Influences Of Transition In Age-Education Structure And Internal Migration On The Labour Market In Brazil[1]

Influências de transição da Estrutura Idade-Educação e Migração Interna no Mercado de Trabalho no Brasil[2]

Ernesto Friedrich de Lima Amaral[3]
Eduardo Luiz Gonçalves Rios-Neto[4]
Joseph E. Potter[5]

Abstract: This study develops a methodology that incorporates internal migration dynamics into models that estimate the impact of demographic and education transitions on the age-education earnings profiles of Brazilian workers over time. Techniques to estimate the level and pattern of migration were integrated. Findings follow initial hypothesis, and indicate that the negative impact of cohort size on earnings is even more negative than estimates that did not take into account population flows. These methodological strategies can be applied to further studies when new data become available, as well as to other countries with the availability of migration data.

Key-words: Age-education transition; Internal migration; Labour market.

Resumo: Este estudo desenvolve uma metodologia que incorpora a dinâmica de migração interna em modelos que estimam o impacto da transição demográfica e educação sobre os perfis de idade-educação ganhos dos trabalhadores brasileiros ao longo do tempo. Técnicas para estimar o nível eo padrão de migração foram integrados. Conclusões seguem hipótese inicial, e indicam que o impacto negativo do tamanho da coorte sobre os ganhos é ainda mais negativo do que as estimativas que não levam em conta os fluxos populacionais. Estas estratégias metodológicas podem ser aplicadas a outros estudos quando novos dados se tornam disponíveis, assim como a outros países com a disponibilidade de dados de migração.

Palavras-chave: Transição na idade-educação; Migração interna; Mercado de Trabalho.

Introduction

Studies have found that the rapid decline in fertility rates in Eastern and Southeastern Asia have generated a proportional increase of workers (15-64 years of age) in relation to the proportion of children (0-14) and elderly (65+) (Bloom and Freeman 1986; Bloom et al. 2003; Williamson 2003; Mason 2005; Mason and Feng 2005). The decrease in the dependency ratio has a positive and significant influence on the economic development of the countries in this region. This indicates the need to evaluate changes in the age structure and its influence on economic development over time. The same demographic changes that occurred in Asia are now occurring in Latin America. However, Latin American countries are different than Asian countries

---

[1] Artigo recebido em novembro de 2011 e aprovado em fevereiro de 2012.
Artigo apresentado no VII Encontro Nacional Sobre Migrações de Tema Central: Migrações, Políticas Públicas e Desigualdades Regionais, realização de 10 a 12 de Outubro de 2011, Curitiba/PR.
[2] This research received support from the Brazilian Association of Population Studies ("Associação Brasileira de Estudos Populacionais" – ABEP) and the Brazilian Institute of Applied Economic Research ("Instituto de Pesquisa Econômica Aplicada" – IPEA), through a research scholarship on Demographic Dynamics and Brazilian Development in the Long Term ("Dinâmica Demográfica e o Desenvolvimento Brasileiro de Longo Prazo", ABEP/IPEA, 001/2009).
[3] Universidade Federal de Minas Gerais (UFMG), Brazil. Email: amaral@fafich.ufmg.br
[4] Universidade Federal de Minas Gerais (UFMG), Brazil. Email: eduardo@cedeplar.ufmg.br
[5] University of Texas at Austin. Email: joe@prc.utexas.edu

because they have lower levels of educational attainment as well as greater levels of socio-economic inequality.

The significance of fertility swings and a shifting age distribution on economic development was analysed in studies of the influence of the "baby boom" on labour market outcomes in the United States (Easterlin 1978; Freeman 1979; Welch 1979; Berger 1985; Triest et al. 2006). Cohorts born during the "baby boom" entered the American labour market between the end of the 1960s and the middle of the 1970s. The new labour force entrants had more schooling than earlier cohorts: (1) the number of persons with 5–8 years of schooling and with 1–3 years of high school fell considerably; (2) the number of high school graduates and those with at least some college education increased significantly.

In Brazil, the decline in fertility rates varied in time, pace and scale among the different states and municipalities (Potter et al. 2002; Potter et al. 2010). Amaral et al. (2007) investigated the influence of demographic and educational transitions on local labour markets in Brazil. That study took into account regional variation in age-education structure, and the results suggest that compositional changes of the labour force have a significant impact on earnings.

The objective of our analysis is to investigate other specificities of demographic and educational changes and economic development in Brazil over time, as a continuation of the previous study undergone by Amaral et al. (2007). In order to conduct a regional study, it is necessary to incorporate the influence of internal migration on workers' earnings. The intention is to include internal migration information in the estimation of models that measure the impact of age-education changes in the Brazilian labour market. The introduction of population movements in this kind of analysis has not been done in previous studies. Therefore, a new methodological procedure is developed, integrating techniques that estimate the level (Stillwell 2005) and pattern of migration (Rogers and Castro 1981).

2. Background

A series of studies have addressed the "demographic dividend", whereby the changing age structure in developing countries resulting from sustained and rapid fertility declines presents a temporary "window of opportunity", during which a reduced dependency ratio yields high rates of growth in income per capita. This positive impact on growth derives from the mechanical link between the size of the working age and total population, increases in labour supply due to higher proportions of women becoming employed, higher savings rates, higher rates of human capital formation, and, possibly, from the impact of an aging population on capital accumulation via capital deepening. It has been demonstrated often that the decline in the dependency ratio, caused by a rapid fertility decline, has substantially influenced economic development in East and Southeast Asia (Bloom and Freeman 1986; Bloom et al. 2003; Williamson 2003; Mason 2005; Mason and Feng 2005). These authors stress the transitory nature of the decrease in the dependency ratio and the conditional nature of the dividend: the drop in the dependency ratio will only result in economic growth in the right policy environment.

While the demographic dividend literature focuses on the ratio of the working-age population to the rest of the population, the analysis employed for this study focuses on the changes in age structure within the working-age population that necessarily accompany the demographic transition, as well as the concurrent change in levels of education that normally accompanies, and may even drive, the demographic transition. Whereas in the dividend literature, the focus is mainly on aggregated outcomes, the concern here is with the distribution of economic outcomes within the labour force and how these may be affected by its changing composition. Discovering how to take advantage of these changes led us to examine studies that had been previously conducted on another major demographic shock: the "baby boom" in the United States (Easterlin 1978; Freeman 1979; Welch 1979; Berger 1985; Triest et al. 2006). Because of the "baby boom" that followed World War II, which peaked from 1955 to 1960, there was an especially significant change in the age structure of the U.S. workforce in the late 1960s and early 1970s. This timeframe was a period when the number of young persons increased very rapidly. The key finding was that the age-earnings profile of male workers appears to be significantly influenced by the age composition of the workforce, i.e., large cohorts do depress earnings. While these studies all refer to the U.S. case, they illustrate the power of the supply-demand framework and the richness of combining age and schooling as basic labour inputs, thus driving wage variations.

As in other developing countries, the age-education transition presented an expressive geographical and temporal heterogeneity across different locations in Brazil. Amaral et al. (2007) used the 1970-2000 Brazilian Censuses to estimate models of the impact of age-education structure on workers' earnings in Brazilian micro-regions over time. Information on age was categorised into four groups: youth (15-24 years); young adults (25-34 years); adults (35-49 years); and older adults (50-64 years). The level of education was classified into three groups using information on completed years of schooling: no more than the first phase of elementary school (0-4 years of schooling); the second phase of elementary school (5-8 years of schooling); and at least some secondary school (9 years of schooling or more). A summary of the results of this study is presented below. Table 1 shows the distribution of the male population by year and age-education group in Brazil. The proportion of males with 0-4 years of schooling decreased between 1960 and 2000. For instance, the proportion of males between 15 and 24 years of age and with 0-4 years of schooling reduced from 30.84 per cent in 1960 to 9.04 per cent in 2000. Moreover, the proportions of those with 5-8 years of schooling and those with at least 9 years of schooling increased in the same period. The greatest increases in groups with at least 9 years of schooling were for 15-24-year-olds (from 1.08 per cent in 1960 to 10.24 per cent in 2000), and for 35-49–year-olds (from 0.91 per cent in 1960 to 8.46 per cent in 2000).

Table 1 - Distribution of the male working-age population by age-education group, Brazil, 1970–2000.

Age-Education Group	1960	1970	1980	1991	2000
15-24 years 0-4 years of schooling	30.84	28.19	20.59	14.61	9.04
15-24 years 5-8 years of schooling	2.63	5.38	10.53	12.09	12.46
15-24 years 9+ years of schooling	1.08	2.74	5.87	5.97	10.24
25-34 years 0-4 years of schooling	22.66	19.71	16.39	12.41	8.82
25-34 years 5-8 years of schooling	1.18	1.98	3.90	6.82	7.63
25-34 years 9+ years of schooling	1.19	2.00	4.77	7.40	8.12
35-49 years 0-4 years of schooling	24.47	22.66	19.02	17.11	13.32
35-49 years 5-8 years of schooling	0.98	1.62	2.39	3.67	6.73
35-49 years 9+ years of schooling	0.91	1.59	2.84	5.54	8.46
50-64 years 0-4 years of schooling	13.21	12.84	11.72	11.49	10.36
50-64 years 5-8 years of schooling	0.43	0.65	0.94	1.16	1.99
50-64 years 9+ years of schooling	0.40	0.62	1.05	1.72	2.84
Total	4,039,104[1]	25,760,594	32,613,947	43,434,534	53,177,964

Source: 1960–2000 Brazilian Censuses.

[1] The 25-percent-sample microdata of the 1960 Census do not have sample weight variables to estimate population size.

In Brazil, regional differences in timing and pace of fertility and education transitions generated substantial disparities in age-education structure across regions, states and municipalities at different points in time (Amaral et al. 2007). On the one hand, there was an increase over the years in the proportion of young adults (25-34) with at least nine years of schooling. Greater proportions of this age-education group are observed in Southeast, South and Central-West regions compared to areas in the North and Northeast. On the other hand, the percentage of male adults (35-49) at the lowest level of education (0-4 years of schooling) has been decreasing in all micro-regions over the last decades. Areas in the southeast and south of Brazil have a greater decrease in the proportion of men in the lowest education group compared to the North and Northeast of the country.

Differences across regions suggest the need to use models that take into account the peculiarities of these micro-regions and to develop the analysis about changes over time. To generate models to verify the impact of the age-education transition on the labour market, Amaral et al. (2007) used the dependent variable as the logarithm of the mean real income by year, micro-region[6], and age-education group. In this study, only the male population was analysed. Such as in Table 1, the population was divided into twelve age-education groups.

[6] These 502 micro-regions differ from those defined by the Brazilian Institute of Geography and Statistics (IBGE) and available in the Census microdata but closely approximate those defined for the 1991 Census (Potter et al. 2002).

