UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA APLICADA

EDUCATION AND LABOUR MARKET OUTCOMES: EVIDENCE FROM BRAZIL

Ricardo Freguglia
Gisele Spricigo
Geraint Johnes
Aradhna Aggarwal

TD. 001/2011
Programa de Pos-Graduação em Economia Aplicada - FE/UFJF

Juiz de Fora

2011

EDUCATION AND LABOUR MARKET OUTCOMES: EVIDENCE FROM BRAZIL

Ricardo Freguglia[1]
Gisele Spricigo[2]
Geraint Johnes[3]
Aradhna Aggarwal[4]

1.Federal University of Juiz de Fora, Juiz de Fora, Minas Gerais, Brazil; ricardo.freguglia@ufjf.edu.br
2.Universidade do Vale do Rio dos Sinos, São Leopoldo, Rio Grande do Sul, Brazil; giseles@unisinos.br
3.Lancaster University Management School, Lancaster, UK; G.Johnes@lancs.ac.uk
4.University of Delhi, Delhi, India; aradhna.aggarwal@gmail.com

This version: January 2011

ABSTRACT

The effect of education on labour market outcomes is analysed using both survey and administrative data from The Brazilian PNAD and RAIS-MIGRA series, respectively. Occupational destination is examined using both multinomial logit analyses and structural dynamic discrete choice modelling. The latter approach is particularly useful as a means of evaluating policy impacts over time. We find that policy to expand educational provision leads initially to an increased take-up of education, and in the longer term leads to an increased propensity for workers to enter non-manual employment.

Keywords: occupation, education, development
JEL Classification: I20, J62, O20

While retaining full responsibility for the contents of this paper, the authors gratefully acknowledge support from the Economic and Social Research Council (grant RES-238-25-0014). Thanks, for useful discussions, are also due to participants at the ESRC Research Methods Festival held in Oxford in July 2010.

1. Introduction

Education has been seen as a route to prosperity by the governments of many countries, both in the developed and developing world. In the former countries, education has been promoted as a means of securing a comparative advantage in the production of goods and services that embed a high degree of human capital. Many less affluent countries have likewise seen education as a route to development. This raises the question of how successful such policies can be: are countries merely leapfrogging one another in a zero-sum game, or does education offer prosperity for all? Particularly interesting in this context are the positions of the BRIC countries – Brazil, Russia, India and China – since these are developing rapidly and offer some contrasting stories.

In Brazil, educational provision, particularly at tertiary level, has expanded rapidly over the last decade and a half. The enhanced skills with which many young people now enter the labour force are likely to impact upon their trajectory through the labour market. In particular, we might expect an increasing proportion of workers to find employment in higher status occupations – typically non-manual jobs in the formal sector. Yet there remains remarkably little evidence specific to the Brazilian context on this issue. This paper represents an attempt to examine the data and come to conclusions about the likely direction and magnitude of future change as the labour market responds to recent developments in education policy.

The paper is structured as follows. We begin with a brief discussion of recent changes in education policy in Brazil. We then survey the literature. There follows a short methodological section. Data sources are then discussed, followed by a presentation of the results of our estimation exercises. The paper ends with a discussion and conclusion.

2. Education Policy in Brazil: A brief overview of recent changes

Recent development of the Brazilian educational system is best viewed as part of the process of democratic consolidation, marked by a new institutional arrangement which is characterized by high degree of autonomy of the three levels of government and hence also the decentralization of educational policy. The Constitution of 1988, Constitutional Amendment No. 14 of 1996 and the new Law of Directives and Bases of Education - LDB, established by Law No. 9394, enacted in 1996, are the major laws governing the current Brazilian educational system .

The structure of the educational system comprises basic education - formed by kindergarten, elementary and secondary education - and higher education. Elementary and early childhood education are the responsibility of the municipalities, while later stages of compulsory education are the responsibility of the district and federal states. The federal government, meanwhile, has broad oversight of educational matters, performing a redistributive and supplementary function, and providing specific technical and financial assistance to States, the federal districts and municipalities. Moreover, the federal government organises the higher education system.

Kindergarten, the first stage of basic education is offered in the form of daycare, for children up to 3 years old and in preschool for children aged 4-6 years. Beyond this, elementary schooling, with a minimum term of eight years, is compulsory and free in public schools, , including those who had no access to it at the proper age. According to the LDB is the duty of parents or guardians to register children at this level of education, from the age of 7 years on. Secondary school, the final stage of basic education, lasts a minimum of three years; this meets the general educational needs of the student and may also include vocational programmes in preparation for the world of work and, optionally, also for professional qualifications.

Beyond the traditional forms of formal schooling, there is provision in the Brazilian system for special education and adult education. Technical education, typically delivered independently of the regular high school system, is a requirement for obtaining the diploma of technician. Higher education includes undergraduate courses in a variety of professional areas, open to candidates who have completed high school or equivalent and who have successfully progressed through the recruiting process.

The structure of the Brazilian educational system is summarised in Table 1. A summary of registration/enrollment data can be seen in Table 2.

3. Literature

In many respects, the most obvious forerunner of the work undertaken in the present paper is a contribution by Duflo (2004) who examines the impact of a policy decision rapidly to expand the education sector in Indonesia. Duflo's work focuses, however, on the wage and labour market participation impacts of the policy on various demographic groups, and is in this sense an analysis of trends at a macro level. In contrast to this, our work drills down to the experience of the individual, and focuses on the choice that individuals make about their activity in each period – whether that activity be schooling, work in one occupation or another, or something else.

Early work in the analysis of occupational choice stems from the seminal contribution of Roy (1951) who provides an admirably lucid exposition of the way in which destination depends upon skills and upon the distribution of returns to skills in each occupation. The empirical implementation of Roy's ideas had to await the development of appropriate econometric tools, however. The multinomial logit model, developed by Nerlove and Press (1973), is in many respects the obvious tool for analysing this type of problem.

