, γ, α , β e σ^2	497,7	0,512	0,461	0,524	1,103	0,823	1.852.698
2 - w=0.50 | education less than 4 | age between 15 - 40							
α , β e σ^2	491,2	0,518	0,474	0,548	1,198	0,834	1.673.243
γ, α , β e σ^2	488,2	0,519	0,475	0,547	1,186	0,837	1.746.710
ψ, γ, α , β e σ^2	487,8	0,518	0,473	0,547	1,196	0,828	1.744.842

	β_{convex}						
	Mean p/c Income	Inequality					Pop
		Gini	E(0)	E(1)	E(2)	V(log)	
Ceara	377,5	0,563	0,583	0,687	1,775	1,023	1.673.243
1 - mean of education equal 7							
α , β e σ^2	435,5	0,534	0,515	0,593	1,386	0,925	1.673.243
γ, α , β e σ^2	418,6	0,536	0,518	0,599	1,412	0,925	1.854.771
ψ, γ, α , β e σ^2	417,2	0,535	0,516	0,594	1,394	0,927	1.852.698
2 - w=0.50 | education less than 4 | age between 15 - 40							
α , β e σ^2	393,4	0,554	0,562	0,660	1,662	0,990	1.673.243
γ, α , β e σ^2	392,3	0,555	0,561	0,658	1,636	0,987	1.746.710
ψ, γ, α , β e σ^2	392,5	0,554	0,561	0,657	1,639	0,986	1.744.842

Fonte: Pnad 99

Source: authors' calculations from the PNAD 1999 dataset.

Table 3
Distribution of household incomes: simulated poverty and inequality

	Mean p/c Income	Inequality					Pop	Poverty Z =68		
		Gini	E(0)	E(1)	E(2)	V(log)		P(0)	P(1)	P(2)
Ceara	158,6	0,619	0,757	0,845	2,334	1,475	6.978.378	45,8	21,4	13,6
ψ	141,0	0,594	0,679	0,796	2,240	1,278	6.978.378	47,8	21,1	12,7
1 - mean of education equal 7										
α , β e σ²	171,8	0,605	0,733	0,784	2,027	1,513	6.978.378	40,9	19,4	12,4
γ, α , β e σ²	169,8	0,581	0,669	0,715	1,780	1,392	6.978.378	38,3	17,7	11,1
ψ, γ, α, β e σ²	178,2	0,564	0,623	0,667	1,607	1,308	6.978.378	34,7	15,1	9,2
2 - w=0.50 | education less than 4 | age between 15 - 40										
α , β e σ²	162,2	0,614	0,748	0,827	2,247	1,480	6.978.378	44,2	20,7	13,1
γ, α , β e σ²	157,1	0,597	0,703	0,772	2,025	1,414	6.978.378	43,5	19,9	12,5
ψ, γ, α, β e σ²	159,0	0,591	0,688	0,757	1,979	1,388	6.978.378	42,2	19,1	11,9

Tabela : Households 2010 - β concavo

	Mean p/c Income	Inequality					Pop	Poverty Z =68		
		Gini	E(0)	E(1)	E(2)	V(log)		P(0)	P(1)	P(2)
Ceara	174,5	0,607	0,739	0,788	2,019	1,526	6.978.378	40,4	19,3	12,4
ψ	136,5	0,608	0,715	0,831	2,372	1,342	6.978.378	51,1	23,4	14,5
1 - mean of education equal 7										
α , β e σ²	184,0	0,600	0,730	0,754	1,848	1,560	6.978.378	38,1	18,3	11,8
γ, α , β e σ²	182,5	0,577	0,667	0,689	1,623	1,436	6.978.378	35,7	16,6	10,5
ψ, γ, α, β e σ²	191,7	0,560	0,621	0,643	1,466	1,349	6.978.378	32,2	14,2	8,7
2 - w=0.50 | education less than 4 | age between 15 - 40										
α , β e σ²	177,1	0,605	0,736	0,778	1,973	1,532	6.978.378	39,6	19,0	12,2
γ, α , β e σ²	172,2	0,589	0,693	0,727	1,779	1,466	6.978.378	38,8	18,2	11,6
ψ, γ, α, β e σ²	174,3	0,583	0,679	0,713	1,738	1,440	6.978.378	37,6	17,4	11,0

Tabela : Households 2010 - β convexo

	Mean p/c Income	Inequality					Pop	Poverty Z =68		
		Gini	E(0)	E(1)	E(2)	V(log)		P(0)	P(1)	P(2)
Ceara	150,1	0,628	0,775	0,885	2,545	1,460	6.978.378	49,1	23,3	14,7
ψ	136,5	0,608	0,715	0,831	2,372	1,342	6.978.378	51,1	23,4	14,5
1 - mean of education equal 7										
α , β e σ²	163,8	0,609	0,738	0,808	2,167	1,485	6.978.378	42,8	20,4	12,9
γ, α , β e σ²	161,4	0,585	0,673	0,735	1,897	1,367	6.978.378	40,5	18,7	11,6
ψ, γ, α, β e σ²	169,3	0,568	0,627	0,688	1,722	1,284	6.978.378	36,5	16,0	9,7
2 - w=0.50 | education less than 4 | age between 15 - 40										
α , β e σ²	153,8	0,622	0,762	0,862	2,435	1,461	6.978.378	47,0	22,3	14,0
γ, α , β e σ²	148,4	0,605	0,715	0,804	2,195	1,394	6.978.378	46,8	21,6	13,4
ψ, γ, α, β e σ²	150,3	0,598	0,699	0,787	2,138	1,369	6.978.378	45,3	20,7	12,7

fonte: Pnad 99

Source: authors' calculations from the PNAD 1999 dataset.

4. Results

The main simulation results are presented in Tables 2 and 3. Table 2 reports the mean income and five different inequality measures, for each of these six simulations on the distribution of labour earnings, among earners with positive labour incomes.[4] Table 3 presents the corresponding results for the distribution of household incomes by individuals and includes, in addition to the same inequality measures as in Table 2, three poverty measures—P(α), for α = 0, 1 e 2. The poverty line was R$68.00, which is the line officially suggested by the Planning Institute of the State Government of Ceará, (IPLANCE). In each of the above tables, the measures presented in the row 'Ceará' *of the panel* β_{99} are those for the actual observed distribution in 1999. The measures presented in the row 'Ceará' in the other two panels arise from imposing the simulated structure of returns (more concave or more convex) on the existing 1999 population—with its actual distributions of education and other characteristics, whether observed or unobserved.[5]

For each of the six combinations of policies and returns, poverty and inequality statistics are presented for three different simulations, denoted by sets of Greek letters. The first of these, denoted by α, β and σ^2 , consists of running the required simulation—of the first or of the second policy—and feeding the simulated distribution of education through the earnings models (2) and (3), either unadjusted (β_{99}), or adjusted (β_{convex} or $\beta_{concave}$). Original residuals are used, and this generates a simulated distribution of earnings, under the required assumption about returns, which corresponds to the new distribution of education. This was in turn obtained from simulating an increase in schooling according to the education discrete choice model in (6). In this simulation, each individual preserves his or her initial (1999) occupation and family composition. All that may change is the amount of education they sell in that market and, for the convex and concave scenarios, the rate at which they do so.

We know, however, that labour force participation and occupational choice are also heavily dependent on education. It is natural to suppose, then, that changes in schooling endowments such as the ones being simulated here for Ceará, are likely to have some impact on who is working, and on where they are working. This is investigated by allowing the simulated distributions of education to feed through the occupational choice model (4), the parameters of which are denoted by γ's. The second row in each cell thus summarizes the inequality and poverty statistics pertaining to the distributions which are simulated when, in addition to the educational endowment being transacted and to the structure of returns, we allow for occupational choices and labour force participation to change.[6]

Finally, the third row allows for family size—driven by the number of children 'demanded' by each family—to change also. This is achieved by allowing the simulated

---

[4] Simulated populations are also included, as a control on changes in participation.

[5] These rows retain their original α's, so that they can be interpreted as what the distribution might look like today, if only the relative returns (β's) had changed.

[6] In order to simulate the earnings of new entrants into the labour force, each needs to be allocated to a sector of activity. We make the conservative assumption that all new entrants are in the occupation with the lowest mean income, namely agriculture.

distributions of education to feed through the demographic choice model (5), the parameters of which are denoted by ψs. This has two second-round effects on household incomes: first, as the number of children in a family changes, the income per capita denominator changes, and it is recalculated accordingly. Second, number of children is, as it must be, an independent variable in the occupational choice multi-logit model (4). In this row of simulations results the γs and ψs interact, since changes in occupational choice reflect not only chances in the educational levels of the individuals (and others in their families) but also changes in the number of under-16s living in the household.

While the aggregated information presented in Tables 2 and 3 tell the basic story, additional insight into the effects of the simulated educational expansions can be gained from looking at the individual impacts on the mean incomes of each percentile of the distribution. Figures 2-13 plot the difference in the logarithms of mean incomes for each percentile, between the simulated distribution and the real 1999 distribution: figures 2-7 refer to the earnings distribution, while figures 8-13 correspond to the distributions of household per capita income. Each distribution is ranked by its own distributed variable. The lines for α, β and σ^2 correspond to simulations where each earner had his or her level of education changed to a level drawn for it in the new distribution of education, as described above. To simulate the concave and convex cases, the βs were changed as appropriate.