Let $log(Y_{git})$ be the logarithm of wages. The twelve indicators of age-education groups (G) interacting with time (θ) are included in the model, taking the first age-education group interacting with time (four variables) as the reference category. Thus, G is a set of age-education-group indicators (dichotomous variables), specifically, G_{11} (15–24 years; 0–4 education), G_{12} (15–24; 5–8), G_{13} (15–24; 9+), G_{21} (25–34; 0–4), G_{22} (25–34; 5–8), G_{23} (25–34; 9+), G_{31} (35–49; 0–4), G_{32} (35–49; 5–8), G_{33} (35–49; 9+), G_{41} (50–64; 0–4), G_{42} (50–64; 5–8), and G_{43} (50–64; 9+). Furthermore, the first model has 2,008 (502*4) area-time-fixed effects (α). Equation (0) replicates the traditional Mincerian labour market equation (Mincer 1958, 1974) for 502 micro-regions (i), four (1970, 1980, 1991 and 2000) Censuses (t), and twelve age-education groups (g):

$$log\left(Y_{git}\right) = \beta_0 + (\beta_1 G_{12} + \cdots + \beta_{11} G_{43}) * \theta_t + \alpha_{it} + \varepsilon_{git}. \qquad (0)$$

The authors include the distribution of the male population in twelve age-education groups (X) interacting with time (θ), as indicated by Equation (1'):

$$log\left(Y_{git}\right) = \beta_0 + (\beta_1 G_{12} + \cdots + \beta_{11} G_{43}) * \theta_t + (\gamma_1 X_{11} + \cdots + \gamma_{12} X_{43}) * \theta_t + \alpha_{it} + \varepsilon_{git} \qquad (1')$$

Because Brazil was divided into 502 micro-regions, twelve age-education groups and four censuses, the maximum possible number of observations in the regressions is 24,096. However, only cells with at least 25 observations are included in the estimations, reducing the maximum number of observations to 19,727 for models that use 1970-2000 Censuses data and 10,782 for models with only 1991 and 2000 data.

The traditional Mincerian labour market model (Mincer 1958, 1974) considers only the direct impact of experience (age) and education on earnings, through the use of the age-education-group-dummy variables (Equation (0)). In this case, fixed effects of area and time were also introduced in the model. Table 2 illustrates this model and indicates that within the same age group, earnings increase with years of schooling. In the same way, within the same education group, earnings increase with age. These results follow the general patterns observed for economic returns in the labour market in relation to experience (age) and education of workers.

Table 2 - Fixed-effects estimates of age-education indicators on the logarithm of monthly earnings[1] (dependent variable), Brazil, 1970–2000.[2]

Variables	Coefficients and standard errors			
Constant	5.33*** (0.004)			
		Interactions with year		
Age-education indicators		**1980**	**1991**	**2000**
15–24 ; 0–4 (reference group)	––	––	––	––
15–24 ; 5–8	0.52*** (0.013)	−0.25*** (0.018)	−0.21*** (0.017)	−0.35*** (0.017)
15–24 ; 9+	1.02*** (0.015)	−0.24*** (0.019)	−0.24*** (0.019)	−0.49*** (0.019)
25–34 ; 0–4	0.35*** (0.012)	0.10*** (0.016)	0.03 (0.016)	−0.01 (0.016)
25–34 ; 5–8	1.26*** (0.013)	−0.19*** (0.018)	−0.39*** (0.017)	−0.48*** (0.017)
25–34 ; 9+	1.93*** (0.014)	−0.22*** (0.018)	−0.40*** (0.018)	−0.57*** (0.018)
35–49 ; 0–4	0.53*** (0.012)	0.13*** (0.016)	0.12*** (0.016)	0.06*** (0.016)
35–49 ; 5–8	1.64*** (0.014)	−0.09*** (0.019)	−0.29*** (0.018)	−0.49*** (0.018)
35–49 ; 9+	2.30*** (0.015)	−0.15*** (0.020)	−0.21*** (0.019)	−0.36*** (0.019)
50–64 ; 0–4	0.54*** (0.012)	0.12*** (0.016)	0.11*** (0.016)	0.14*** (0.016)
50–64 ; 5–8	1.79*** (0.016)	−0.09*** (0.022)	−0.26*** (0.022)	−0.40*** (0.021)
50–64 ; 9+	2.35*** (0.018)	−0.08** (0.024)	−0.07** (0.024)	−0.11*** (0.023)

Number of observations	19,727
Number of groups	2,008
Fraction of variance due to the υ_i	0.81
F (44; 17,675): All coefficients=0	5,538***
F (2,007; 17,675): Area*Time fixed effects=0	25.46***

Source: 1970–2000 Brazilian Censuses.

* Significant at p<0.05; ** Significant at p<0.01; *** Significant at p<0.001.

[1] Nominal income was converted to base 1 in January 2002, taking into account changes in currency and inflation.

[2] This is the traditional labour market model (Mincerian), including all age-education groups for males.

Table 3 shows results for the regression model that allows for impacts of own-effects from cohort size (distribution of male population in age-education groups), as illustrated by Equation (1'). As in the Mincerian model, the age-education-group indicators demonstrate earnings increases with age and years of schooling. The coefficients of distribution of the male population in age-education groups indicate that greater negative impacts occur for the higher educated groups. Interactions of these proportions with time show that the negative impact is decreasing over time. This pattern is mainly observed in 1991 and 2000, in which the positive coefficients counterbalance the negative impacts of 1970 (reference category). The proportion coefficients have the greater impacts for the oldest age group (50-64 years) with higher education (5-8 and 9+ years of schooling) in 1991 and 2000. Other groups also present significant positive coefficients, mainly in the last two years.

Table 3 - Fixed-effects estimates of age-education indicators, and proportion of population in age-education groups on the logarithm of monthly earnings[1] (dependent variable), Brazil, 1970–2000.[2]

Variables	**Coefficients and standard errors**			
Constant	5.23*** (0.017)			
		Interactions with year		
Age-education indicators		**1980**	**1991**	**2000**
15–24 ; 0–4 (reference group)	––	––	––	––
15–24 ; 5–8	1.09*** (0.066)	–0.39*** (0.082)	–0.42*** (0.079)	–0.68*** (0.089)
15–24 ; 9+	1.45*** (0.065)	–0.32*** (0.078)	–0.39*** (0.074)	–0.79*** (0.072)
25–34 ; 0–4	0.45*** (0.090)	0.07 (0.112)	0.07 (0.099)	–0.04 (0.095)
25–34 ; 5–8	1.70*** (0.064)	–0.33*** (0.079)	–0.54*** (0.074)	–0.71*** (0.074)
25–34 ; 9+	2.34*** (0.064)	–0.37*** (0.077)	–0.63*** (0.072)	–0.95*** (0.070)
35–49 ; 0–4	0.63*** (0.137)	0.33* (0.161)	0.28 (0.154)	0.16 (0.146)
35–49 ; 5–8	2.01*** (0.064)	–0.19* (0.078)	–0.39*** (0.073)	–0.68*** (0.071)
35–49 ; 9+	2.63*** (0.063)	–0.31*** (0.076)	–0.38*** (0.071)	–0.67*** (0.069)
50–64 ; 0–4	0.96*** (0.083)	0.11 (0.103)	–0.19 (0.101)	–0.16 (0.095)
50–64 ; 5–8	2.12*** (0.066)	–0.23** (0.080)	–0.39*** (0.077)	–0.68*** (0.074)
50–64 ; 9+	2.61*** (0.066)	–0.24** (0.080)	–0.23** (0.075)	–0.42*** (0.072)
		Interactions with year		
Proportions in age-education groups		**1980**	**1991**	**2000**
15–24; 0–4	0.75*** (0.181)	–0.28 (0.226)	–0.28 (0.213)	–0.84*** (0.214)
15–24; 5–8	–7.23*** (0.470)	3.58*** (0.543)	4.66*** (0.548)	5.20*** (0.652)
15–24; 9+	–8.49*** (0.805)	2.76** (0.899)	4.13*** (0.934)	6.66*** (0.849)
25–34; 0–4	0.70 (0.367)	–0.43 (0.511)	–1.03* (0.456)	–1.33** (0.432)
25–34; 5–8	–14.40*** (1.220)	7.79*** (1.401)	10.89*** (1.279)	11.04*** (1.291)
25–34; 9+	–12.36*** (1.134)	7.68*** (1.248)	10.76*** (1.193)	11.78*** (1.187)
35–49; 0–4	0.63 (0.537)	–1.38* (0.664)	–1.43* (0.642)	–1.86** (0.608)
35–49; 5–8	–13.22*** (1.529)	2.46 (1.923)	6.30*** (1.696)	9.56*** (1.577)
35–49; 9+	–10.58*** (1.493)	6.86*** (1.727)	8.24*** (1.563)	9.89*** (1.523)
50–64; 0–4	–1.31** (0.434)	–0.83 (0.594)	1.05 (0.592)	0.27 (0.558)
50–64; 5–8	–23.51*** (4.258)	9.60 (5.325)	10.25* (5.130)	19.88*** (4.526)
50–64; 9+	–14.32** (4.592)	11.18* (5.369)	11.91* (4.890)	15.45** (4.681)

Number of observations	19,727	F (92; 17,627): All coefficients=0	2,902***
Number of groups	2,008	F (2,007; 17,627): Area*Time fixed effects=0	18.80***

Source: 1970–2000 Brazilian Censuses.

* Significant at p<0.05; ** Significant at p<0.01; *** Significant at p<0.001.

[1] Nominal income was converted to base 1 in January 2002, taking into account changes in currency and inflation.

[2] This is the model based on Equation (1'), including all age-education groups for males.

These results indicate that cohort size is an important factor in the determination of earnings. Coefficients of proportions of people in age-education groups have negative impacts on earnings, with greater intensity in groups with higher education. These findings follow the theory that age-education groups are not perfect substitutes, generating negative impacts of cohort size on earnings. Even with changes in the Brazilian labour market demand in the last decades, this variation was not sufficient to compensate for the supply variation.

Effects and magnitudes of technological and institutional changes are suggested by the positive interactions of the proportion with the year in Table 3. Even with positive interaction terms, only some of them were strong enough to compensate for the negative impacts shown in the reference period (1970). Results also indicate that institutional and demand changes were insufficient to compensate for the negative pressures that supply changes (distribution of male population in age-education groups) have been creating on groups with medium (5-8) and higher (9+) education. The only exception occurs in the older group (50-64) with the highest education (9+), wherein there is a positive impact in 2000 that exceeds the estimation in the reference year (1970).

3. Data

Because there was great geographical heterogeneity in the declining fertility rate and educational improvements in Brazil, the estimated labour-market models introduced area- and time-fixed effects (Amaral et al. 2007). In a sub-national study, it is necessary to introduce migration flows among the analysed areas. Previous studies about the influence of changing age distribution on economic development did not incorporate the migration variable because the analysis was done at the national level. Thus, aside from the information used by Amaral et al. (2007), migration variables were also collected from the 1970, 1980, 1991 and 2000 Brazilian Censuses microdata for the present study. Specifically, information on the state of birth and the number of years that the respondent lived in the municipality was obtained from the 1970–2000 Censuses. Furthermore, the 1991 and 2000 Censuses were used to provide information on which municipality and state the person lived exactly five years before the Census. The following sections detail how the data were used to estimate the impact of migration flows on workers' earnings.