This model involves the use of maximum likelihood methods to choose the appropriate parameter estimates in the expressions

$$P(Y=j) = \frac{e^{\delta'_j z_i}}{1+\sum_{k=1}^{J} e^{\delta'_k z_i}}, \ j=1,2,...,J$$

$$P(Y=0) = \frac{1}{1+\sum_{k=1}^{J} e^{\delta'_k z_i}} \qquad (1)$$

where the δ terms are parameters and the z are the explanatory variables.

The multinomial logit method, while instructive, does suffer some drawbacks. The first, well documented in the literature, is that it makes an assumption of the independence of irrelevant alternatives. That is, it is assumed that the relative odds between two alternative outcomes are unaffected by augmenting the set of possible outcomes. In some contexts – particularly where the qualitative characteristics of the added regime are close to one but not the other of the two alternatives under study – this assumption is clearly absurd. Several partial fixes for this problem have been suggested in the literature, including nested logit and mixed logit methods.[1] In the present paper we adopt a different approach – that of dynamic discrete choice modelling. The dynamic model links theory to empirical application by adopting a structural approach in which all possible regime choices are included, and, at each date, experience in each regime determines the instantaneous returns to each regime.

A second, rather obvious, feature of the static multinomial logit analysis that is unappealing in the present context is that it is poorly equipped to investigate the impact of policy changes. In particular, the long term impact of an instantaneous change in education policy – where education is usefully regarded as an investment in an individual's future labour market performance – is not readily captured in a static analysis. For this reason too, use of a dynamic approach is appealing.

The essentially dynamic nature of occupational choice was first addressed by Willis and Rosen (1979) who model the decision of when to leave education as an optimal stopping problem. In their model, there is only one post-school outcome, rather than a multiplicity of destinations (including various occupations and life outside the labour force). A solution to this type of problem is offered also by Rust (1987) who developed the nested fixed point algorithm as a means of solving such dynamic stopping models. The extension of this type of model to the case in which, at each point in time, agents make decisions across a multiplicity of options, and where these decisions are conditioned upon decisions made in the past (and determine the nature of options available in the future) is due to Keane and Wolpin (1994, 1997). In effect, the Keane and Wolpin method provides a means of empirically estimating models that combine the salient features of the contributions of Roy, on the one hand, and Willis and Rosen, on the other.

The essence of the problem identified by Keane and Wolpin is very simple. In each period, individuals choose between activities. The instantaneous return to each activity depends upon past experience which is made up of the schooling and labour market choices that the individual has made in the past. In each period the choice made by the individual therefore impacts on the returns that she can make not only in that period but in every subsequent period. For an individual seeking to maximise her lifetime returns, the state space is therefore huge. Empirical evaluation of such a model requires the adoption of approximation methods. Keane and Wolpin propose the evaluation of expected future returns at a sample of points in the state space, fitting a regression line on the basis of this sample, and using this line to estimate expected future returns for points outwith the sample. Using these estimates allows us then to proceed to estimate the parameters of the model in the usual way, using maximum likelihood. We use the variant of the Keane and Wolpin model that allows for regime-specific shocks to be serially correlated.

A feature of the structural modelling approach used here is the close relationship between the theoretical model and the empirical implementation. The analyst begins with an assumed specification of the model, and estimates this model.[2] For this reason, empirical applications of this kind are often referred to as structural models. While attractive in the sense that this approach involves the estimation of the parameters of the theoretical model itself, there are some disadvantages. First, a reader might wish to quibble with the precise specification being assumed in the theoretical model; since the empirical implementation is so closely linked to

[1] Soopramanien and Johnes (2001) offer an example of the use of such methods in the context of occupational choice.
[2] This contrasts with more usual practice, which is to develop some theory and then use regression analysis to test whether or not a particular variable influences another in a particular direction consistent with that theory.

that particular specification, such a quibble assumes empirical importance. Secondly, the close link between theory and estimation means that generic software cannot be developed to estimate models of this kind. In effect, the whole program must be rewritten from scratch each time the specification of the model is subject to a minor modification. These issues have been widely discussed in the literature. Keane (2010), for example, has noted that 'structural econometric work is just very hard to do' – and so is not fashionable. We recognise this; we invite the reader therefore to go along with our story while appreciating that no small aspect of the story can be easily tweaked.

In one important respect, our task has been easier than that of earlier researchers in this area. A recent survey of structural dynamic discrete choice models by Aguirregabiria and Mira (2010) is accompanied by a website[3] that offers software that has been used by earlier researchers to estimate these models.[4] The software is written in high level languages (the Keane and Wolpin program, for example, is in fortran), and requires considerable adaptation before being used to estimate even models that are very similar to those evaluated in the original applications. It nevertheless provides a useful starting point.

Both static and dynamic models of occupational choice have been widely applied to the analysis of occupational choice in developed economies. Variants of the static model have been employed by, *inter alia*, Boskin (1974), Schmidt and Strauss (1975), Ham (1982), Makepeace (1996) and Johnes (1999). In these examples, the emphasis has been on the development of a structural model in which wages explicitly play a key role. In the present paper, we finesse this issue by opting to model a reduced form in which wages do not explicitly appear, but where the determinants of wages are included as explanatory variables. The seminal contribution in the area of dynamic modelling is that of Keane and Wolpin (1997), but other important papers include Stinebrickner (2000, 2001a, 2001b).