As indicated above, in these simulations, people are selling more education on the labour market, but are still working in the same occupation as before, and have exactly the same family composition. The lines that include a γ simulate the additional effect of those changes in years of schooling on people's labour force participation and/or occupational choices. And those that include a ψ as well, also incorporate the effect of those extra years of schooling on the number of children each family is likely to have, and any subsequent additional impact that may have on occupational choice.

4.1 Effects on earnings

Policy One, which consisted of raising mean years of schooling in Ceará from 4.5 to seven, in a manner which was consistent with individual propensities to acquire education, turns out to be both (i) income-increasing and (ii) generally equalizing. The rise in mean earnings can be seen from a comparison of the simulated mean under Policy One, with the 'Ceará' mean, in Table 2. In fact, mean incomes are higher than the actual 1999 mean (R$413.70) for all simulations, in all three returns scenarios.[7] However, the differences in income gains are quite marked across the returns scenarios. This is particularly evident from inspection of Figures 2-4: whereas the educational expansion would result in very large gains indeed for the very poor if returns to the low skilled rise, they would leave them hardly better off if returns convexify through a fall at the bottom rungs of the schooling ladder. The effects on inequality are naturally also driven by this behaviour with respect to returns. When compared to the observed earnings Gini in 1999, of 0.542, the fall arising from the educational expansion is of less than one Gini point if returns become more convex, of about two points if they remain unchanged, and of some three points if they

[7] Note, however, that all simulations do include the effects of exogenous economic growth at an average rate of 2.5 percent p.a. for eleven years, in addition to the change in educational endowments.

become more concave. The other inequality measures confirm this assessment, and the magnitude of the inequality changes are similarly unimpressive.

Finally, the impact of this large expansion in education on participation and demographic behaviour does not translate into changes in inequality in the earnings distribution. This can be seen both from looking at the inequality measures across the simulations that incorporate occupational choice (γ), and demographic effects (ψ) of greater education, with the basic earnings simulation for each policy-return configuration, as well as from the Figures 2-4, where all three lines lie very close together. It would be wrong to conclude, however, that limited impact on aggregate earnings inequality equals no impact at all. The labour force expanded by almost two hundred thousand people each time the educational effect on occupational choice was taken into account. But, as Figure 14 illustrates, the composition of the net entrants into the labour force is biased towards the poor (in household income terms)—who are gaining most education—and so this lowers overall mean earnings, and have ambiguous effects on overall earnings inequality.[8]

The effects of Policy Two, which consisted of a targeted effort at reducing illiteracy, by halving the proportion of persons with four years of schooling or less, were similar in nature, but considerably more muted. The rows for simulations under Policy Two in Table 2 reveal smaller increases in mean earnings and smaller reductions in inequality than for Policy 1. Figures 6 and 7 show that, for this particular policy, the configuration of returns is absolutely crucial: if returns to the unskilled rise, then the impact of having a little more education on the welfare of those who at the bottom of the distribution will be positive and substantial. Most people in the bottom quintile of the distribution would have between 20 percent and 50 percent higher earnings. If, on the other hand, Policy Two is combined with a decline in the returns to lower levels of schooling, as in Figure 7, then it will patently not be sufficient to reverse the impoverishing effect of the change in returns.

4.2 Effects on household incomes

Turning to the effects on household incomes, we are faced immediately with two very large differences from the earnings distribution. First, occupational choice and demographic effects matter a great deal more. Second, changes in the (exogenous) structure of returns matter a great deal less. Both have the same fundamental explanation: the individuals who are considered alone in the earnings distribution, are not organized into families in a random manner. Hence, many of the entrants into the labour force turn out to reside in poor families (as seen in Figure 14, bearing in mind footnote 7), and their new labour supply becomes highly equalizing in the distribution of household incomes. It is also chiefly among the poor that the effect of more schooling on fertility—to reduce the number of children in the household, thus raising per capita incomes—is particularly pronounced.

It is thus that Figures 8-10 have the curves with γ indicating larger income increases for the poor than the basic earnings simulation, and those with both γ and ψ, higher still. The sharp downward turn in these log-income difference curves for the top decile should signal a substantial equalizing effect. And this is confirmed by inspecting the Gini coefficients, or other measures, in Table 3: from an observed 1999 level of 0.62, the Gini could fall by a

[8] Note that Figure 14 is ranked by the percentiles of the distribution of household incomes.

full 6 points if returns helped (and became more concave). More surprisingly, even if returns convexified—which we saw was a powerful unequalizing force on the distribution of earnings—we would still see a fall in Gini of five points, after all effects are taken into account. If returns were identical to those of 1999, the fall would be midway between the two extremes, with the simulated Gini at 0.564. This illustrates the two points made earlier: not only do occupational and demographic decisions by households acquire greater importance when we look at the household income distribution, but also the different returns scenarios generate less variation in outcomes than was the case for earnings. Households pool resources, and provide insurance to individual members: even if assortative mating is very pronounced in Brazil[9] (and we suspect, in Ceará), education levels still do differ across individuals in the same household, so that changes in returns hurt or benefit the pooled family less than it might hurt or benefit each member.

The combination of rising mean incomes and falling inequality should spell good news for poverty reduction, as a result of educational expansion Policy One. Indeed, with respect to the state's poverty line of R$68.00 per capita per month, we observe declines in poverty headcount (or incidence) as large as 13.6 percentage points, with returns flattening out. But once again, the structure of returns matters less here, so even if they convexified, Policy One would still lower P(0) from the 1999 level of 45.8 percent, to 36.5 percent. Each of these results take into account all simulated effects of the greater endowment of education, and in particular its labour supply, occupational and demographic impacts. Their importance is once again highlighted by the fact that, in their absence, the poverty reduction effect of the educational expansion would be approximately halved. Specifically, with constant returns, the fall excluding these effects would be of five percentage points, rather than eleven.

Policy Two turns out to be a weaker substitute for Policy One. All of its effects have the same sign, but all have weaker magnitudes. Mean incomes still rise (although, if returns convexify, only for about the poorest 40 percent of the population). Inequality falls, although by less. And poverty also falls, as you would expect, but by much less, and with a greater sensitivity to what happens to returns: if they convexify, the headcount would fall a mere 0.5 percentage point. At the other extreme, with returns rising for the low-skilled, it would fall eight points.

Table 4 below shows the effects of these policies on the composition of the poor, rather than just on their level. The first panel presents a simple poverty profile, as observed in Ceará in the 1999 PNAD. The profile is by location, gender of the household head and schooling of the household head. The P(0) column indicates incidence in the subgroup, while the 'perfil' column indicates contribution of that subgroup to overall poverty. The second panel presents the simulated profile for Policy One, under each of the three different returns scenarios.

[9] The simple correlation coefficient between the years of schooling reported by household heads and their spouses in the 1999 PNAD (for all of Brazil) is 0.73, which compares with 0.63 in the US, for instance.

Table 4
Actual and simulated poverty profiles for Ceará

	%	Observed values		Rising the mean of distribution of schooling β_{99}		$\beta_{concave}$		β_{convex}		Reducing 50% of β_{99}		$\beta_{concave}$		β_{convex}	
		P(0)	Composition	P(0)	Composition	P(0)	Composition	P(0)	Composition	P(0)	Composition	P(0)	Composition	P(0)	Composition
Ceará		45,81		34,66		32,19		36,47		42,20		37,58		45,35	
metropolitan	38,80	27,14	22,99	15,34	17,16	12,57	15,15	17,25	18,35	22,36	20,56	17,41	17,98	26,08	22,32
urban	29,00	43,23	27,37	28,21	23,60	25,17	22,68	31,27	24,86	39,16	26,91	32,13	24,79	43,44	27,78
rural	32,10	70,69	49,53	63,84	59,11	62,22	62,05	64,38	56,66	68,91	52,41	66,85	57,11	70,34	49,79
Men	48,70	47,02	49,99	35,58	49,99	33,10	50,07	37,36	49,88	43,24	49,90	38,51	49,91	46,28	49,70
Women	51,30	44,66	50,01	33,79	50,01	31,33	49,93	35,64	50,12	41,21	50,10	36,69	50,09	44,47	50,30
Years of schooling															
0	38,10	57,87	48,13	44,11	48,48	41,23	48,81	45,85	47,89	52,99	47,84	48,10	48,77	56,37	47,36
1	6,50	63,50	9,01	44,85	8,41	43,19	8,72	46,56	8,30	57,35	8,83	52,01	9,00	60,07	8,61
2	8,00	56,93	9,94	43,45	10,03	41,11	10,22	45,04	9,88	52,61	9,97	46,97	10,00	56,29	9,93
3	7,80	52,66	8,97	36,20	8,15	33,70	8,17	39,02	8,35	47,23	8,73	41,44	8,60	51,50	8,86
4	10,20	43,38	9,66	33,19	9,77	30,79	9,76	35,42	9,91	40,41	9,77	35,08	9,52	43,97	9,89
5	5,30	41,00	4,74	32,34	4,95	27,71	4,56	35,23	5,12	38,33	4,81	32,10	4,53	42,49	4,97
6	3,60	35,33	2,78	27,77	2,88	24,71	2,76	31,27	3,09	34,08	2,91	28,89	2,77	38,64	3,07
7	3,50	29,70	2,27	24,85	2,51	22,23	2,42	26,55	2,55	28,83	2,39	25,18	2,35	31,90	2,46
8	4,50	19,24	1,89	14,97	1,94	12,83	1,79	16,34	2,02	18,20	1,94	15,24	1,82	20,47	2,03
9	1,80	22,34	0,88	17,57	0,91	16,32	0,91	18,83	0,93	22,26	0,95	18,58	0,89	23,18	0,92
10	1,60	16,76	0,59	14,00	0,65	12,57	0,62	14,35	0,63	17,21	0,65	14,71	0,63	17,92	0,63
11	6,60	7,05	1,02	6,13	1,17	5,30	1,09	6,58	1,19	6,71	1,05	5,55	0,97	7,65	1,11
12	0,40	0,77	0,01	0,77	0,01	0,77	0,01	1,53	0,02	0,77	0,01	0,77	0,01	1,53	0,01
13	0,30	1,03	0,01	1,03	0,01	1,03	0,01	1,03	0,01	1,03	0,01	1,03	0,01	1,03	0,01
14	0,20	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1,44	0,01
15	1,00	0,29	0,01	0,29	0,01	0,29	0,01	0,29	0,01	0,29	0,01	0,29	0,01	0,29	0,01
16	0,50	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
17	0,10	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