4. Models

Because the models were estimated at the local level, it would have been important to account for internal migration in the equations. Internal migration in Brazil is an important demographic component because significant population streams from rural to urban areas occurred in previous decades. This migration is characterised by streams from areas of higher fertility rates to those of lower fertility rates. In other words, internal migration might reduce the differential in birth rates between rural and urban areas. However, this process might also increase the difference in dependency ratios because migrants are concentrated within working ages. These specificities indicate that models would have to take into account the migration variable to surmise the influence of age structure on economic development.

Because internal migration in Brazil is influenced by the availability of jobs and level of earnings (Oliveira and Jannuzzi 2005), people move to those areas with better income opportunities due to relative declines in the size of the labour force in a particular age-education group. If there were no migration flows, the sending areas (which already have lower relative earnings) would have even lower earnings in specific age-education groups, and the receiving areas (which already have higher relative earnings) would experience increases in earnings. In such a scenario, migration biases the estimated negative effects on wages toward zero. By not controlling for migration in the models, the results of Table 3 underestimate the negative effect of group size on earnings. The hypothesis is that by controlling for migration flows, the negative impacts of age-education-group proportions will be even more negative than those presented by Equation (1').

Borjas (2003) suggests that studies about the impact of immigration on the labour market have mainly been based on the comparison of employment opportunities between immigrants and natives across regions. The main conclusion is usually that immigrants do not lower native wages. Using a new approach based on the assumption that similarly educated workers with different levels of experience are not perfect substitutes, Borjas suggests that immigration reduces the wage and labour supply of competing native workers. However, because internal population flows in Brazil are influenced by the availability of jobs and levels of income in sending and receiving areas, migration is an endogenous variable that cannot be simply introduced as an exogenous variable in the estimation of labour outcomes.

However, migration flows cannot just be introduced as independent (exogenous) variables in the models. The introduction of migration models to estimate earnings would generate endogeneity problems because the level of earnings in one area also explains immigration flows. As a strategy to correct for endogeneity problems, a methodology was developed by congregating the estimation of the migration level proposed by Stillwell (2005) and the correction of migration schedules with mathematical models (Rogers and Castro 1981). In the following sub-sections, there is an explanation of this methodological strategy.

4.1. Estimation of migration level

Gravity models, taking into account distances among areas, are used to control for migration flows. More than only distance, these gravity models take into account the population in the micro-region of origin (at the beginning of the period), the population in the micro-region of destination (at the end of the period), as well as the proportion of migrants already living in a specific area at the time of a given Census. Distance is constant over time, but the micro-regions' populations at the beginning and end of the period change their out- and in-migration trends over time. This study used a matrix of kilometre distances between all Brazilian micro-regional centroids[7]. The combination of this matrix with information on population size of micro-regions as well as age and education of migrants was used to generate attraction and repulsion measures of population flows among micro-regions.

[7] This matrix of kilometer distances between Brazilian micro-regions is available on (http://schmert.net/KnoxCox) by the project "Knox meets Cox: Adapting Space-Time Epidemiological Statistics to Demographic Studies", developed by Carl P. Schmertmann (Florida State University), Renato M. Assunção (Federal University of Minas Gerais, Brazil), and Joseph E. Potter (University of Texas-Austin).

Specifically, Stillwell (2005) proposes a series of statistical strategies to model inter-regional migration flows. He indicates that "log-normal" models are limited because they assume that the dependent variable and the error terms have a "log-normal" distribution. Moreover, these models consider that error variances are constant in the different sizes of estimated flows. Statistical models based on the Poisson distribution would generate better estimates because they take into account that the dependent migration variable is measured in discrete units (integer counts of people), and these models have a discrete probability distribution. In the case of migration flows between the 502 Brazilian micro-regions, the Poisson estimation is used because there is a great number of small flows in the origin-destination matrix as well as a small number of larger flows of people. The Poisson regression equation is:

$$M_{ij} = exp\left(b_0 + b_1 \log P_i + b_2 \log P_j + b_3 \log d_{ij}\right) + \varepsilon_{ij}, \qquad (2)$$

where b_o is the constant; b_1 is the regression coefficient associated with the population at the beginning of the period (P_i); b_2 is the coefficient associated with the population at the end of the period (P_j); b_3 is related to the distance between micro-regions (d_{ij}); and ε_{ij} is the random error term associated with all pairs of micro-regions.

To estimate this Equation (2), it is necessary to obtain information about the population at the beginning of the migration period (P_i) as well as at the end of the period (P_j). Because migration is a rare demographic event, the best strategy is to use observed rates from age groups with higher migration flows and to estimate migration rates for the other age groups. The notion of age selectivity in migration is discussed by Rogers and Castro (1981). The authors argue that people between 20 and 25 years of age have the highest rates of migration. The lowest rates are for migrants between 10 and 19 years of age. Migration rates for people between zero and nine years of age are a response of the rates for people between 26 and 40 years.

Because migration flows between the Brazilian micro-regions (502*501=251,502) have a low number of migrants in several cases, a group with a high level of migration rates among all age groups was selected. The group between 20 and 24 years of age was selected to estimate migration flows between micro-regions (*ij*) as well as to calculate the population exposed to the risk of migration at the beginning of the period (P_i) and end of the period (P_j). With the purpose of generating these flows, it is necessary to use migration information that indicates the micro-region (or municipality) of residence in a specific previous moment. Information about the municipality of previous residence (whether the person has lived less than ten years in the present municipality) is available in the 1980 and 1991 Censuses. Information was collected in 1991 and 2000 about the municipality of residence five years before the Census. Because one of the objectives is to develop a methodology that can generate comparative results over time and can be used in future studies, this analysis used information on municipality of residence five years before the 1991 and 2000 Censuses. This migration information allowed the estimation of: (1) the population at the beginning of the period with 15-19 years of age by micro-region of origin, sex, and education group; (2) the population at the end of the period with 20-24 years of age by micro-region of destination, sex, and education group; and (3) migrants at the end of the period with 20-24 years of age by micro-region of origin and destination, sex, and education group. Technically, it would be necessary to estimate those people who died between 1986 and 1991, as well as between 1995 and 2000, to include them as the population at risk of migration. However, this calculation is not made, based on the assumption that the 20-24 age group used for this analysis has a low mortality rate, which does not comprise the estimated results.

Then, a dataset containing information on: (1) micro-region of origin; (2) micro-region of destination; (3) kilometre distances between micro-regions *i* and *j*; (4) population aged 15-19 years at the beginning of the period by sex, education group, and micro-region of origin (1991 and 2000 Censuses); (5) population aged 20-24 years at the end of the period by sex, education group, and micro-region of destination (1991 and 2000 Censuses); and (6) migrants aged 20-24 years at the end of the period by micro-region of origin and destination, sex, and education group (1991 and 2000 Censuses). Following Equation (2), the logarithm of the variables was estimated before the estimation of the Poisson regression. The cells with no migration flows or no population were replaced by zero to be considered in the regression. Furthermore, regressions were estimated only for the cases in which the micro-regions of origin were different than the micro-regions of destination (502*501=251,502 flows) and for men. The results of these models by year and education groups are presented in Table 4.

Table 4 - Poisson estimates of population at the beginning and end of the period, and distance between micro-region centroids on the logarithm of migration flows for men aged 20–24 years (dependent variable) by education group, Brazil, 1991 and 2000.

Variables	1991			2000		
	0–4 years of schooling	5–8 years of schooling	9+ years of schooling	0–4 years of schooling	5–8 years of schooling	9+ years of schooling
Constant	-6.848*** (0.0692)	-5.541*** (0.0642)	-5.325*** (0.0692)	-7.696*** (0.0823)	-6.947*** (0.0677)	-6.400*** (0.0667)
Log of population aged 15–19 years at the beginning of the period in micro-region *i*	0.648*** (0.00507)	0.525*** (0.00434)	0.557*** (0.00468)	0.710*** (0.00615)	0.593*** (0.00457)	0.534*** (0.00408)
Log of population aged 20–24 years at the end of the period in micro-region *j*	0.831*** (0.00513)	0.719*** (0.00399)	0.688*** (0.00409)	0.780*** (0.00590)	0.753*** (0.00417)	0.698*** (0.00376)
Log of KM distances between micro-regions	-1.093*** (0.00503)	-0.965*** (0.00568)	-0.963*** (0.00661)	-1.063*** (0.00560)	-0.973*** (0.00547)	-0.933*** (0.00574)
Observations (n)	251,502	251,502	251,502	251,502	251,502	251,502

Source: 1991 and 2000 Brazilian Censuses.
* Significant at $p<0.1$; ** Significant at $p<0.05$; *** Significant at $p<0.01$.
Standard errors in parentheses.

Based on the coefficients of Table 4, 20- to 24-year-old male migrants in each micro-region of destination (at the end of the period) were estimated by year and education group. Because the migration level is later used to estimate the effects of population flows on the workers' earnings at the end of the period, the immigration rate for each combination of micro-regions of origin and destination are calculated by year and education group. These estimates are considered the migration level between the micro-regions for each year and education group, which are then combined with estimations of migration schedules. Explanations of the methodology used to estimate migration schedules are made in the following sub-sections.

4.2. Estimation of migration schedule

To estimate migration schedules, it is necessary to have age-group-specific migration rates. The idea is to have the most specific migration schedule for each micro-region and period analysed. However, age-specific migration rates for each combination of micro-regions and Census year would generate migration curves with very low rate levels, or even null rates. The solution to this problem is to estimate migration schedules for population flows among the major Brazilian regions (North, Northeast, Southeast, South and Central-West) for each Census year (1991 and 2000). This procedure computes a total of 50 population flows (five regions of origin, five regions of destination, and two Census years). As a method for standardising the information used to estimate the level and pattern of migration, these rates were estimated with information on the municipality of residence five years before each Census. This variable allows for the estimation of inter-major-regional migration as well as intra-major-regional because the information is given at the municipality level.

Such as explained by Amaral (2008), the age-specific emigration rate ($ASER_{x,ij}$) by age group can be estimated using data on place of residence at a fixed time prior to the Census:

$$ASER_{x,ij} = \frac{\Sigma\left(K_{ij}^{x}\right)}{t * \Sigma\left[\frac{\left(K_{i.}^{x}+K_{ii}^{x}\right)+\left(K_{i}^{x}\right)}{2}\right]}, \qquad (3)$$

where $ASER_{x,ij}$ is the age-specific emigration rate from region i to region j for age group x; $K_{x,ij}$ refers to migrants that lived in region i at the beginning of the period and moved to region j at the end of the period for age group x; $K_{x,i.}$ refers to migrants that lived in region i at the beginning of the period and live in another region at the end of the period for age group x; $K_{x,ii}$ is the population that lived in region i at the beginning as well as at the end of the period for age group x; $K_{x,i.} + K_{x,ii}$ is the total population at the beginning of the period for age group x; $K_{x,i}$ is all the population that lived in region i at the end of the period (this is the total population at the end of the period) for age group x; $[(K_{x,i.} + K_{x,ii}) + (K_{x,i})]/2$ is the estimated population at the middle of the period for age group x; and t is the number of years between the date of reference of the Census and the fixed prior time available in the migration question (1991 and 2000 Brazilian Censuses ask where people lived exactly five years before the Census, t=5). An assumption is made to calculate migration rates using this procedure. The rate of migration is the same between those who died during the five years before the Census and those who survived during this same period.