Despite the availability of high quality household data, there have been relatively few analyses of occupational destination in Brazil. In an early study that uses census data, Arriagada and Ziderman (1992) investigate the extent to which vocational education raises earnings. They find that, where there is a good match between the nature of the vocational education and the characteristics of the occupation in which a worker is employed, the rate of return to education is high, with a Mincerian rate of return of around 22 per cent. This does not differ significantly from the rate of return to academic education.

Behrman *et al* (1996) show that the impact of schooling quality as well as quantity is important in labour market outcomes. Specifically, using data from the 1980 census, their findings suggest that there is a significant independent impact ofschooling quality on wages in Brazil that operates over and above the effect of schooling quantity.
With the aim of understanding the effects of economic shocks on employment and schooling, Duryea, Lam & Levison (2007) analyse the relationship between household economic shocks and child employment in Brazil's six largest cities. Brazil has had relatively high levels of child employment, especially considering the country's relatively high *per capita* income. The authors used the Monthly Employment Survey (PME) from 1982 to 1999, to estimate probit regressions. The regressions indicate that an unemployment shock increases the probability with which a child enters the labour force, drops out of school and fails to advance in school.

[3] http://individual.utoronto.ca/vaguirre/wpapers/program_code_survey_joe_2008.html
[4] Another useful recent survey is provided by Keane and Wolpin (2009).

Curi and Menezes-Filho (2007) have examined the relation between school performance and wages of young Brazilians. After correcting for selection bias problems caused by migration and by the high educational level of the sample, their most important results indicate that the average test scores of a generation has a significant impact on its wages 5 years later, with a positive elasticity of 0.3.

In a more recent study, Curi and Menezes-Filho (2008) investigated the effect of stature - viewed as a proxy for socioeconomic, demographic and health conditions - on both wages and education in Brazil. They find that human capital deficiencies in infancy have very important effects over the life cycle so that public investments in health, education, housing and nutrition early on have high returns. Specifically, they examine the relationship between height and school cycles, on occupation allocation and on the earnings of individuals when in the labour market, separately for men and women. They find that height has a positive effect on the probability of completion of the schooling cycles for men and women and on labour market earnings, independently of its effects on occupation and on education. The occupations that require greater ability attract individuals that are, on average, taller.

Madeira *et al*. (2010) have used a difference-in-difference methodology to evaluate the change in probability of older workers (of pensionable age) being engaged in the labour market as employers. They find that the impact of a 2003 law that, for the first time, allowed commercial banks to offer social security recipients loans whose repayments can be made as direct deductions from their salaries was to raise the probability with which such individuals engage in the labour market. Specifically, since such loans are often used for business purposes, the authors find that the probability with which such individuals engage in entrepreneurial activities – and so are classified as ‘employers’ rose by almost 1 per cent in response to the change in the law.

Arguably the most relevant study in the present context, albeit one that uses a somewhat different methodology, is that of Ferreira and Leite (2002). These authors conduct an analysis of the impact of educational expansion on the incidence of poverty in the state of Ceará, using data from the 1999 round of the Pesquisa Nacional por Amostra de Domicilios (PNAD). Their model involves separate estimation of a number of separate ‘blocks’, each of which explains an aspect of individual behaviour such as occupational choice and education choice. They then use the estimates from these models to simulate the impact of policy change over time. Our model differs from that of Ferreira and Leite in that we model all decisions within a single, dynamic, framework, and consider the impact of policy changes within this framework. But certain aspects of Ferreira and Leite’s work – especially their use of multinomial logit as a means of modelling choice – are similar to the approach we take below, and so comparisons between our work and theirs are particularly instructive.

4. Data

The two types of analysis conducted in this paper call on the use of two distinct datasets. For the static multinomial logit analysis, we employ the standard large scale Brazilian household survey, namely the Pesquisa Nacional por Amostra de Domicilios (PNAD) for the years of 1993, 1999 and 2005. This dataset has been widely used in the literature; see, for example, Arbache *et al*. (2004) and Ribas and Machado (2007). It contains information concerning, *inter alia,* work experience, education and other personal characteristics. We choose to analyse the data at six year intervals over a period running from the early 1990s through the mid-2000s, this period corresponding with a rapid rise in educational participation, as we could see in Table 2)

For the longitudinal analysis, we use data from the RAIS-MIGRA data set over the period 1995-2006.[5] This is a large longitudinal administrative data set which takes the form of an annual census of all formal sector workers. In view of the large size of this data set, and of the computer intensive nature of the estimation procedure being used, we have taken a random sample of 2509 male workers, all of whom pass through the school leaving age of 14 at some point during the 1995-2006 window.

It should be noted that the RAIS-MIGRA data provide information only for years in which the worker is employed in the formal sector.[6] A little over a half of all employment in Brazil is in this sector (Hoek, 2007). However, it is possible to infer activity in some other periods from the data that are provided. In particular, we know from RAIS-MIGRA the individual's highest level of education and so (on the assumption that education is uninterrupted) we know the individual's age when he leaves education. We therefore know that he is in school at all ages younger than this. Beyond this age, if he is not observed in RAIS-MIGRA, he must belong to the 'other activity' category (which may include employment in the informal sector or a state of not being employed). In this way, we can construct a complete, balanced, panel of data for our sample.

It is worth noting explicitly that the way in which these data are constructed inevitably leads to some measure of selection bias. Since we have data only for workers who, at some stage, have been employed in the formal sector, the data do not represent a random sample of the Brazilian population. While it would be possible to obtain such a random sample from other surveys (such as PNAD), these other surveys do not have the longitudinal properties needed in order to carry out the research attempted here.

Wage data are available for periods when a worker is employed in the formal sector, and these are used in the estimation as a means of identifying the coefficients. Using the consumer price index, these wage data have been deflated to 2005 values. There is a small number of observations where, while the respondent is known to be in formal sector work, wage data are absent. In these cases, the occupation specific average of the real wage is used.