Fonte: PNAD/IBGE 1999
Source: authors' calculations from the PNAD 1999 dataset.

13

One striking revelation of this profile is that the simulated educational expansion policies, although they were not designed in a spatially sensitive manner, appear to reinforce the prevalence of rural poverty. This is largely because living in rural areas is currently associated with having lower educational attainment and, as a result, the multinomial logit that assigns the distribution of extra years of schooling among individuals, allocates them more often to urban residents, ceteris paribus. Hence, whereas 50 percent of Ceará's poor today live in rural areas, between 56 percent and 62 percent (of a smaller number) might do so in ten or twenty years, if special care is not taken to encourage faster enrollment and good school supply in rural areas.

5. Conclusions

As with most uses of econometric estimates to make out-of-sample predictions, the results of our microsimulation exercise should be treated with considerable circumspection. Probably even more than most. Household data is measured with substantial error. Educational data based on years of schooling, in particular, is famously a very poor measure for quality-adjusted human capital stocks. Our models of fertility and occupational choices are acceptable only as very reduced forms. And their parameters, as indeed all others, may very well change over time.

Having said all this, the following five conclusions appear to receive some support from our analysis, and might be of some use to those concerned about the impact of educational policies on the distribution of economic welfare in the State of Ceará in the next ten to twenty years.

First, a broad-based expansion of enrollment and reduction in evasion rates, which raised average endowments of education (say, to seven years), would be very likely to make a substantial contribution to poverty reduction, of the order of a quarter to a third of the state's current incidence.[10] These policies would not, however, have the same impact on inequality. While educational expansion would be equalizing, most inequality measures would not record substantial declines as a result. The Gini coefficient for household incomes would lose at best a tenth of its value, and would still remain at a very high level, above 0.55.

Second, results depend on what happens to returns to education, which are determined by the interaction between the relative supply of and demand for different skills. In this paper, we did not model the determinants of the demand side of the labour market. While we provided estimates for three possible scenarios, effectively considering a range for the variation in returns, there is no guarantee that actual changes over the next one or two decades will stay within that range. Given that gains in labour earnings to the poor are sensitive to these changes, a stagnation of demand for unskilled labour should cause particular cause for concern.

Third, if our analysis shed any light at all on the impact of educational expansion on the distribution of income in Ceará, it was on the crucial role played by household dynamics in

[10] And closer to a third for the distribution-sensitive FGT(2).

the process. We saw that the State appears to have something of a 'reserve army', awaiting conditions to enter paid or self-employment. Like in other places where educational levels rose rapidly, this is to a large extent composed of women.[11] As they acquire education and enter the labour force, their fertility behaviour also changes, reducing the number of children in the family.

In income terms, each of these tendencies is positive for the families to which they belong. In fact, the participation and demographic changes arising from educational expansion account for around half of its overall poverty reduction impact. Figures 8, 9 and 10 illustrate the great importance of these gender-sensitive effects on the overall welfare of poor families. In the labour market, however, a large inflow of women into relatively underprivileged segments may generate downward wage pressure or enhance job competition. The extent to which Ceará will be able to capitalize on a more educated labour force depends, in large measure, on how effectively it ensures a level playing field for its women.

Fourth, the changes predicted by our model suggest that gains in average education might reinforce the rural nature of the state's poverty profile. The negative statistical association between rural residence and educational attainment, which essentially reflects past policies, leads the model to predict that gains in education would be less widespread in the countryside than in urban areas. This need not be the case, if there is a concerted public effort to make quality education more accessible in rural areas, and possibly to support programmes that make school attendance relatively less costly in terms of forgone earnings from the child's labour. But the results do suggest that, in the absence of such concerted efforts, inroads into rural deprivation may be less than expected.

It should also be noted that a number of important choices, or dimensions of household or worker behaviour, remained outside the scope of our analysis. Amongst these is the key decision to migrate. Greater endowments of education might affect the flows of migrants within the state—say, from rural areas to metropolitan Fortaleza—or outwards from the state. These decisions are likely to be determined by the relative conditions of labour demand, and thus wages, in these areas, and in other states. This falls outside the scope of this simple model, but this does not make it any less important a concern for policy-makers.

Finally, the relatively muted impact of Policy Two—the focused effort to reduce functional illiteracy by half—on the incomes of the poor suggests that such a targeted exercise should not be seen as a substitute, but rather as a complement, to a broader expansion of educational opportunities across the board.

[11] See Bourguignon et al. (2000) on the key role played by changes in female participation in the Taiwanese development process.

References

Barros, R., Henriques, R., Mendonça, R. (2000) 'Pelo Fim das Décadas Perdidas: Educação e Desenvolvimento Sustentado no Brasil', in Henriques, R. (ed.) *Desigualdade e Pobreza no Brasil*, IPEA: Rio de Janeiro.

Bourguignon, F., Ferreira, F.H.G., Lustig, N. (1998) 'The Microeconomics of Income Distribution Dynamics in East Asia and Latin America' (mimeo), World Bank, DECRA: Washington DC.

Bourguignon, F., Fournier. M., Gurgand, M. (2000) 'Distribution, Development and Education in Taiwan, 1979-1994' (mimeo), World Bank DECVP: Washington DC.

Ferreira, F.H.G., Barros, R. (1999) 'The Slippery Slope: Explaining the Increase in Extreme Poverty in Urban Brazil; 1976-96', *Revista de Econometria*, 19 (2): 211-96.

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

Langoni, C. (1973) *Distribuição de Renda e Desenvolvimento Econômico do Brasil*, Expresso e cultura: Rio de Janeiro.

Juhn, C., Murphy, K., Pierce, B. (1993) 'Wage Inequality and the Rise in Returns to Skill', *Journal of Political Economy*, 101 (3): 410-42