The approach in this study requires the estimation of in-migration, rather than out- migration as a method for obtaining the impact of population flows on earnings at the end of the period. Thus, the age-specific immigration rate ($ASIR_{x,ij}$) by age group was estimated among the five major Brazilian regions in two Censuses (5*5*2=50 flows). The equation is adjusted in the denominator to estimate the population at the middle of the period for the region of destination:

$$ASIR_{x,ij} = \frac{\Sigma\left(K_{ij}^{x}\right)}{t * \Sigma\left[\frac{\left(K_{j.}^{x}+K_{jj}^{x}\right)+\left(K_{j}^{x}\right)}{2}\right]}. \quad (4)$$

Because the intention is to estimate the pattern of migration by age group, the proportional $ASIR_{x,ij}$ were generated, which equal one unit for the sum of all age-group rates, considering one region of origin, one region of destination, and one Census year. It is important to note that the $ASIR_{x,ij}$ were estimated for a total of ten age groups (15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59, and 60–64), which are the scope of this study.

4.3. Modelling migration schedules

After the estimation of immigration rates by age group, the mathematical models proposed by Rogers and Castro (1981) were implemented for these rates. The authors indicate that mortality and fertility models can furnish a suitable foundation to create model-migration schedules. Model schedules are important for all demographic components because they arrange different observed data by similarities in level and pattern. Model schedules are made using information from many observed data collected in various populations. The creation of models for demographic variables is possible because mortality, fertility and migration rates behave according to some predefined limits. Fertility, mortality and migration rates for an age group are highly correlated to the rates of other age groups. The mathematical expressions of the relationships between age groups are the foundation to estimating the model schedules. The development of hypothetical patterns, based on the regularities of different populations, generates models that can be used to correct observed migration schedules by age.

Migration presents selectivity by sex, such as argued by Rogers and Castro (1981). Sex has an influence on the shape and level of population movements. Men's migration rates are higher than women's rates. The highest migration rates for women occur in younger ages because they marry before men. However, age selectivity is even greater, with a great fraction of migration occurring in young ages. The regularities in the migration pattern by age persist for flows between areas of different sizes. For a longer period of time, return migration presents higher rates, as well as rates for non-surviving migrants, which contribute to the underestimation of the number of migration movements that really occurred. The notion of selectivity in migration is applied by Rogers and Castro in their model-migration schedules. The shape and level of migration schedules vary depending on the age of analysis. Migrants between 20 and 25 years of age present the highest rates. The lowest rates are presented for those people between 10 and 19 years of age. Migration rates between zero and nine years of age are a response of rates between 26 and 40 years of age. Migration rates to metropolitan areas (with high levels of services and cultural activities) around the age of 65 years are high. The regularities found in the migration schedules by age helped the development of hypothetical migration models that can be used in population studies with limited or inadequate data.

The main objective of Rogers and Castro (1981) was the estimation of a model-migration schedule, called a "multiexponential function". These migration models enable the classification of different migration patterns. There is also the possibility of analysing the highest migration rates. Through these schedules, it is possible to analyse the influence of older migration to younger migration.

The estimated migration curve is composed of four components related to the labour market. The pre-labour curve is a negative exponential curve from zero to 19 years of age (with a descendent indicator called α_1). The curve for migrants of labour age has a parabolic shape (with μ_2 as the mean age indicator; λ_2 as the ascendant indicator; and α_2 as the descendent indicator). The first ascending half of this parabola represents migration for those between 20 and 25 years of age. The second descending half of the parabola represents migration for those between 26 and 40 years of age. The post-labour curve is a small parabola around the age of 65 years (the mean age indicator is represented by μ_3; λ_3 is the ascendant indicator for this curve; and α_3 is the descendent one). The last parameter of the model schedule is a constant that adjusts the migration rates to the mathematic expression (the indicator is the constant value named *c*). This proposition establishes that migration is highly influenced by economics because the curves indicate different moments of an individual's entrance into the labour market.

Rogers and Castro (1981) created three different model-migration schedules. The most complete schedule is the one named the "basic model migration schedule". This schedule has all four labour-migration components listed above. The curve has a total of 11 parameters. The second migration schedule is defined as "reduced form". This model does not present the post-labour curve around the age of 65 years. Without this last parabola, the reduced form has seven parameters. "Model migration schedule with an upward slope" is the third schedule and has a linear curve in post-labour ages, instead of a parabola. The total number of parameters in this migration curve is nine.

The first model ("basic migration model") has a parabola in post-labour ages and has the following mathematical form:

$$S_x = a_1 * exp(-\alpha_1 x) + a_2 * exp\{-\alpha_2(x-\mu_2) - exp[-\lambda_2(x-\mu_2)]\} + a_3 * exp\{-\alpha_3(x-\mu_3) - exp[-\lambda_3(x-\mu_3)]\} + c, \quad (5)$$

where S(x) denotes the conditional migration rate at age x.

Another model-migration schedule is the "reduced model". This model has a constant value in post-labour ages. The next equation details the mathematical form of this model:

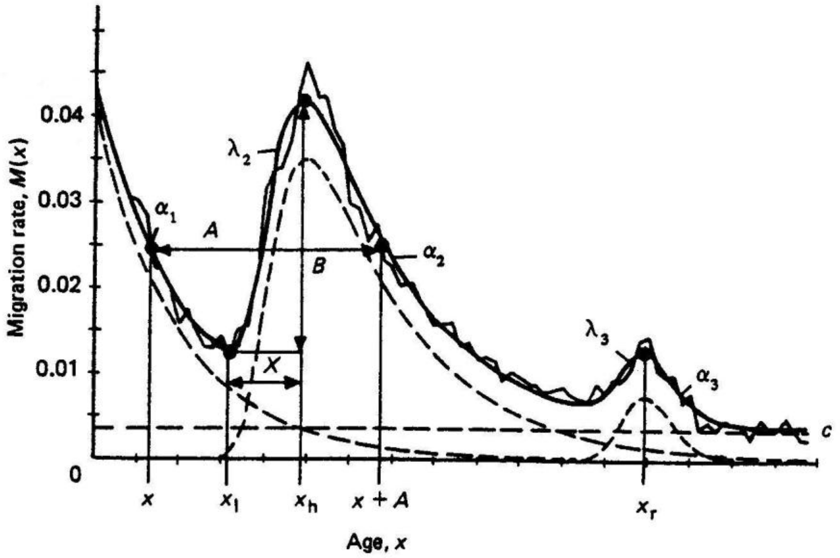
$$S_x = a_1 * exp(-\alpha_1 x) + a_2 * exp\{-\alpha_2(x-\mu_2) - exp[-\lambda_2(x-\mu_2)]\} + c. \quad (6)$$

The last model is the "migration model with an ascending inclination". This model has a linear function in post-labour ages. The following equation illustrates the mathematical form of this model:

$$S_x = a_1 * exp(-\alpha_1 x) + a_2 * exp\{-\alpha_2(x-\mu_2) - exp[-\lambda_2(x-\mu_2)]\} + a_3 * exp(\alpha_3 x) + c. \quad (7)$$

The appropriate model-migration schedule is selected after the analysis of the shape of migration rates among regions. Furthermore, the level of migration rates has to be examined among all regions. Figure 1 illustrates all parameters in the curve for the "basic migration model".

Figure 1 - The model-migration schedule elaborated by Rogers and Castro.



Source: Rogers and Castro (1981, p.6).

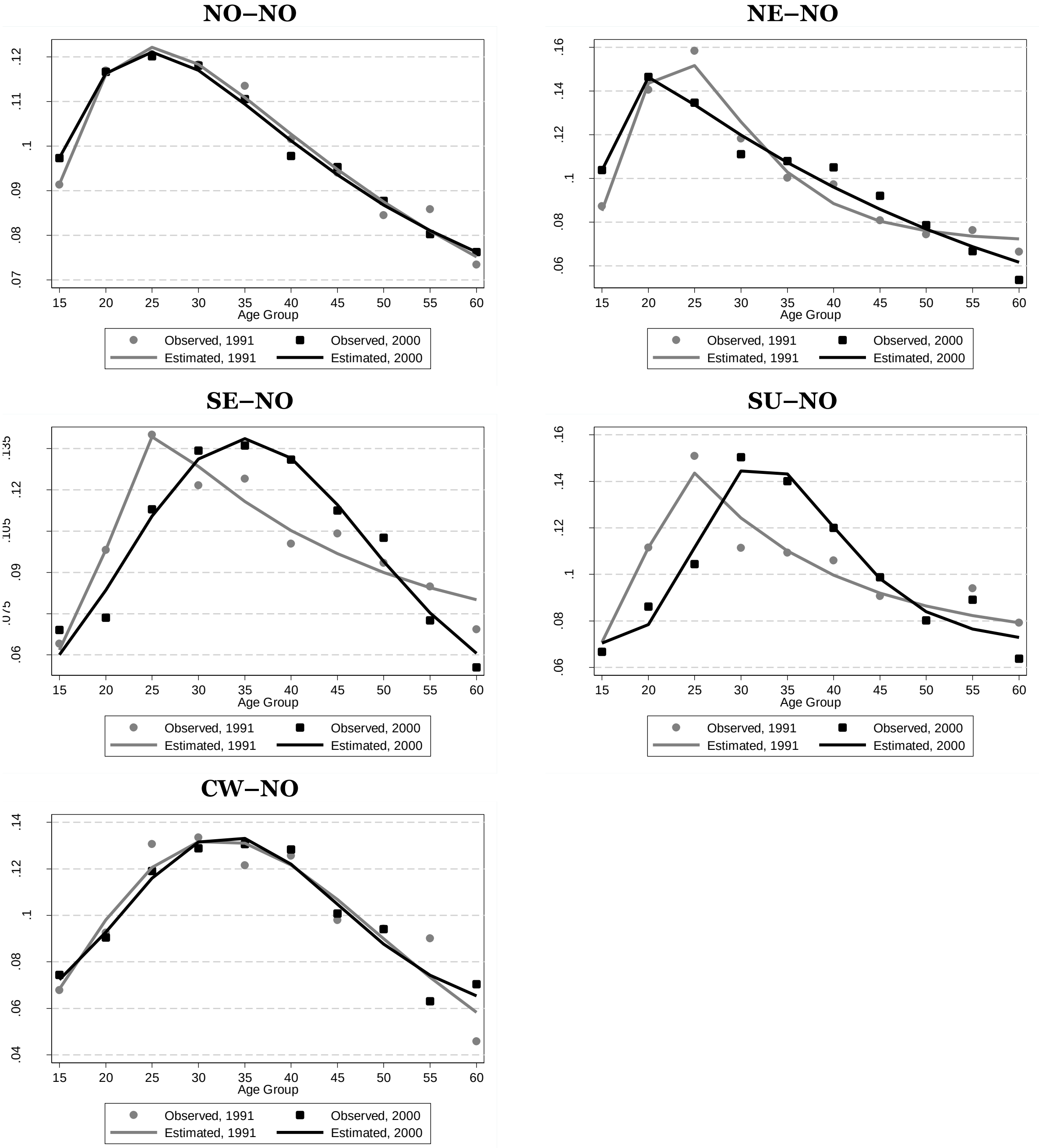
In the migration model, "basic measures" are divided into two different groups. The first group includes the eleven "fundamental parameters". Levels of migration rates are measured by different indicators. The first indicator is the level of migration in pre-labour ages (a_1). The level of migration in labour ages is estimated by a_2. The level of migration in post-labour ages is called a_3. The constant (c) is the last-level estimator. Placements of migration rates are calculated using two means: the mean age in labour ages (μ_2) and the mean age in post-labour ages (μ_3). Model-migration schedules have some parameters for estimating curve slopes. The negative slope of the pre-labour curve is named α_1. Another parameter is the negative slope of the labour curve (α_2). The third indicator is the negative slope of the post-labour curve (α_3). λ_2 is the positive slope of the labour curve. Finally, the positive slope of the post-labour curve is estimated by λ_3.