The choice of education policy variable (educpol) presented something of a challenge in that consistent time series for many of the conventional measures (such as public expenditure on education as a percentage of GDP) are not readily available for all years in our study. Commonly used sources of data such as the World Development Indicators have gaps for certain years. We have therefore used the gross enrolment rate in tertiary education for the relevant age group (18-22), calculated from PNAD data (and, for 2003 – when there was no PNAD – from the census). The series for this variable is reported in Figure 1 and shows a marked increase in the enrolment rate over time. Indeed the enrolment rate has more than doubled over the course of little more than a decade.

[5] RAIS-MIGRA is a data set produced by the Ministry of Labour and Employment (Ministério do Trabalho e Emprego). RAIS (Relação Anual de Informações Sociais) is the annual social information data set; RAIS-MIGRA refers to an extension of this data set to enable workers to be tracked through time, primarily for the purpose of analysing migration.

[6] The formal labour market is subject to a plethora of regulations, in particular covering issues such as severance payments and the requirement to provide notice of termination of contract.

The sample that we use comprises men aged 15-35 inclusive for the years 1993, 1999 and 2005. The six possible outcomes for our dependent variable (*y*) are agriculture (*status=1*), other manual (*status=2*), non-manual (*status=3*), self-employed outside agriculture (*status=4*), in education (*status=5*), and not in work or education (*status=6*)[7].

The number of observations in the various occupation and activity categories differs. According to Table 3, the most common activities are other manual activities, non-manual, and not in work or education. Analyzing the change from 1993 to 2005, it is clear that both non-manual and education categories increased in importance through the period. On the other hand, about 11 percent of the individuals are registered as agriculture. This percentage declined from 1993 to 2005. The proportion of people that are not in work or education also declined over these years.

We summarize the explanatory variables considered in this study and present brief descriptions and basic statistics of them in Table 2 for 1993, Table 3 for 1999 and in Table 4 for 2005. The explanatory variables can broadly be grouped into personal characteristics (age, age squared, number of years of study, ethnicity dummies – race), family and household characteristics (number of children younger than age 15, number of males working, number of females working, number of males older than 60, number of females older than 60), and regional dummies (27 Brazilian states).

The ethnicity dummies are separated into: (1) aboriginal Brazilians, (2) white, (3) African Brazilians, (4) Asian Brazilians and (5) pardo Brazilians. In Brazil, there are 27 federal units, comprising 26 states and one federal district. The data are also separated by: number of boys and girls in the household; males and females of working age; and males and females older than 60 years of age. The set of explanatory variables also includes household composition: number of boys (aged under 15); number of girls (aged under 15); males of working age; females of working age; males older than 60; females older than 60.
For the years of 1993, 1999 and 2005 pooled we have a total of 399.153 observations. There are about 51% of female individuals. In general, the medium age is around 24 years old. Regarding the number of years of study, the average is around 8.2 years. About the race, the

[7] "In education category" includes not only those in full-time education, but also those who are working part-time in manual activities while being in education. It is important to highlight that the occupational classifications adopted by the Brazilian Institute of Geography and Statistics (IBGE) in the Brazilian Classification of Occupations (CBO) have changed over the past two decades, in order to be closer to the standards established by the International Labour Organisation, while other occupations have emerged or have lost relevance in the work force. Hence, we developed a means of reconciling these data to a broad categorisation between manual and non-manual. Our reference was the International Standard Classification of Occupations (ISCO-88), using the one digit classification: (1) Legislators, senior officials and managers; (2) Professionals (e.g. physical, mathematical and engineering science professionals); (3) Technicians and associate professionals; (4) Clerks; (5) Service workers, shop and market sales workers; (6) Skilled agricultural and fishery workers; (7) Craft and related trades workers; (8) Plant and machine operators and assemblers; (9) Elementary occupations.

The five first categories – from (1) to (5) – are non-manual activities, while the last four categories – from (6) to (9) are manual activities. When the descriptions in our data have a different code in the IBGE 91classification and in the ISCO classification, we adopted the criteria of the most prevalent category. This is the case to the years of 1993 and 1999. For instance, if a particular code is prevalent on manual activities, we will consider all the individuals as manual workers, and vice-versa. Our sample for 2005 has occupations based on the CBO, which is follows the *ISCO-88*.

observations are divided into White and Brown: 0.16% is Indian, 51.31% are white, 5.78% are Black, 0.42% is Yellow and 42.33% are Brown. The most part of observations are located in these three states: 22.34% in São Paulo, 10.63% in Minas Gerais and 8.01% in Bahia.

5. Statistical modelling

We report, first, the results of the static multinomial logit modelling exercise; results obtained using the dynamic discrete choice model follow later.

(i) Multinomial logit models

We consider six labour market outcomes: (i) agricultural work; (ii) other manual employment; (iii) non-manual employment; (iv) self-employment outside agriculture; (v) in education; and (vi) not in work or education. Explanatory variables are: years of schooling, age, age squared, ethnicity, a full set of region dummies, and a variety of household composition variables. The latter comprise counts of: working age males; working age females; males older than 60; females older than 60; boys; and girls in the household. The standard regions[8] are: Rondônia; Acre; Amazonas; Roraima; Pará; Amapá; Tocantins; Maranhão; Piauí; Ceará; Rio Grande do Norte; Paraíba; Pernambuco; Alagoas; Sergipe; Bahia; Minas Gerais; Espirito Santo; Rio de Janeiro; Paraná; Santa Catarina; Rio Grande do Sul; Mato Grosso do Sul; Mato Grosso; Goiás; and Distrito Federal.