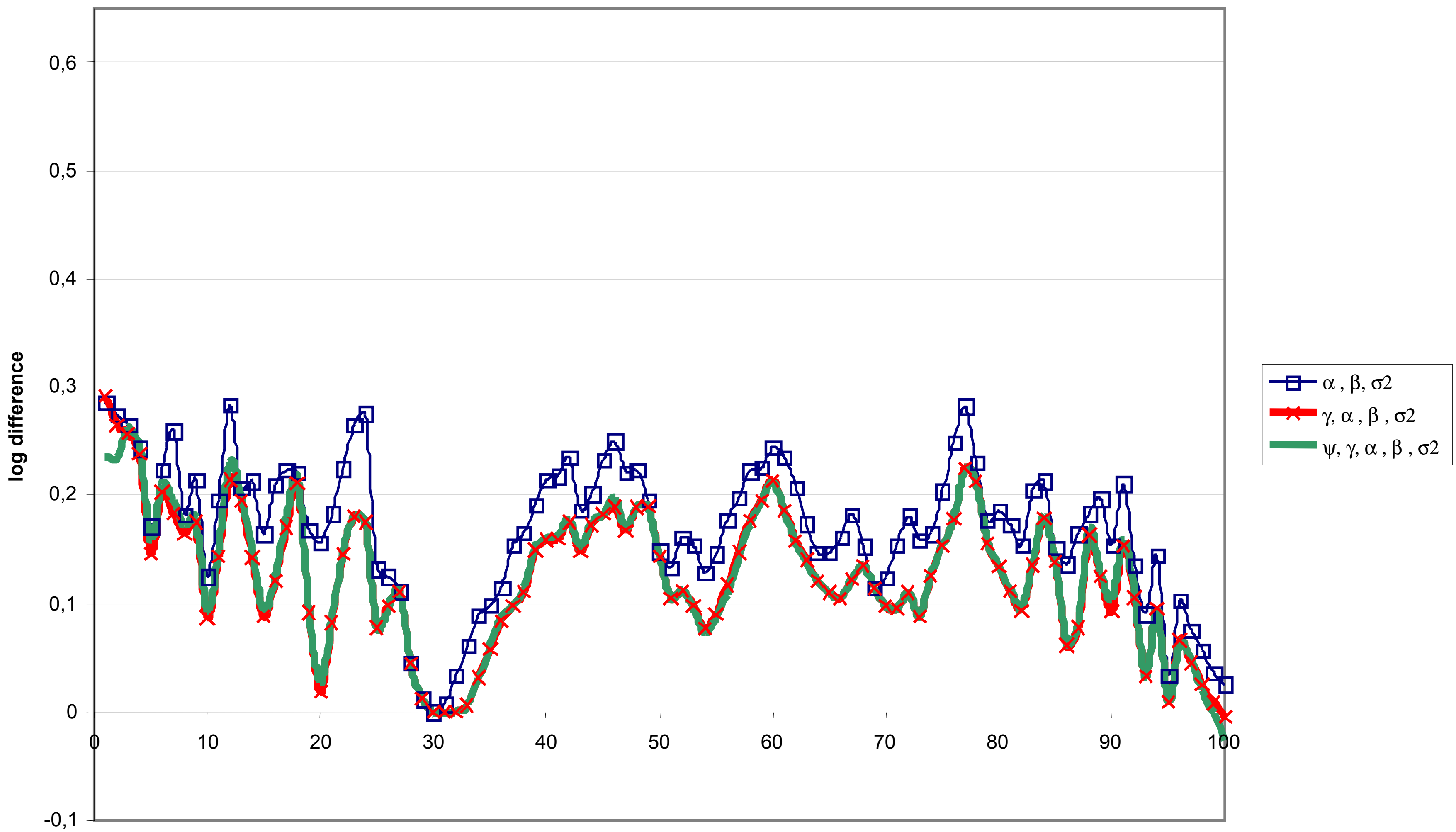
Katz, L., Murphy, K. (1992) 'Changes in the Wage Structure 1963-87: Supply and Demand Factors', *Quarterly Journal of Economics*, 107: 35-78

Reis, J. A., Barros, R. (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: 117-43

Tinbergen, J. (1975) *Income Differences: Recent Research*' North Holland Publishing: Oxford.