The second group of "basic measures" includes six "ratios of parameters". The labour force dependency of population is measured by $\delta_{1c}=a_1/c$. The measure $\delta_{12}=a_1/a_2$ is the youth dependency ratio of population. The old-age dependency ratio of population is given by $\delta_{32}=a_3/a_2$. The relationship between child migration and adult migration is calculated by $\beta_{12}=\alpha_1/\alpha_2$. The ratio $\sigma_2=\lambda_2/\alpha_2$ estimates a similarity measure between the first and second halves of the labour curve. Finally, similarity between the first and second halves of the post-labour curve is estimated by $\sigma_3=\lambda_3/\alpha_3$.

"Derivative measures" are divided into three groups to analyse the migration pattern. The first group includes the calculation of four "areas below the migration curve". The total area below the curve is the gross migra-production rate (GMR). The area below the line for age 0-14 is the percentage of migrants in pre-labour ages. The area below the curve for 15-64 years of age is the percentage of migrants in labour ages. The area below the curve for age of 65 until the oldest age is the percentage of migrants in post-labour ages. The second group is related to "four locations in the migration curve": $\bar{n}$ is the mean age of migration; x_1 is the low point, located at the intersection of the descending pre-labour force component and the ascending labour force component; x_h is the high peak, located at the intersection of the ascending labour force component and the descending labour force component; and x_r is the retirement peak, located at the intersection of the ascending post-labour force component and the descending post-labour force component. Finally, the third group of derivative measures refers to three "distances among parameters". The distance between the middle pre-labour rate and middle labour rate is represented by the parameter *A* (parental shift). The vertical distance between the lowest pre-labour rate and highest labour rate is calculated by the parameter *B* (jump). The horizontal distance between the lowest pre-labour rate and highest labour rate is calculated by *X* (labour force shift).

Rogers and Jordan (2004) indicate that migration flows are usually modelled with the "reduced model" in Equation (6) because there is no evident parabola or linear function in post-labour ages. Because this present study has the purpose of correcting the income estimates of 15- to 64-year-old men, the modelling of migration rates was conducted based on the second component (the parabola for migrants in labour ages) and the fourth component (constant to adjust the mathematical expression). In other words, Equation (6) was used without the first component (pre-labour curve) because this curve refers to the population of less than 15 years of age. The use of these mathematical models was developed with the computer software "Table Curve 2D". Figures 2 to 6 illustrate the proportional age-specific immigration rates ($ASIR_{x,ij}$), estimated with Equation (4) and modelled with Equation (6), for the 50 population flows among the major Brazilian regions. The vertical axes are not uniform to allow a better visualisation of the observed and modelled migration schedules for each population flow.

Figure 2 - Observed and estimated proportional age-specific immigration rates (ASIR) of flows to the North region[1], 1991 and 2000.[2]

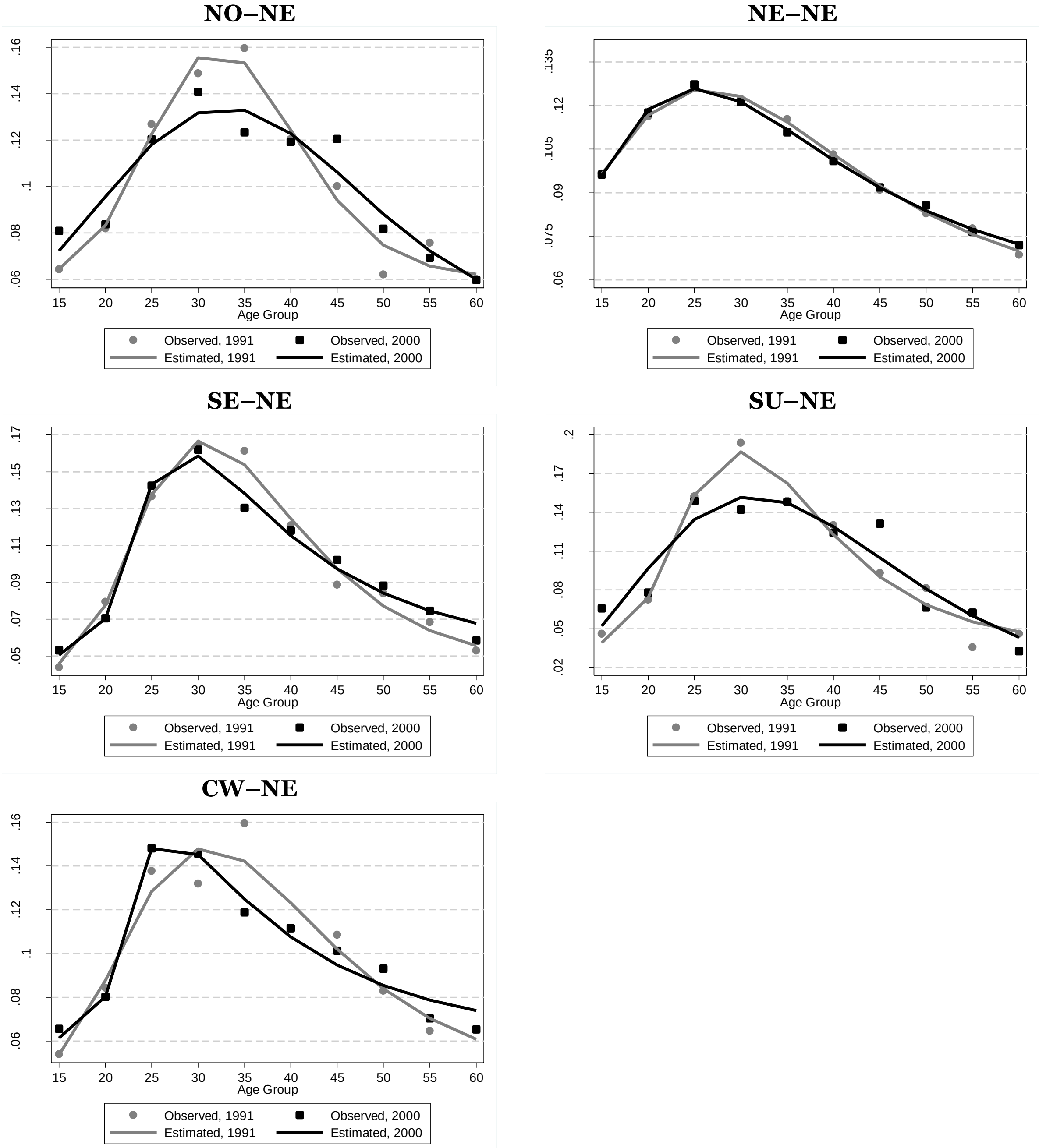


Source: 1991 and 2000 Brazilian Censuses.

[1] North region (NO); Northeast region (NE); Southeast region (SE); South region (SO); Central-West region (CW).

[2] Migration flows were estimated with information on municipality of residence five years before each Census reference date. This allows for the estimation of intra-regional migration in each period.

Figure 3 - Observed and estimated proportional age-specific immigration rates (ASIR) of flows to the Northeast region[1], 1991 and 2000.[2]

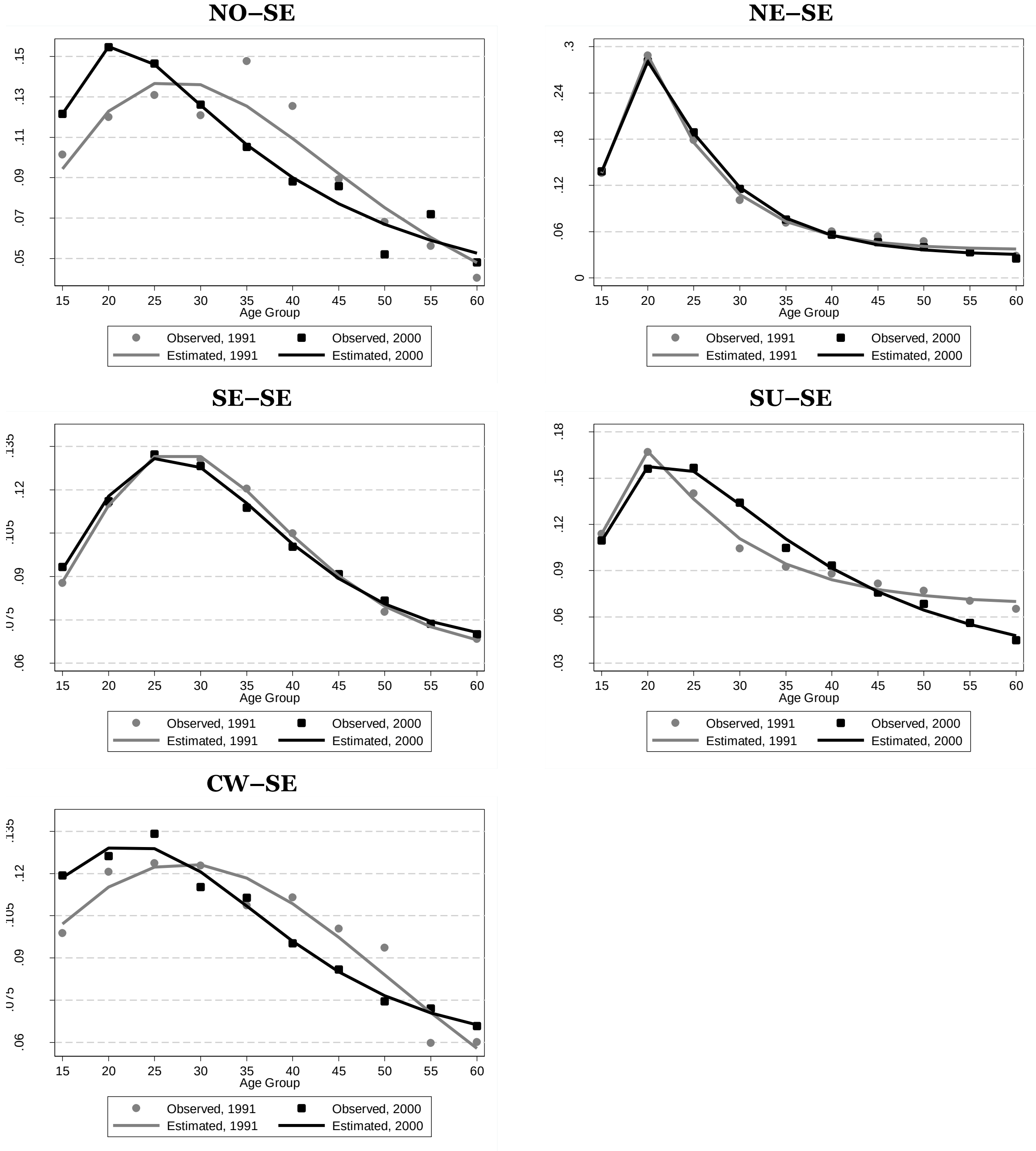


Source: 1991 and 2000 Brazilian Censuses.