In common with Ferreira and Leite (2002), we model occupational choice as a reduced form, choosing not to include an earnings variable as a determinant of choice, but rather including characteristics typical of those found in Mincerian earnings functions as measures of earnings potential. The adoption of a reduced form approach allows us to finesse issues of endogeneity and sample selection bias.

In Tables 5-7, we report the marginal effects on the years of schooling variable, separately for each year, and separately for males, females, and all respondents. It is clear that in all years, schooling raises the probability with which an individual enters non-manual work, and reduces the probability with which an individual enters manual work. Schooling also raises the probability of continuing in education. For women, in most years, schooling raises the probability of entering self-employment outside of agriculture. This could conceivably reflect gender discrimination; if highly educated women find that their opportunities as employees are limited, they may decide to set up their own businesses.

For reasons of space, we do not report the marginal effects on the other variables in full; we do, however, report the results, pooled across men and women, for each year separately in the appendix. It is readily observed that, almost without exception, these marginal effects are highly significant, and that they affect outcomes in the expected direction.

In Table 8 we report the results of an analysis in which data from all three rounds are pooled, but the schooling variable is interacted with a round index so that we can investigate how the impact of schooling has changed over time.[9]

[8] See Figure A.1 with the Brazilian States in the appendix.
[9] The schooling interaction to the year of 1993 is the omitted dummy.

On one hand, we find that the impact of schooling is increasing the probability that the respondent will be in education over time (as expected), as well as (more surprisingly) in the other manual employment and in agricultural employment categories.[10] On the other hand, these effects of schooling on occupational outcomes are diminishing over time for non-manual employment and for self-employed.
The results reported above make clear that an increased incidence of education raises the probability with which individuals remain in education (unsurprisingly), and the probability with which they enter employment as non-manual workers. It is clear therefore that national investment in education has a direct impact on occupational outcomes, leading to more workers entering non-manual jobs. We investigate this further as we turn to consider the dynamic modelling of destination.

(ii) Dynamic discrete choice model

In this section we evaluate the dynamic model, taking seriously the starting point provided by Keane and Wolpin. We thus begin with the following instantaneous reward functions:

$$
\begin{aligned}
R_{1t} &= w_1 = \alpha_{10}+\alpha_{11}s_t+\alpha_{12}x_{1t}+\alpha_{13}\,x_{2t}+\varepsilon_{1t} \\
R_{2t} &= w_2 = \alpha_{20}+\alpha_{21}s_t+\alpha_{22}x_{1t}+\alpha_{23}\,x_{2t}+\varepsilon_{2t} \\
R_{3t} &= \beta_0+\beta_1 I(s_t \geq 12)+\beta_2 \mathrm{educpol}+\varepsilon_{3t} \\
R_{4t} &= \gamma_0+\varepsilon_{4t}
\end{aligned} \tag{2}
$$

Here s refers to years of schooling received prior to the current period t, x_1 is years of experience in occupation 1, and x_2 is years of experience in occupation 2. The terms R_1 through R_4 denote respectively the instantaneous returns to working in occupation 1 (non-manual occupations), occupation 2 (manual occupations), or schooling, or other activity (which may include other work, unemployment, or absence from the labour force). In the case of the first two outcomes, we observe the wages, w_1 and w_2 respectively, and these are incorporated into the modelling procedure. The ε terms represent alternative-specific, period-specific, random shocks. These are crucial in determining why some workers take certain paths through their career while others take others. The first term in the instantaneous reward for schooling equation indicates that we expect the one-period 'reward' associated with schooling at tertiary level, β_1, to be negative owing to the payment of tuition fees. The second term in that equation is intended to capture the effect of education policy (educpol) on the decision to stay on at school, and the sign and magnitude of the coefficient attached to that variable, β_2, is therefore of primary interest in the present study.

As with any approximation method, a number of parameters need to be set by the analyst in order to proceed. For the simulation used to evaluate the regime that yields the greatest expected future return, we use 500 draws; we evaluate the expected return at 300 randomly chosen points in the state space and use the interpolation method for all other points. The discount parameter is set at 0.95. The convergence toward the maximum likelihood solution is deemed to be complete when further iterations fail to achieve an improvement in the log likelihood that exceeds 0.001%.

Parameter estimates are reported in Table 9, and are broadly in line with our prior expectations. The key finding is that educpol raises the propensity of respondents to stay in

[10] The impact of agricultural employment is not significant in 1999.

education. Moreover, educational attainment increases the propensity to be in formal sector work relative to other destinations – though surprisingly it has a greater effect on entry to manual as opposed to non-manual work in the formal sector. The high value of the ρ_{33} parameter indicates that there is a considerable amount of unobserved heterogeneity across individuals, and that this impacts on the returns that are available to education; it may be the case that this could be modelled by separately evaluating coefficients for respondents that come from different family backgrounds, but this is an exercise that we leave for further work.

Following Keane and Wolpin (1994, 1997) we evaluate standard errors using the outer product of numerical first derivatives. Keane and Wolpin note that there may be a downward bias associated with these standard errors. The high t statistics reported in Table 1 for most of the coefficients seem to be quite typical for this type of model. Moreover, we note that the educpol variable is clustered across all observations in a given year. We are not aware of any literature that allows correction for such clustering in this context, but note that this too will likely bias the standard error downwards. Hence our central result concerning the impact of educational policy needs to be interpreted with some measure of caution.

It is possible to use the estimates reported in Table 9 as a starting point in an exercise which aims to evaluate how future changes in educational policy are likely to affect occupational outcomes. The software provided by Keane and Wolpin includes a program that, given the estimated parameter values, enables us to compute the within period probabilities with which a randomly selected observation is expected to appear in each regime in each period of the time frame under consideration; we can thus calculate these probabilities for an assumed time series of the educational policy variable. This is, once again, a rather computationally intensive exercise: for each individual in each period it is necessary to evaluate the expected returns at each point in a large state space. We do so using Keane and Wolpin's default values. Raising the educational policy variable from 5% to 15% has the effect of raising the unconditional mean value of years spent in non-manual formal sector work from 0.0570 to 0.0717. The value of these means is small (since many individuals in the sample are of an age still to be in compulsory education), but the direction of change is very much in line with intuition.