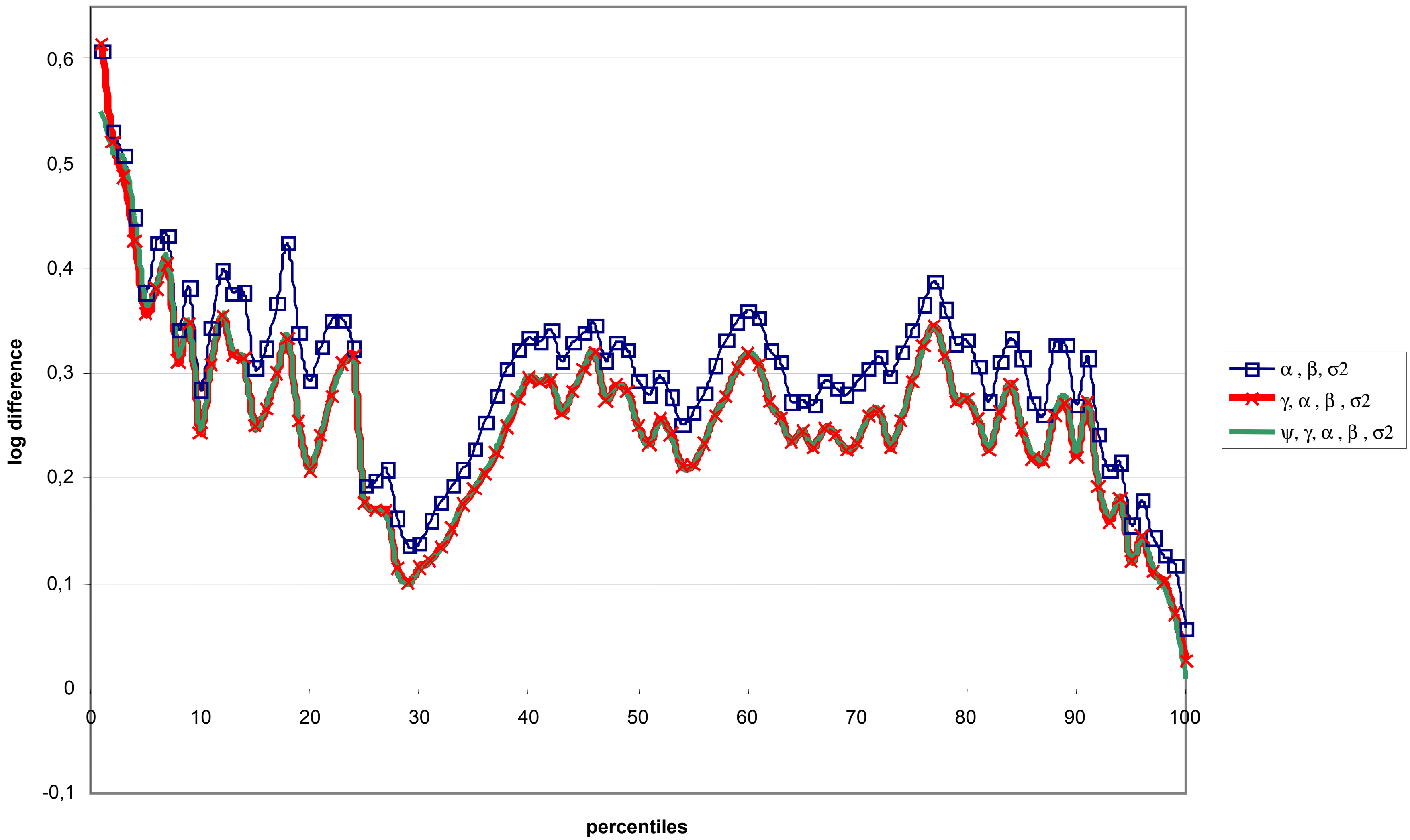
Figure 2

Earnings: mean of education equals 7, β_{99}



17

Figure 3
Earnings: mean education equal 7, β **concave**



18

Figure 4

Earnings: mean of education equal 7, β **convex**

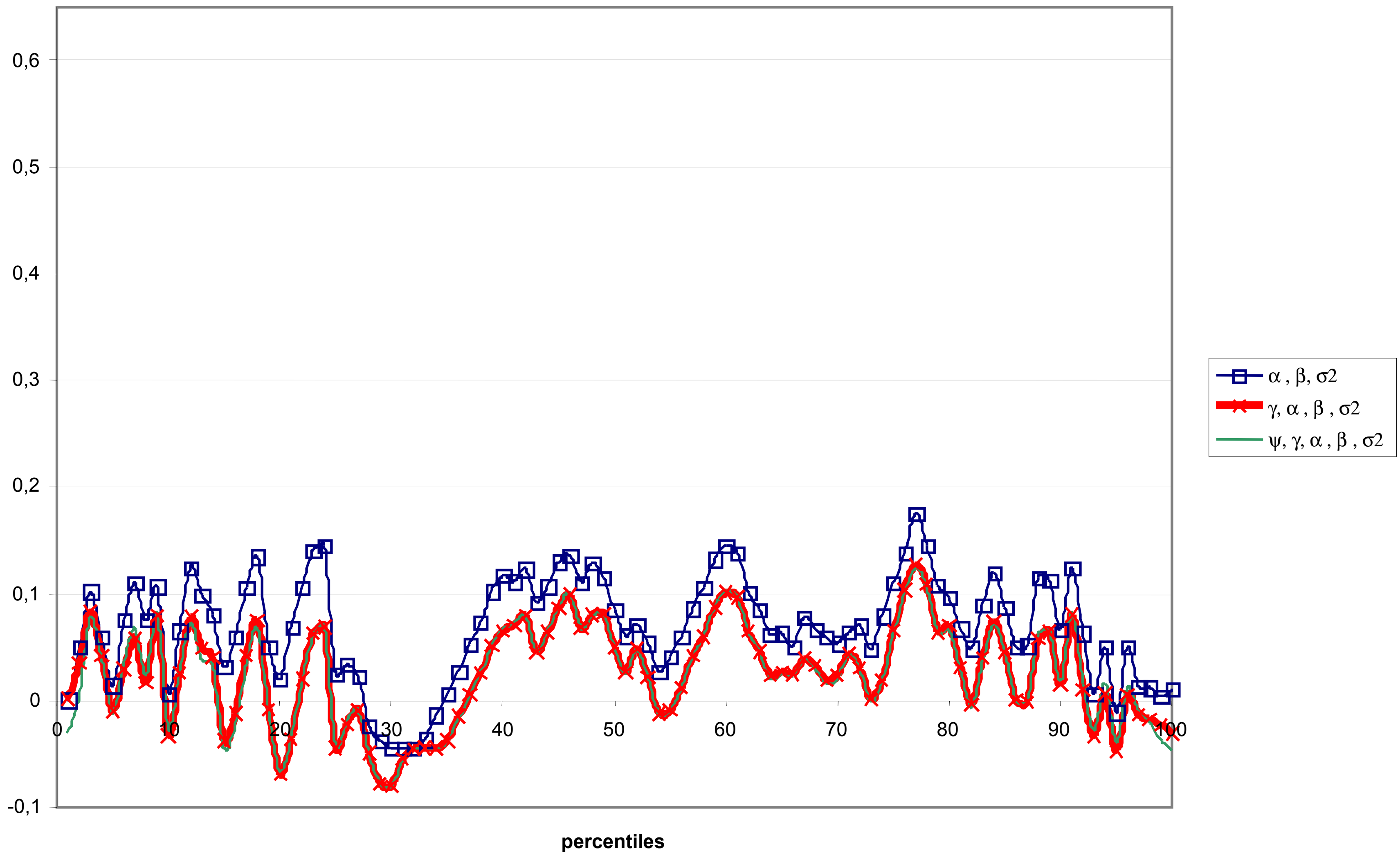
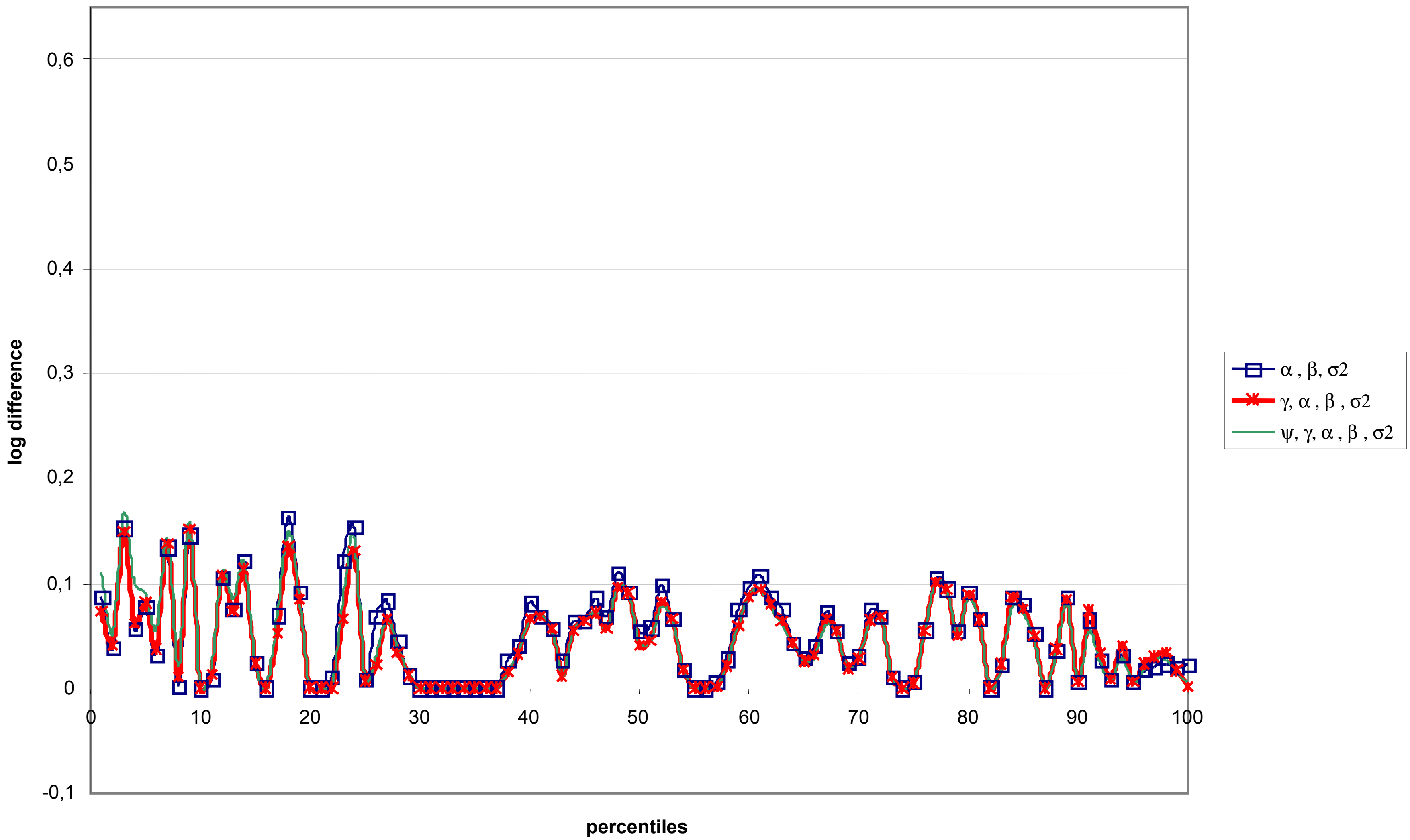


Figure 5

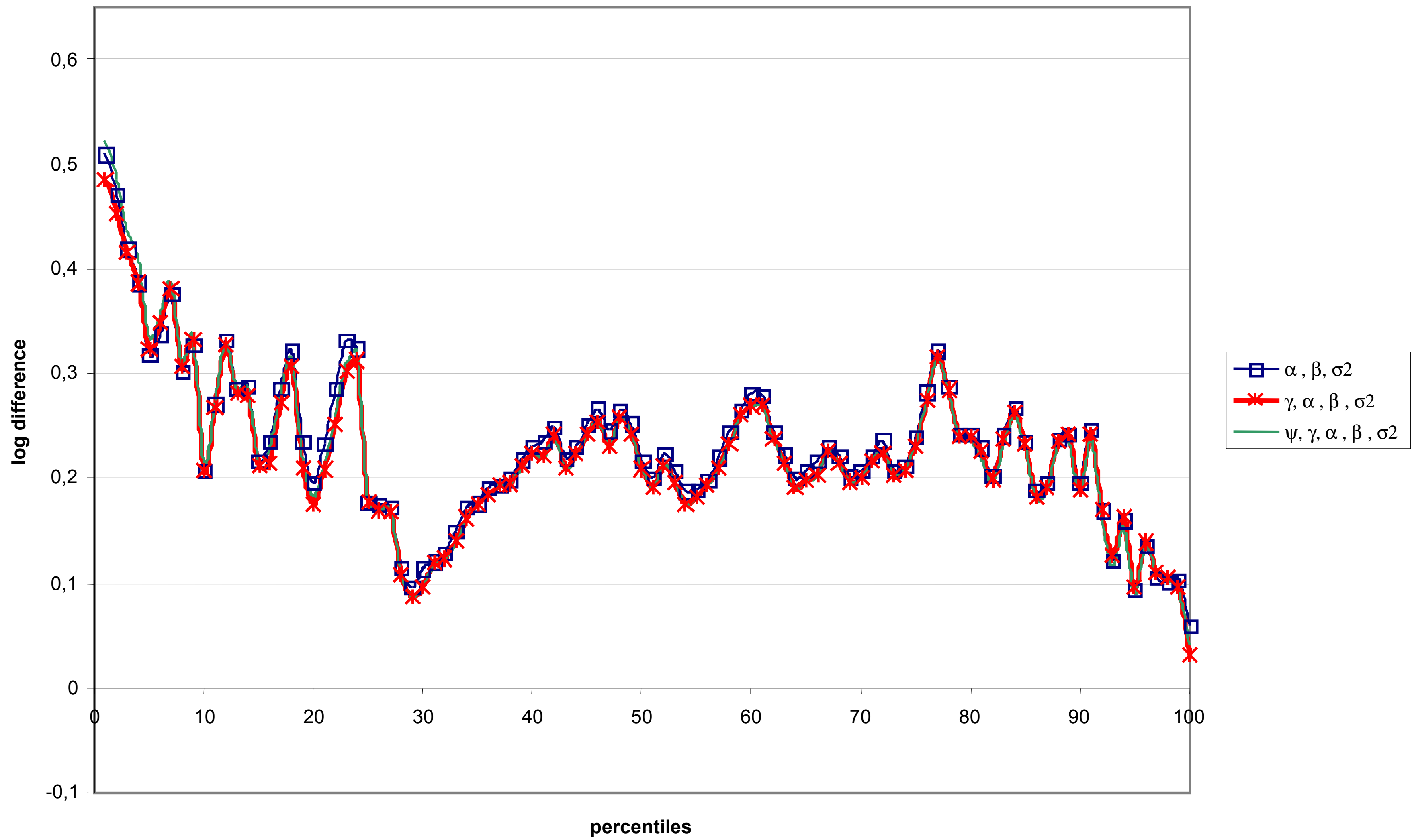
Earnings: min of education equal 4 w=0.50, β_{99}