[1] North region (NO); Northeast region (NE); Southeast region (SE); South region (SO); Central-West region (CW).

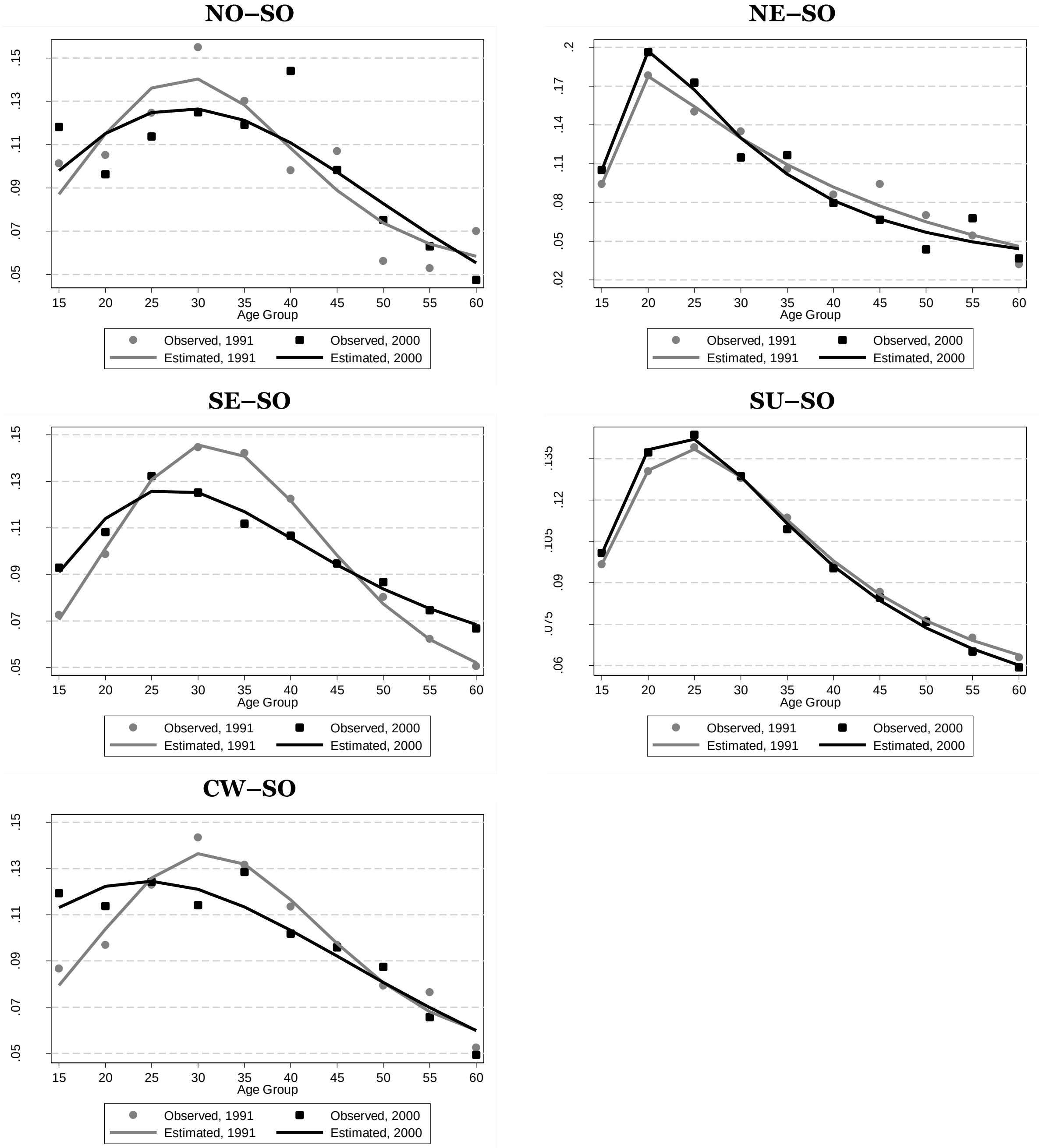
[2] Migration flows were estimated with information on municipality of residence five years before each Census reference date. This allows for the estimation of intra-regional migration in each period.

Figure 4 - Observed and estimated proportional age-specific immigration rates (ASIR) of flows to the Southeast region[1], 1991 and 2000.[2]



Source: 1991 and 2000 Brazilian Censuses.
[1] North region (NO); Northeast region (NE); Southeast region (SE); South region (SO); Central-West region (CW).
[2] Migration flows were estimated with information on municipality of residence five years before each Census reference date. This allows for the estimation of intra-regional migration in each period.

Figure 5 - Observed and estimated proportional age-specific immigration rates (ASIR) of flows to the South region[1], 1991 and 2000.[2]

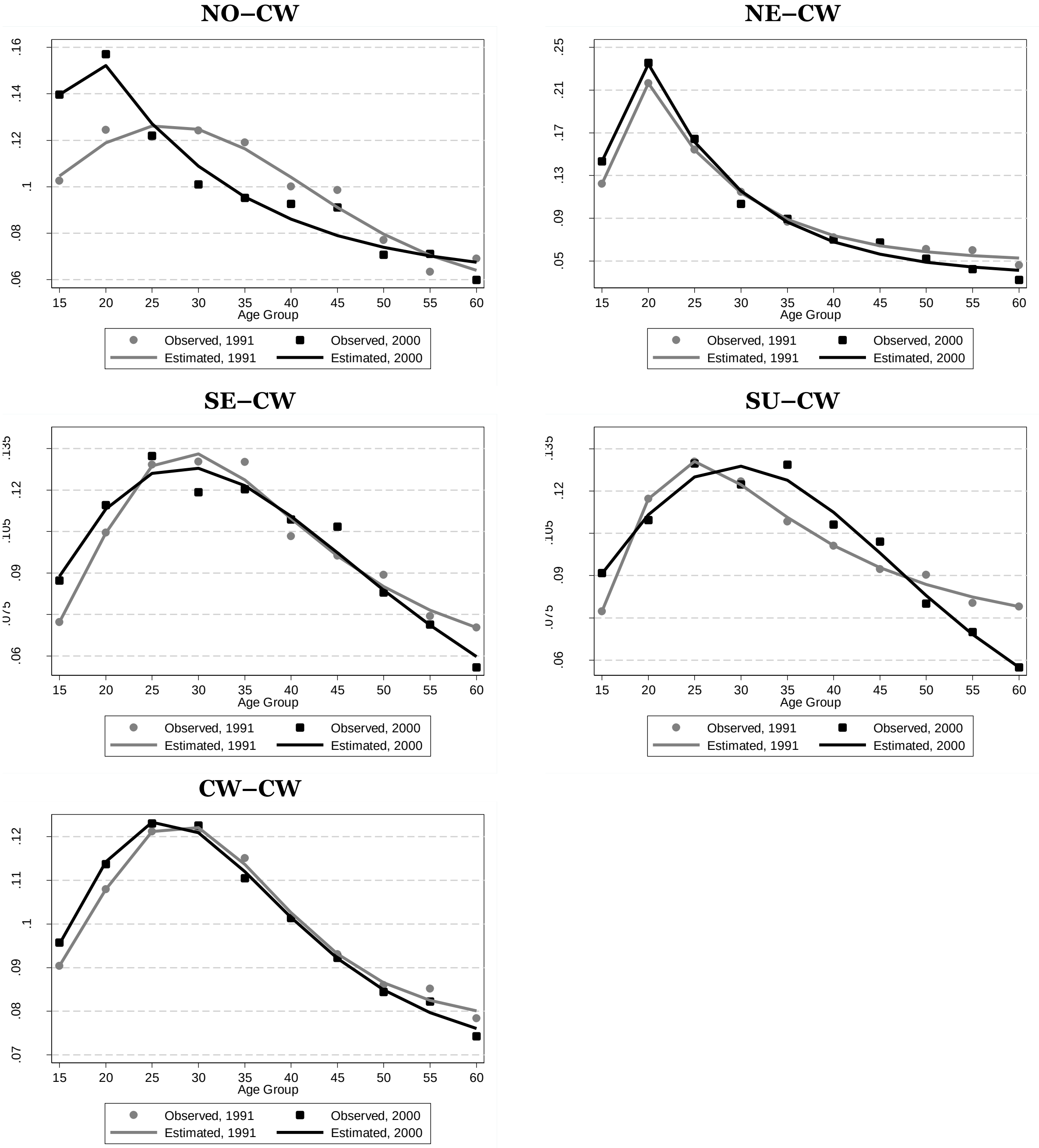


Source: 1991 and 2000 Brazilian Censuses.

[1] North region (NO); Northeast region (NE); Southeast region (SE); South region (SO); Central-West region (CW).

[2] Migration flows were estimated with information on municipality of residence five years before each Census reference date. This allows for the estimation of intra-regional migration in each period.

Figure 6 - Observed and estimated proportional age-specific immigration rates (ASIR) of flows to the Central-West region[1], 1991 and 2000.[2]



Source: 1991 and 2000 Brazilian Censuses.
[1] North region (NO); Northeast region (NE); Southeast region (SE); South region (SO); Central-West region (CW).
[2] Migration flows were estimated with information on municipality of residence five years before each Census reference date. This allows for the estimation of intra-regional migration in each period.

4.4. Integrating level and schedule of migration

After the procedures detailed above, it was possible to estimate the level of migration for 20- to 24-year-old men between the 502 micro-regions by education group (0-4, 5-8, and 9+) and year (1991 and 2000). Moreover, the age-specific immigration rates ($ASIR_{x,ij}$) were modelled for each of the 50 population flows among the five major regions by year. Then, the ratio between the level of migration for 20- to 24-year-old men (for each one of the pairs of micro-regions of origin and destination) and the age-specific immigration rates of the 20- to 24-year age group ($ASIR_{20\text{-}24,ij}$) was calculated. This ratio took into account the migration pattern by major region of origin and destination corresponding to the migration level by micro-regions of origin and destination. The ratio was then multiplied by each $ASIR_{x,ij}$ of the other age groups, considering the education group and year. In other words, the migration level between the micro-regions (502 micro-regions of origin * 501 micro-regions of destination * 3 education groups * 2 Census years) was applied to the migration patterns of the major Brazilian regions (5 major regions of origin * 5 major regions of destination * 2 Census years).