The result is subject to a number of caveats. In particular, the 'not in schooling or formal sector employment' category is broad; were it possible to disaggregate this category, we might conceivably find that expanding educational opportunity results in substantially lower levels of non-employment. A number of studies that employ dynamic discrete choice methods divide the population into subgroups (based, for example, on family income). To preserve simplicity, we have not done this, and it may be the case that a more refined specification of the model could identify stronger policy effects. Finally, it should be remembered that the way in which the RAIS-MIGRA data are collected inevitably result in some selection bias, since only workers that are at some stage employed in the formal sector are included in the data.

6. Conclusions

An increase in spending on education leads, not surprisingly, to an increase in the propensity for young people to undertake education. Later in the life cycle, the human capital that they have acquired equips these young people to undertake jobs that are qualitatively different from those in which they would otherwise have become employed. Put simply, more people get better jobs. This should be expected to tilt the economy’s comparative advantage toward the production of goods and services that are more skill intensive and hence more remunerative.

Our results are plausible, but should be treated with a measure of caution. In particular, it should be noted that the approach taken in the dynamic modelling assumes that, at the outset of their working lives, individuals differ only in the random shocks that they encounter. It may well be the case that different types of individual can be identified, and that improvements to the model fit can be secured by modelling these types in a distinct fashion. This is left for further work.

7. References

Aguirregabiria, Victor and Pedro Mira (2010) Dynamic discrete choice structural models: a survey, *Journal of Econometrics*, 156, 38-67.

Arbache, Jorge Saba, Dickerson, Andy and Green, Francis (2004) Trade liberalisation and wages in developing countries, *Economic Journal*, 114, F73-F96.

Arriagada, Ana-Maria and Ziderman, Adrian (1992) Vocational secondary schooling, occupational choice and earnings in Brazil, World Bank Working Paper Series 1037, http://siteresources.worldbank.org/BRAZILINPOREXTN/Resources/3817166-1185895645304/4044168-1186326902607/45pub_br65.pdf

Behrman, Jere R., Birdsall, Nancy and Kaplan, Robert (1996) The quality of schooling and labor market outcomes. in Birdsall, Nancy and Sabot, Richard H. (eds) Opportunity Foregone: education in Brazil. Washington, Inter-American Development Bank, Chapter 8.

Boskin, Michael (1974) A conditional logit model of occupational choice, *Journal of Political Economy*, 82, 389-398.

Curi and Menezes-Filho (2007) The relationship between school performance and future wages in Brazil, Brazilian Meeting of ANPEC.

Curi and Menezes-Filho (2008) A relação entre altura, escolaridade, ocupação e salários no Brasil, Brazilian Meeting of ANPEC.

Duflo, Esther (2004) The medium run consequences of educational expansion: evidence from a large school construction program in Indonesia, *Journal of Development Economics*, 74, 163-197.

Dureya, Suzanne, Lam, David and Levison, Deborah (2007) Effects of economic shocks on children's employment and schooling in Brazil, *Journal of Development Economics*, 84, 188–214.

EDUDATABRASIL - Sistema de Estatísticas Educacionais. **Matrículas->Educação de Jovens e Adultos->Perfil do estabelecimento.** Available at: http://www.edudatabrasil.inep.gov.br/. Access in November 1st, 2010.

Ferreira, Francisco H.G. and Leite, Phillippe George (2004) Educational expansion and income distribution: a microsimulation for Ceara, in Shorrocks, Anthony and van der Hoeven, Rolph (eds) Growth, inequality and poverty: prospects for pro-poor economic development, Oxford: Oxford University Press, http://www.econ.puc-rio.br/pdf/td456.pdf

Hoek, Jasper (2007) Labor flows in formal and informal markets in Brazil, IZA/World Bank Conference on Employment and Development, Bonn, 8 June, http://www.iza.org/conference_files/worldb2007/hoek_j1567.pdf

Johnes, Geraint (1999) Schooling, fertility and the labour market experience of married women, *Applied Economics*, 31, 585-592.

Keane, Michael P. (2010) Structural v atheoretic approaches to econometrics, *Journal of Econometrics*, 156, 3-20.

Keane, Michael P. and Wolpin, Kenneth I. (1994) The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence, *Review of Economics and Statistics*, 76, 648-672.

Keane, Michael P. and Wolpin, Kenneth I. (1997) The career decisions of young men, *Journal of Political Economy*, 105, 473-522.

Keane, Michael P. and Wolpin, Kenneth I. (2009) Empirical application of discrete choice dynamic programming models, *Review of Economic Dynamics*, 12, 1-22.

Madeira, Gabriel, Rangel, Marcos A. And Rodrigues, Mauro (2010) Occupational choice and commitment power: inferential evidence from changes in the availability of credit contracts, http://mitsloan.mit.edu/neudc/papers/paper_361.pdf

Makepeace, Gerald H. (1996) Lifetime earnings and the training decisions of young men in Britain, *Applied Economics*, 28, 725-735.

Nerlove, Marc and Press, S. James (1973) Univariate and multivariate log-linear and logistic models, R-1306-EDA/NIH, Santa Monica: Rand, available at http://www.rand.org/pubs/reports/2006/R1306.pdf.