20

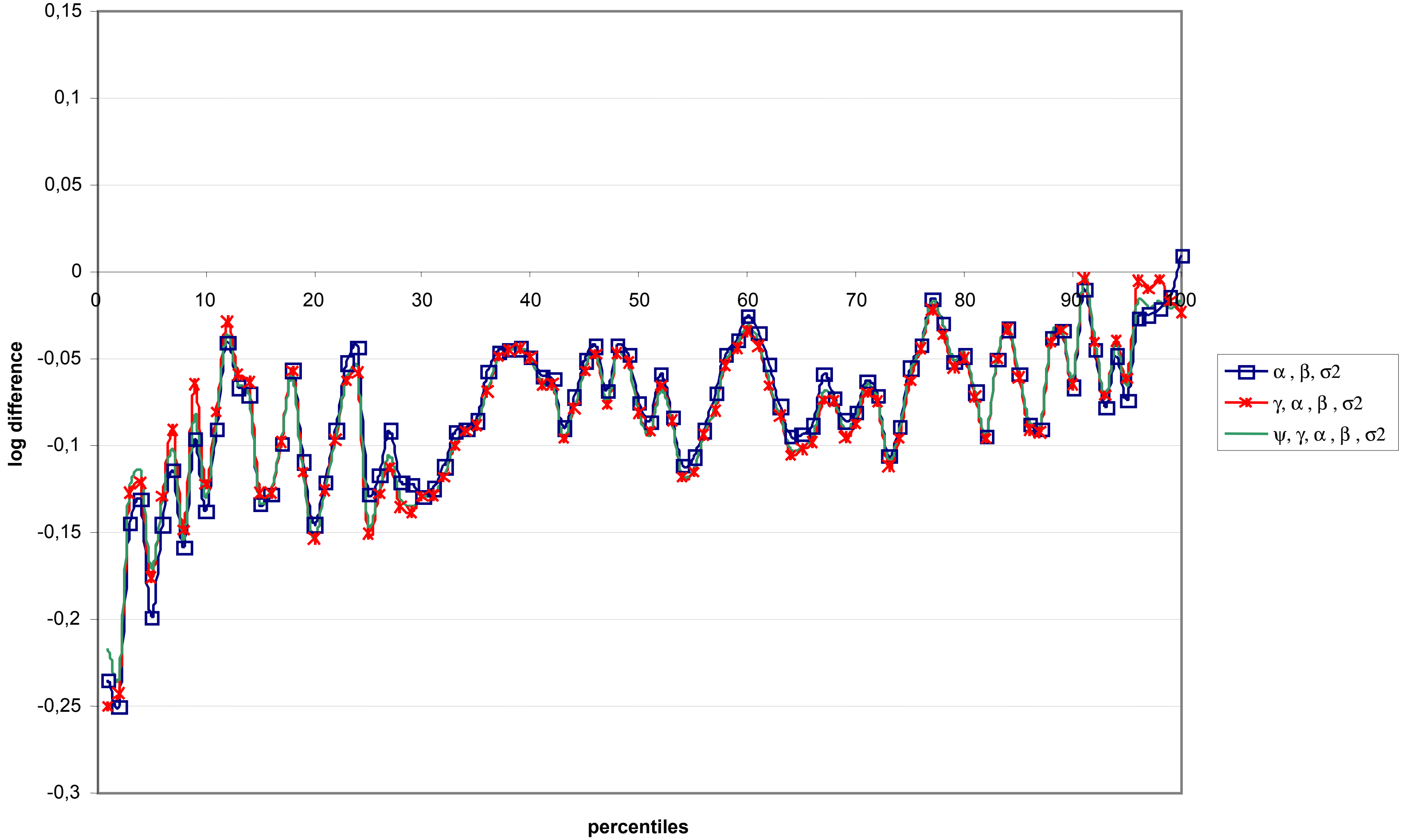
Figure 6

Earnings: min of education equal 4 w=0.50, β **concave**



21

Figure 7
Earnings: min of education equal 4 w=0.50, β **convex**



22

Figure 8
Households: mean of education equal 7, β99

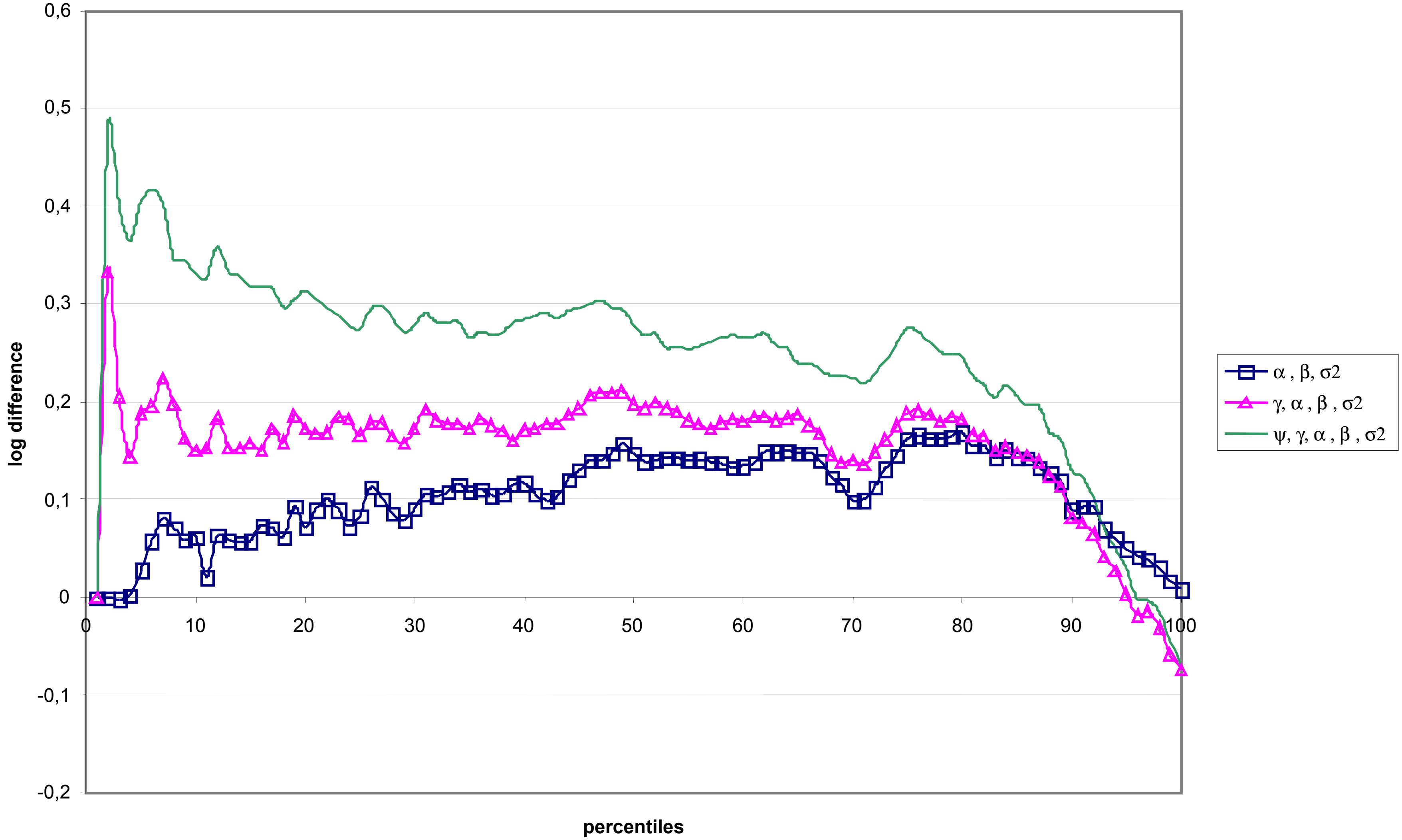