Following this procedure, the total non-immigration rate ($TNIR_{ij}$) was estimated for each one of the twelve age-education groups related to the earning models, for each Census and each combination of micro-region of origin and destination. Based on Amaral (2008), this rate is estimated using the $ASIR_{x,ij}$:

$$TNIR_{g,ij} = exp\left(-\sum ASIR_{x,e,ij}\right), \tag{8}$$

where g represents the combination of age (15–24, 25–34, 35–49, and 50–64) and education (0–4, 5–8, and 9+) groups, and there are twelve combinations; x is the five-year age group; and e is the education group.

The total immigration rate (TIR_{ij}) was calculated using the $TNIR_{ij}$:

$$TIR_{ij} = 1 - TNIR_{ij}. \tag{9}$$

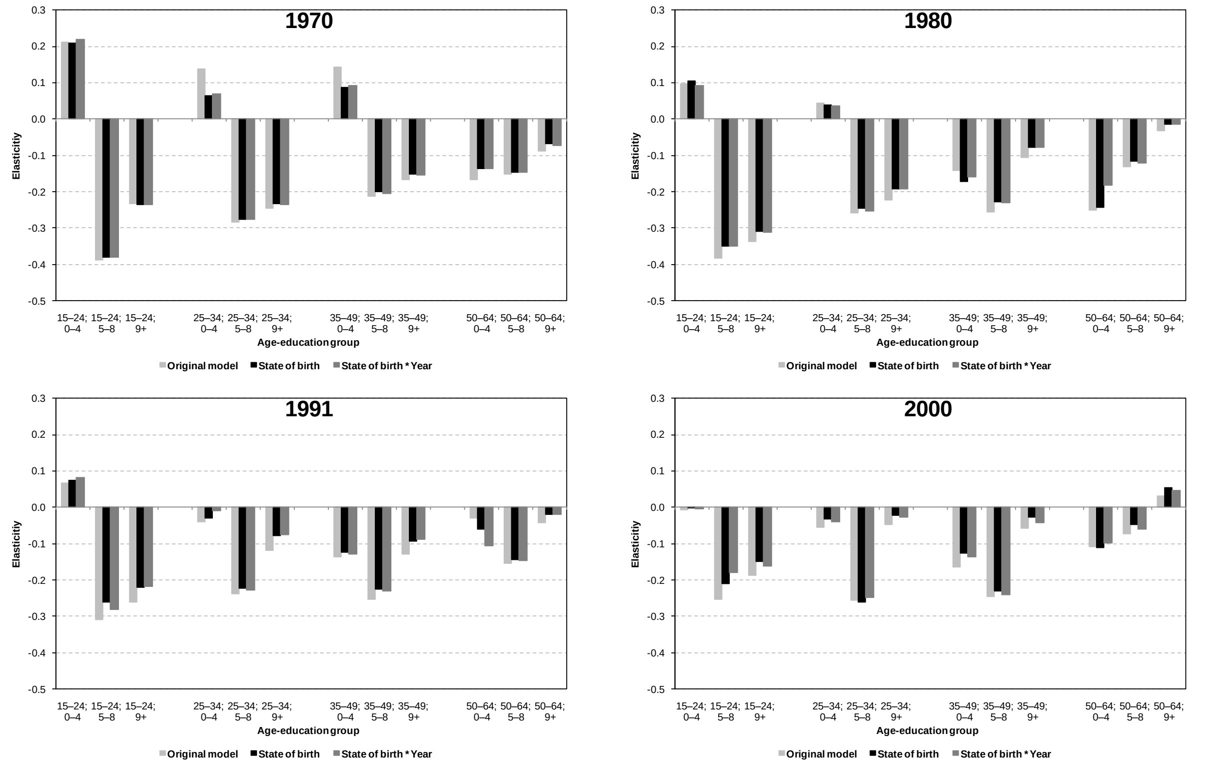
Finally, a measure of force of migration was estimated as the sum of all total immigration rates for each micro-region of destination. In other words, all TIR_{ij} from micro-regions of destination to a specific micro-region of destination were added, considering each age-education group and year. This procedure generated a total of 12,048 measures of force of migration (502 micro-regions * 12 age-education groups * 2 Census years). These estimated migration rates were merged with the data on workers' earnings (Amaral et al. 2007). The new measures of migrants' distributions by age-education groups allow for the estimation of exogenous impacts of migration on the income of workers, removing the endogeneity of the original migration variables.

5. Results

To verify whether the estimated migration rates generate more suitable results than the ones provided by observed migration variables, a set of models will be compared in this section. Figures 7 to 10 present the estimated elasticities from the original model (Table 3) and from a set of models that include migration proportions by age-education groups. The elasticities were calculated as the product of age-education-proportion coefficients and the distribution of men by age-education groups over time (Table 1).

Figure 7 illustrates estimated elasticities of age-education-group proportions, after controlling for the model by proportion of migrants in age-education groups, using the 1970-2000 Censuses. State of birth was used to categorise individuals as migrants or naturals in each micro-region. The light grey bars are the estimates from the original model (Table 3). The black bars are the estimates from the model using proportion of migrants in twelve age-education groups. The dark grey bars are the estimates using the same proportions of migrants interacted with the year. Although the hypothesis was that after controlling for migration, the negative impacts of proportions of people in age-education groups would increase negatively, Figure 7 indicates a lower negative impacts on earnings for the migration models. Because the state of birth takes into account migration movement over a long period of time, the elasticities of these models might not have had greater negative impacts because, as proposed by Rogers and Castro (1981), these migration movements are underestimated, due to return migration and non-surviving migrants.

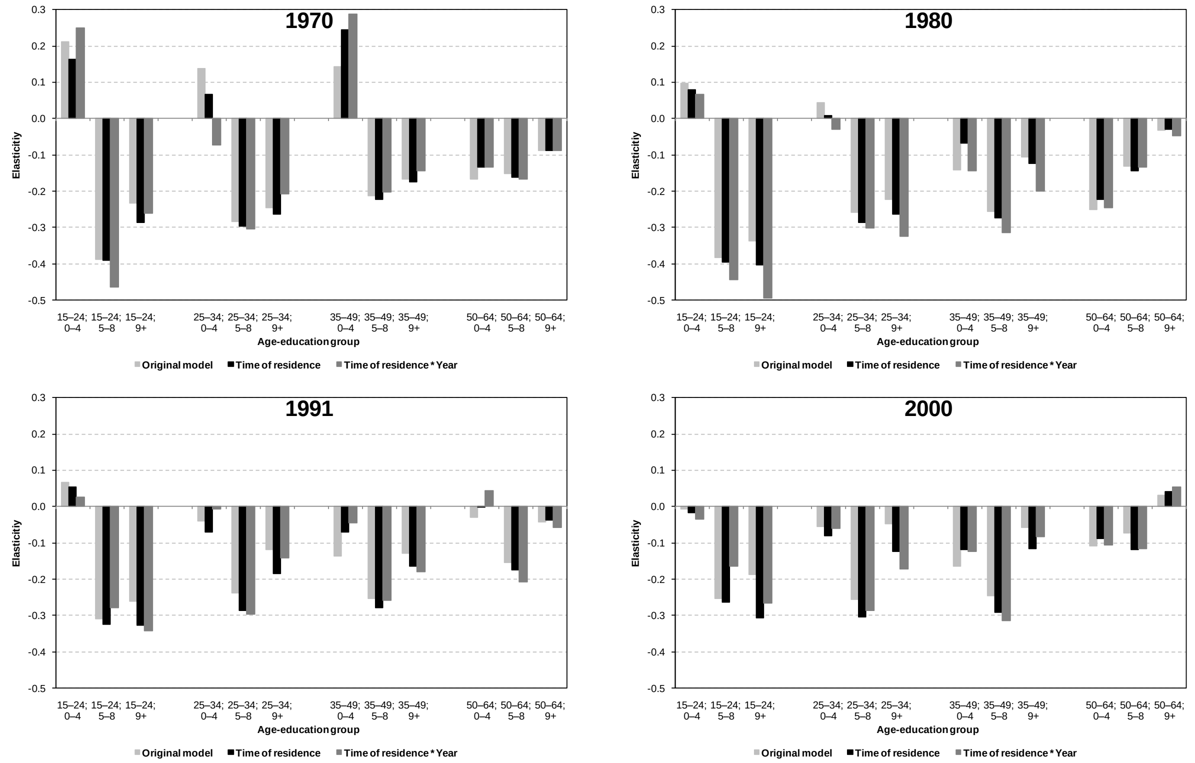
Figure 7 - Estimated elasticities of population proportions in age-education groups from original model (Table 3) and migration models (state of birth), 1970–2000.



Source: 1970–2000 Brazilian Censuses.

Figure 8 shows elasticities using the original model (Table 3), the model with the proportion of migrants in age-education groups, taking into account the classification of migrants as those that live less than five years in the municipality (time of residence), as well as interactions of these proportions of migrants with the year. In general, elasticities presented in Figure 8 have a greater negative impact on earnings than the ones of the original model. However, this trend is not consistent across all age-education groups, which might be an indication that more precise migration variables should be used in the models.

Figure 8 - Estimated elasticities of population proportions in age-education groups from original model (Table 3) and migration models (time of residence), 1970–2000.

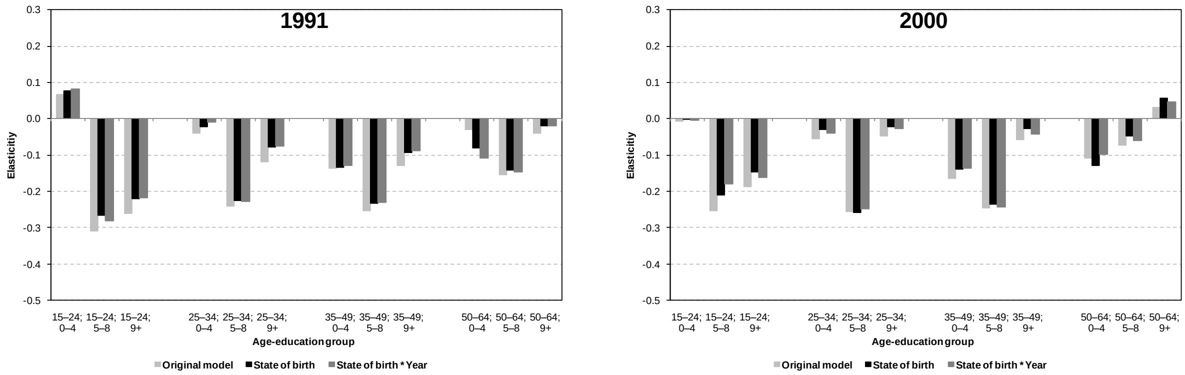


Source: 1970–2000 Brazilian Censuses.