OEI – Organización de Estados Iberoamericanos para la Educación. **Sistema Educativo Nacional de Brasil.** Available at: http://www.oei.es/quipu/brasil/estructura.pdf. Ribas, Rafael Perez and Machado, Ana Flavia Machado (2007) Distinguishing chronic poverty from transient poverty in Brazil: developing a model for pseudo-panel data, http://nip-lac.org/uploads/Rafael_Perez_Ribas.pdf

Roy, Andrew Donald (1951) Some thoughts on the distribution of earnings, *Oxford Economic Papers*, 3, 135-146.

Rust, John (1987) Optimal replacement of GMC bus engines: an empirical model of Harold Zurcher, *Econometrica*, 55, 999-1035.

Schmidt, P and Strauss, R.P. (1975) The prediction of occupation using multiple logit models, *International Economic Review*, 16, 471-486.

Soopramanien, Didier and Johnes, Geraint (2001) A new look at gender effects in participation and occupation choice, *Labour*, 15, 415-443.

Stinebrickner, Todd R. (2000) Serially correlated variables in dynamic discrete choice models, *Journal of Applied Econometrics*, 15, 595-624.

Stinebrickner, Todd R. (2001a) Compensation policies and teacher decisions, *International Economic Review*, 42, 751-779.

Stinebrickner, Todd R. (2001b) A dynamic model of teacher labor supply, *Journal of Labor Economics*, 19, 196-230.

Willis, Robert J. and Rosen, Sherwin (1979) Education and self-selection, *Journal of Political Economy*, 87, S7-S36.

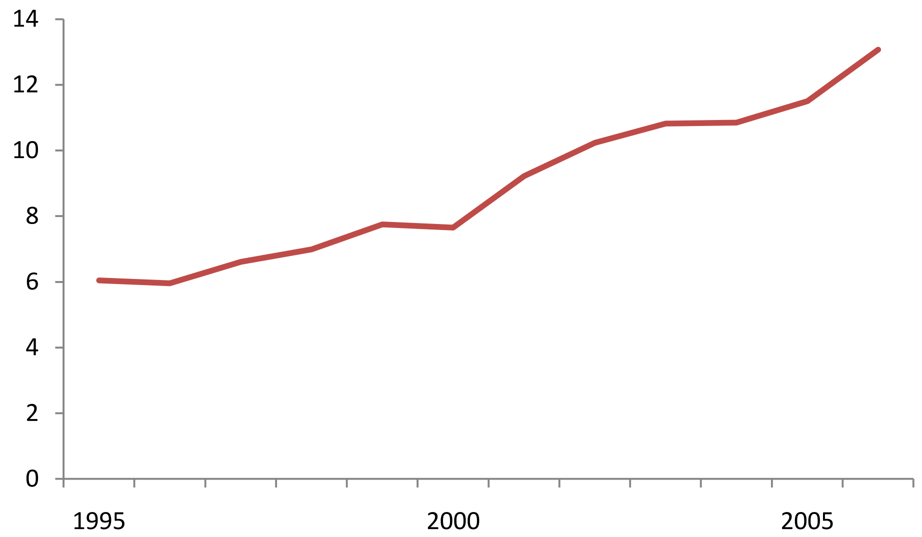

Figure 1: Percentage gross enrolment rate of 18-22 year olds in tertiary education
Source: Census (2000); PNAD (all other years).

Table 1: Structure of education in Brazil

Level	**Ages**	**Number of Years**
Kindergarten	0 – 3 (day care) 4 – 6 inclusive	3 2
Elementary education	7 to 14 inclusive	8*
Secondary education	From 15 years old	3 to 4
Higher education	Variable.	3 to 6

Source: adapted from OEI (2010).
Note: *A recent change, effective from 2010, raises the duration of basic education to 9 years.

Table 2: Registration/enrollment in Brazil

Year	Kindergarten - day care center	Kindergarten - pre-school	Basic education	Secondary education	Higher education	Special education	Youth and adult education
1999	831,978	4,235,278	36,059,742	7,769,199	not available	165,325	3,071,906
2000	916,864	4,421,332	35,717,948	8,192,948	2,694,245	30,052	1,693,786
2001	1,093,347	4,818,803	35,298,089	8,398,008	3,030,754	323,399	3,777,989
2002	1,152,511	4,977,847	35,150,362	8,710,584	3,479,913	337,897	3,779,593
2003	1,237,558	5,155,676	34,438,749	9,072,942	3,887,022	358,898	4,403,436
2004	1,348,237	5,555,525	34,012,434	9,169,357	4,163,733	371,383	4,577,268
2005	1,414,343	5,790,670	33,534,561	9,031,302	not available	331,814	4,619,409
2006	1,427,942	5,588,153	33,282,663	8,906,820	not available	326,994	5,296,050

Source: adapted from EDUDATABRASIL - Sistema de Estatísticas Educacionais.

Table 3: Occupation Categories by Age Group

Y	1993			1999			2005		
	15-22	23-35	Total	15-22	23-35	Total	15-22	23-35	Total
Agriculture	7,378	7,265	14,643	6,371	6,756	13,127	6,173.38	7,432.71	13,606.10
	14.99	11.00	12.70	11.30	9.57	10.34	9.86	8.56	9.11
Other manual	9,418	14,101	23,519	8,894	15,285	24,179	9,825.99	20,802.62	30,628.61
	19.13	21.35	20.40	15.77	21.66	19.05	15.69	23.97	20.50
Non-manual	7,914.66	16,791.70	24,706.36	8,624.29	18,149.10	26,773.40	11,494.15	25,801.48	37,295.63
	16.08	25.43	21.43	15.29	25.72	21.09	18.35	29.73	24.96
Self-employed	2,476.13	9,890.46	12,366.59	2,628.99	10,550.58	13,179.57	2,654.33	11,135.10	13,789.43
outside agriculture	5.03	14.98	10.73	4.66	14.95	10.38	4.24	12.83	9.23
In education	12,500.33	1,149.29	13,649.61	19,664.96	2,135.62	21,800.58	21,420.54	2,922.15	24,342.69
	25.39	1.74	11.84	34.87	3.03	17.17	34.20	3.37	16.29
Not in work or	9,546.73	16,842.11	26,388.83	10,206.52	17,682.33	27,888.85	11,056.77	18,699.79	29,756.55
education	19.39	25.50	22.89	18.10	25.06	21.97	17.66	21.55	19.91
Total	49,234.14	66,039.86	115,274	56,389.68	70,558.32	126,948	62,625.15	86,793.85	149,419
	100	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**

Source: PNAD.