Figure 9
Households: mean of education equal 7, β concave

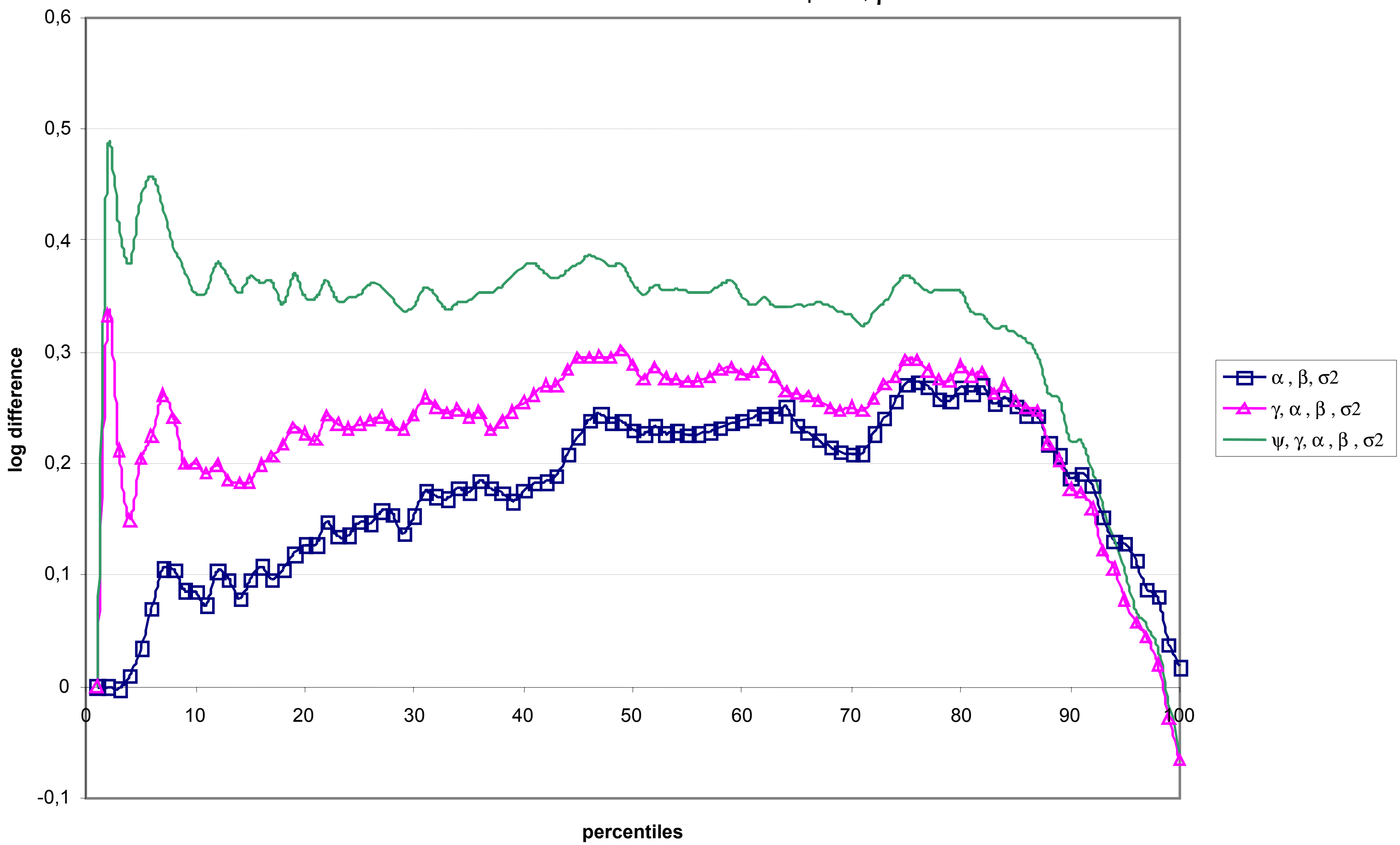


Figure 10
Households: mean of education equal 7, β convex

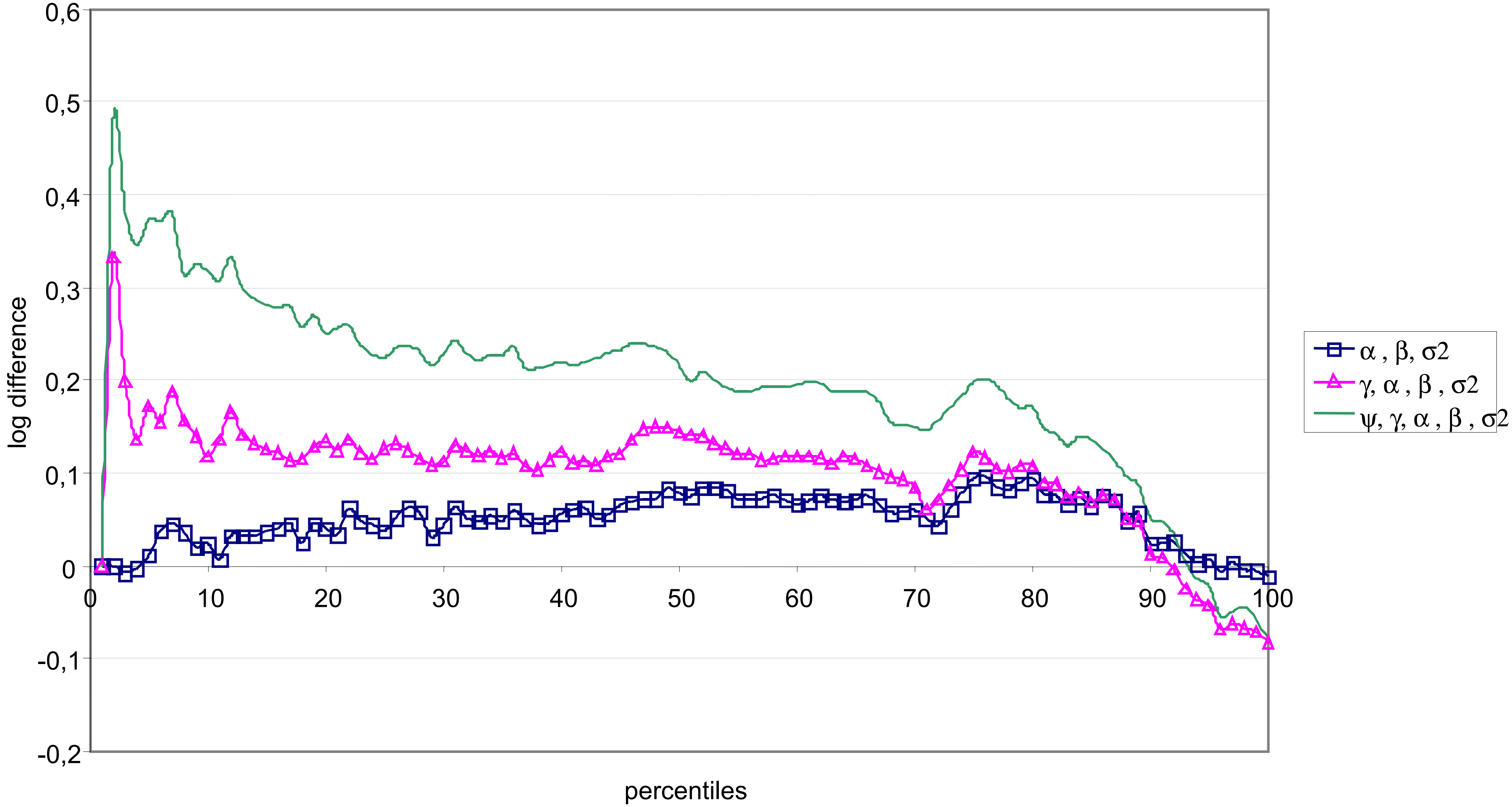
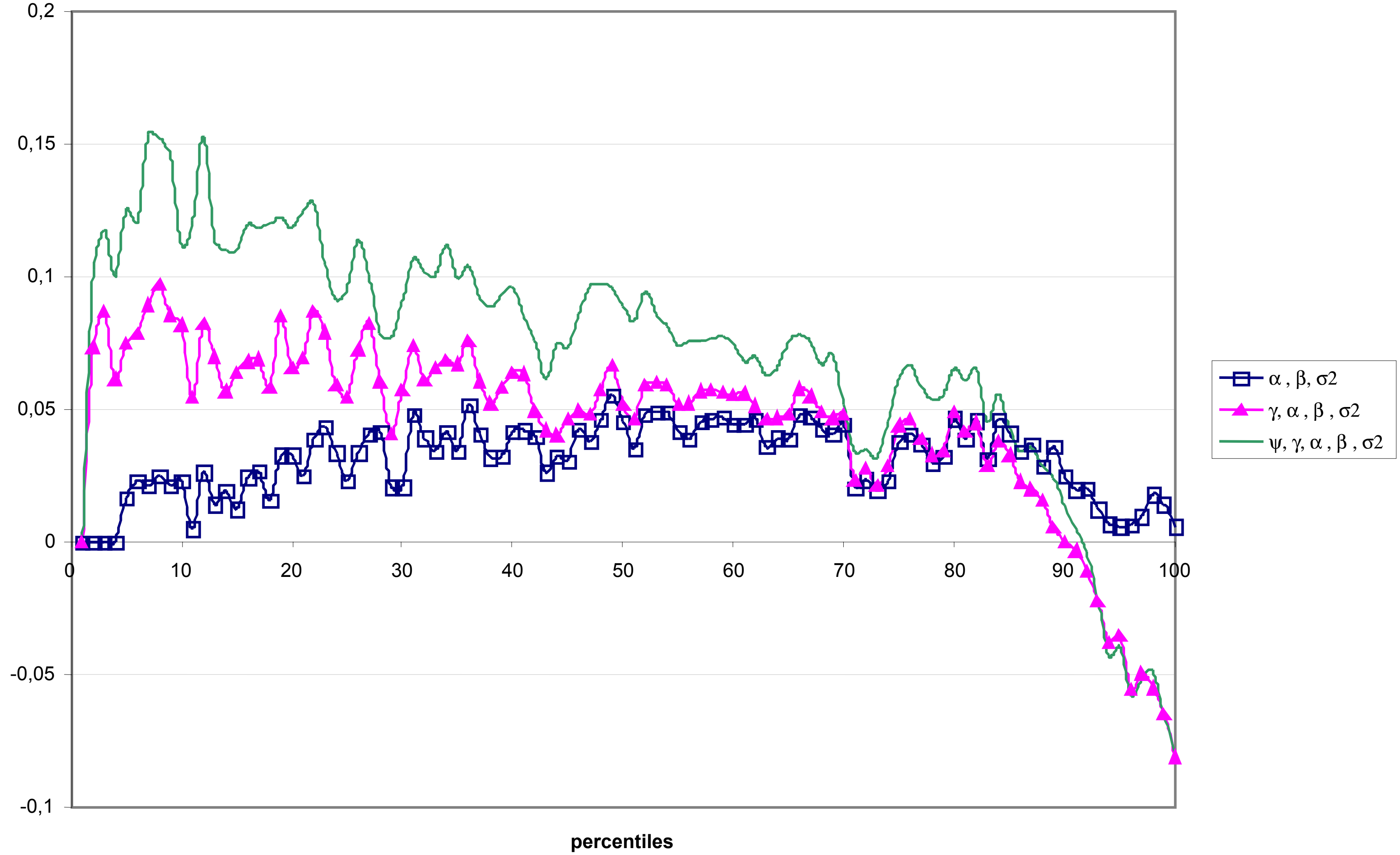