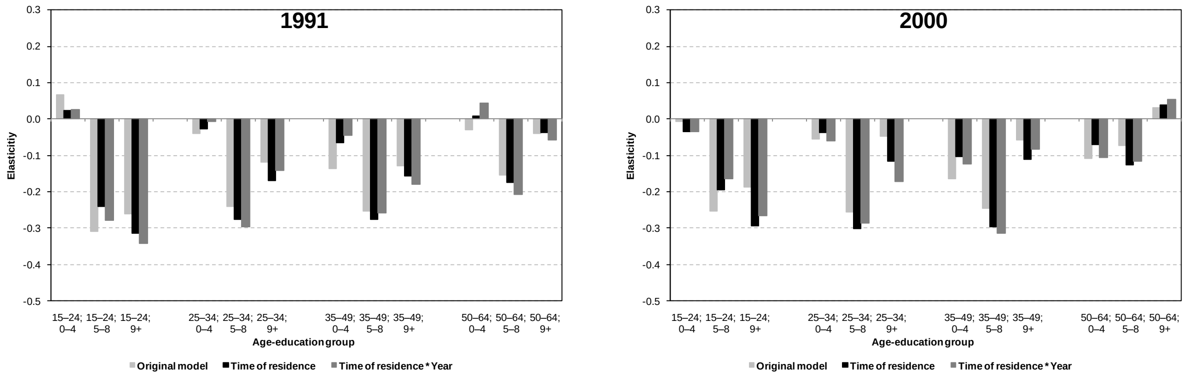
Such as explained in previous sections, the correction of the level and pattern of migration is possible only for the 1991 and 2000 Censuses because of the inclusion of information on municipality of residence five years before the Census. Before the introduction of this migration variable into the models, it was important to verify whether models using only the 1991 and 2000 Censuses had the same level and composition of estimated elasticities by age-education groups as those ones presented in Figures 7 and 8. Figure 9 was generated with this purpose and illustrates that the behaviour of elasticities for the models with information on state of birth have the same pattern as those ones presented by the two bottom graphs of Figure 7. The same pattern is also observed for elasticities generated using information on time of residence, comparing the two bottom graphs of Figures 8 and 9.

Figure 9 - Estimated elasticities of population proportions in age-education groups from original model and migration models (state of birth, and time of residence), 1991 and 2000.

State of birth



Time of residence

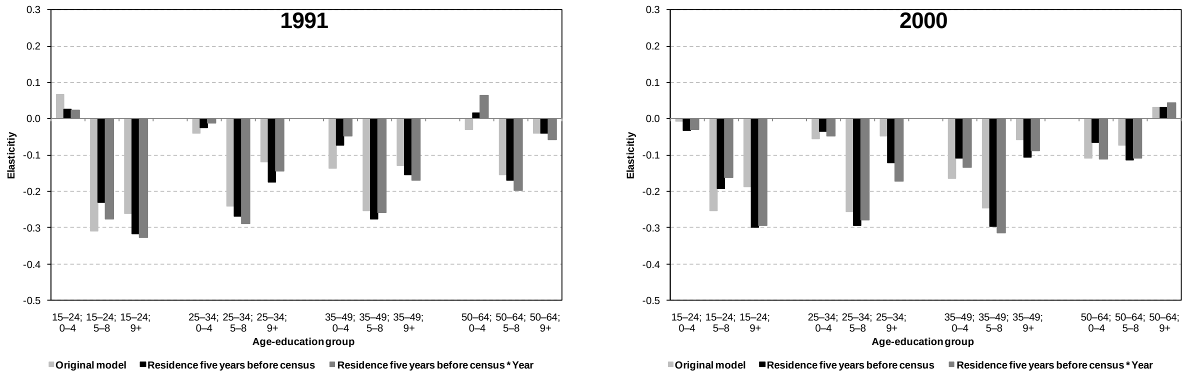


Source: 1991 and 2000 Brazilian Censuses.

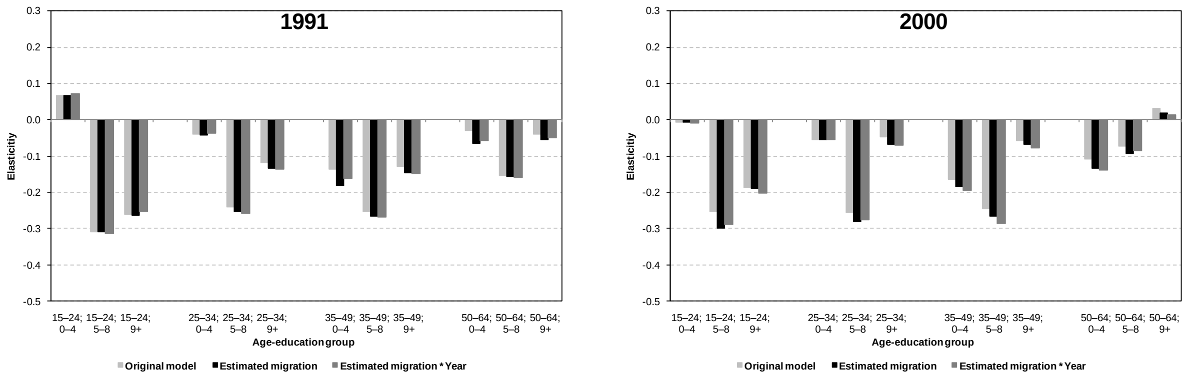
Information was included in the estimated models on municipality of residence five years before the Censuses. The top two graphs in Figure 10 illustrate results from the inclusion of the proportion of migrants in age-education groups, taking into account municipality of residence in a fixed prior time as well as interactions of these proportions of migrants with the year. As expected, these two graphs have similar results to the two bottom graphs in Figure 9 because migration information on time of residence (whether the individual lived less than five years in the municipality) obtains the same kind of migration movement as information on municipality of residence five years before the Census.

Figure 10 - Estimated elasticities of population proportions in age-education groups from original model and migration models (residence five years before census, and estimated migration), 1991 and 2000.

Residence five years before census



Estimated migration



Source: 1991 and 2000 Brazilian Censuses.

Finally, the estimated migration using information on municipality of residence five years before the Census, corrected by level (Stillwell 2005) and pattern (Rogers and Castro 1981) techniques, is used to generate the elasticities presented in the two bottom graphs of Figure 10. Elasticities are more negative than the original ones (Table 3) for all age-education groups with no inconsistencies, following the hypothesis of this study. The only positive elasticities are observed in the first age-education group in 1991 and the last age-education group in 2000, which are also consistent with the original results.

Final considerations

The purpose of this paper was to include internal migration flows in models that had previously estimated the impact of age and educational changes on workers' earnings in a micro-regional analysis (Amaral et al. 2007). The findings of this study follow the initial hypothesis, which stated that by controlling for migration flows, the negative impact of cohort size on earnings would be even more negative than estimates that did not take into account population flows. Moreover, the inclusion of internal migration estimates in the model has consistent results in the labour market models only when they have had their level and pattern previously adjusted. A methodological approach was developed in this analysis, integrating strategies proposed by Stillwell (2005) and Rogers and Castro (1981). Although these adjusted-migration estimates were made only for the 1991 and 2000 Censuses, these strategies were designed in such a way that they can be used in further studies when new data become available, as well as in the context of other countries with the availability of migration data.

References

AMARAL, E. F. L. Improvements of techniques to estimate migration rates: an application with Brazilian censuses data, **Population Review** 47(2): 1-24, 2008.

AMARAL, E. F. L.; HAMERMESH D. S.; POTTER, J. E.; RIOS-NETO, E. L. G. Demographic change and the structure of wages: a demand-theoretic analysis for Brazil, **NBER Working Paper Series** 13533. (online: http://www.nber.org/papers/w13533), 2007.

BERGER, M. The effect of cohort size on earnings growth: a reexamination of the evidence, **The Journal of Political Economy** 93(3): 561-573, 1985.

BLOOM, D. E; CANNING, D.; SEVILLA, J. **The Demographic Dividend: a New Perspective on the Economic Consequences of Population Change**. Santa Monica: RAND, 2003.

BLOOM, D. E.; FREEMAN, R. B.. The effects of rapid population growth on labor supply and employment in developing countries, **Population and Development Review** 12(3): 381-414, 1986.

BORJAS, G. J. The labor demand curve is downward sloping: reexamining the impact of immigration on the labor market, **The Quarterly Journal of Economics** November: 1335-1374, 2003.

EASTERLIN, R. A. What will 1984 be like? Socioeconomic implications of recent twists in age structure, **Demography** 15(4): 397-432, 1978.

FREEMAN, R. B. The effect of demographic factors on age-earnings profiles, **The Journal of Human Resources** 14(3): 289-318, 1979.

MASON, A. Demographic transition and demographic dividends in developed and developing countries, **Population Division, Department of Economic and Social Affairs**, *United Nations* UN/POP/PD/2005/2.

MASON, A.; FENG, W. Demographic dividends and China's post reform economy, **XXV IUSSP General Population Conference**. Tours: IUSSP, 2005.

MINCER, J. Investment in human capital and personal income distribution, **The Journal of Political Economy** 66(4): 281-302, 1958.

MINCER, J. **Schooling, Experience, and Earnings**. New York: National Bureau of Economic Research, 1974.

OLIVEIRA, K. F.; JANNUZZI, P. M. Motivos para migração no Brasil e retorno ao Nordeste [Reasons of migration in Brazil and return to the Northeast], **São Paulo em Perspectiva** 19: 134-143, 2005.

POTTER, J. E.; SCHMERTMANN, C. P.; CAVENAGHI, S. M. Fertility and development: evidence from Brazil, **Demography** 39: 739-761, 2002.

POTTER, J. E.; SCHMERTMANN, C. P.; ASSUNÇÃO, R. M.; CAVENAGHI, S. M. Mapping the timing, pace, and scale of the fertility transition in Brazil. **Population and Development Review** 36(2): 283-307, 2010.

ROGERS, A.; JORDAN, L. Estimating migration flows from birthplace-specific population stocks of infants, **Geographical Analysis** 36(1): 38-53, 2004.

ROGERS, A.; CASTRO, L. J. **Model Migration Schedules**. Laxenburg: International Institute for Applied Systems Analysis, 1981.

STILLWELL, J. Inter-regional migration modeling: a review and assessment, **Forty Fifth Congress of the European Regional Science Association (ERSA).** Vrije: ERSA, 2005.

TRIEST, R. K.; SAPOZHNIKOV, M.; SASS, S. A. Population aging and the structure of wages, **Manuscript from the Center for Retirement Research at Boston College,** 2006.

WELCH, F. Effects of cohort size on earnings: the baby boom babies' financial bust, **The Journal of Political Economy** 87(5)2: S65-S97, 1979.

WILLIAMSON, J. G. Demographic change, economic growth, and inequality, in N. Birdsall, A. C. Kelley, and S. W. Sinding (eds.), **Population Matters**. Oxford: Oxford University Press, pp. 106-136, 2003.