Figure 11

Households: min of education equal 4 w=0.50, β_{99}



26

Figure 12

Households: min of education equal 4 w=0.50, β **concave**

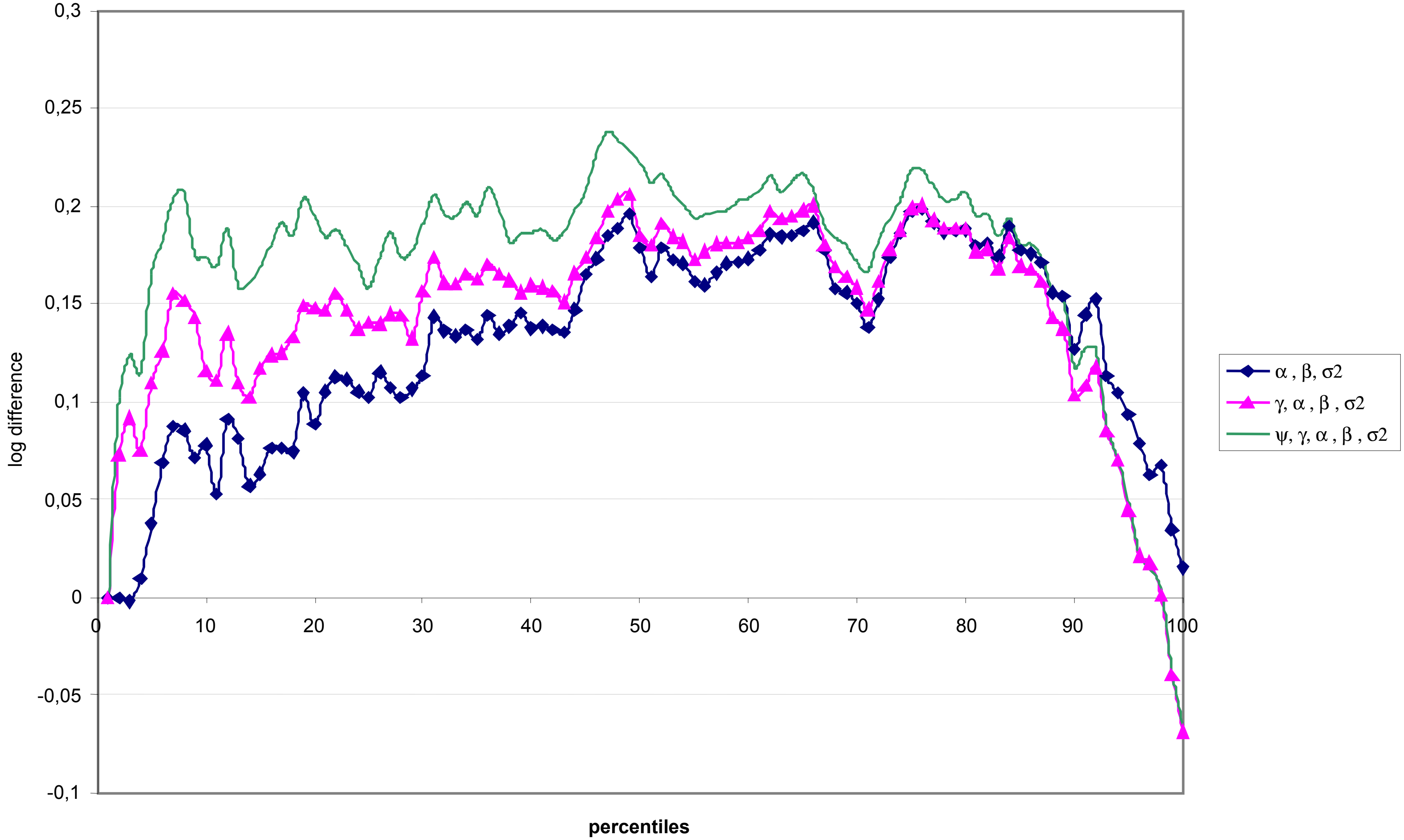
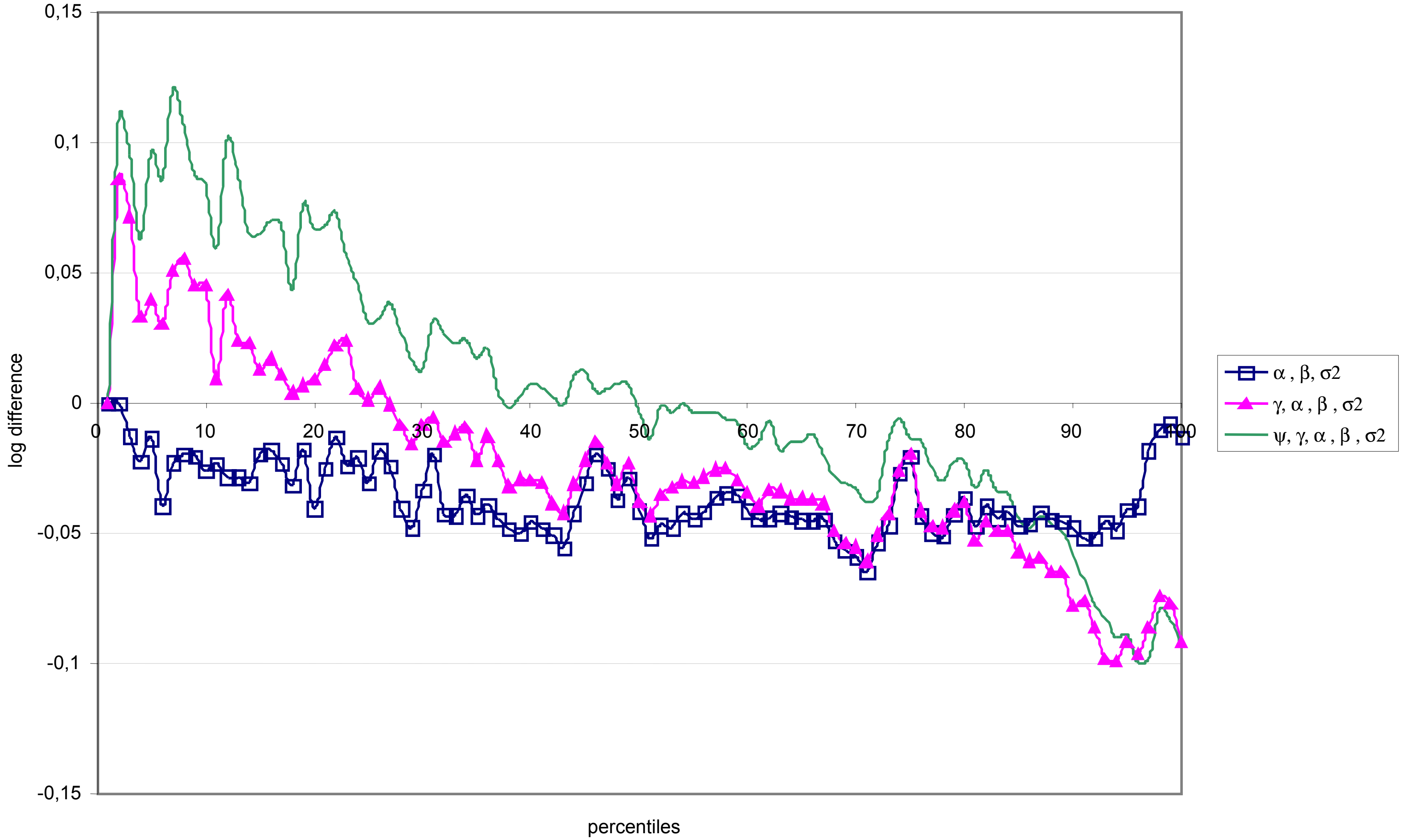


Figure 13
Households: min of education equal 4 w=0.50, β **convex**



28

Figure 14
Net entrances into the labour force (all sectors) per percentile, frequency