Table 4: Summary statistics - Pooled

Variable	Description	Total Obs.	Mean	Std. Dev.	Min.	Max.
Occupational choice (y)	y includes 6 categories: (a) agriculture, (b) other manual, (c) non-manual, (d) self-employed outside agriculture, (e) in education and (f) not in work or education.	391,641	3.6380	1.6907	1	6
Age	Age	399,153	24.3097	6.0633	15	35
Agesqr	Age squared	399,153	627.7224	301.8335	225	1225
Schooling	Number of years of study	396,142	8.1978	3.8710	1	16
Female		399,153	0.5074	0.4999	0	1
Race						
Race1	Aboriginal Brazilians	399,106	0.0016	0.0394	0	1
Race2	White	399,106	0.5131	0.4998	0	1
Race3	African Brazilians	399,106	0.0578	0.2335	0	1
Race4	Asian Brazilians	399,106	0.0042	0.0646	0	1
Race5	Pardo Brazilians	399,106	0.4233	0.4941	0	1
State - Federal Unit						
State1	Rondônia	399,153	0.0068	0.0822	0	1
State2	Acre	399,153	0.0027	0.0519	0	1
State3	Amazonas	399,153	0.0149	0.1210	0	1
State4	Roraima	399,153	0.0017	0.0407	0	1
State5	Pará	399,153	0.0276	0.1637	0	1
State6	Amapá	399,153	0.0028	0.0527	0	1
State7	Tocantins	399,153	0.0069	0.0830	0	1
State8	Maranhão	399,153	0.0329	0.1785	0	1
State9	Piauí	399,153	0.0171	0.1295	0	1
State10	Ceará	399,153	0.0431	0.2032	0	1

State11	Rio Grande do Norte	399,153	0.0168	0.1286	0	1
State12	Paraíba	399,153	0.0207	0.1424	0	1
State13	Pernambuco	399,153	0.0476	0.2130	0	1
State14	Alagoas	399,153	0.0170	0.1294	0	1
State15	Sergipe	399,153	0.0111	0.1046	0	1
State16	Bahia	399,153	0.0801	0.2714	0	1
State17	Minas Gerais	399,153	0.1063	0.3082	0	1
State18	Espírito Santo	399,153	0.0191	0.1369	0	1
State19	Rio de Janeiro	399,153	0.0808	0.2725	0	1
State20	São Paulo	399,153	0.2234	0.4165	0	1
State21	Paraná	399,153	0.0576	0.2330	0	1
State22	Santa Catarina	399,153	0.0321	0.1761	0	1
State23	Rio Grande do Sul	399,153	0.0570	0.2319	0	1
State24	Mato Grosso do Sul	399,153	0.0128	0.1124	0	1
State25	Mato Grosso	399,153	0.0161	0.1257	0	1
State26	Goiás	399,153	0.0311	0.1737	0	1
State27	Distrito Federal	399,153	0.0139	0.1172	0	1
Boy	Younger than 15 years old	399,153	0.5938	0.8376	0	8
Girl	Younger than 15 years old	399,153	0.5814	0.8285	0	9
Males of working age		399,153	1.4322	0.9859	0	9
Females of working age		399,153	1.4175	0.8387	0	9
Males older than 60		399,153	0.0735	0.2635	0	3
Female older than 60		399,153	0.0761	0.2715	0	4

Source: PNAD/IBGE

Table 5: Multinomial logit marginal effects of years of schooling, men and women aged 15-35

agricultural employment or self-employment	other manual employment	non-manual employment	self-employed outside agriculture	in education	not in work or education
		1993			
-0.0239	-0.0231	0.0532	-0.0020	0.0110	-0.0152
(-72.81)	(-45.09)	(106.48)	(-5.89)	(44.00)	(-29.50)
n=115,253		LL=-70,826,214		Pseudo R^2=0.2248	
		1999			
-0.0181	-0.0199	0.0472	-0.0038	0.0152	-0.0206
(-62.72)	(-42.13)	(100.30)	(-13.07)	(49.40)	(-41.32)
n=126,930		LL=-77,647,869		Pseudo R^2=0.2205	
		2005			
-0.0149	-0.0183	0.0517	-0.0059	0.0102	-0.0228
(-59.56)	(-41.65)	(104.41)	(-22.55)	(29.17)	(-51.18)
n=149,412		LL=-91,335,137		Pseudo R^2=0.2077	

Note: z values in parentheses; control variables are as described in the text.

Table 6: Multinomial logit marginal effects of years of schooling, men aged 15-35

agricultural employment or self-employment	other manual employment	non-manual employment	self-employed outside agriculture	in education	not in work or education
		1993			
-0.0313	-0.0160	0.0506	-0.0046	0.0064	-0.0051
(-56.24)	(-21.32)	(74.06)	(-8.00)	(21.02)	(-9.32)
n=55,357		LL=-77,506.85		Pseudo R^2=0.1874	
		1999			
-0.0248	-0.0133	0.0436	-0.0074	0.0112	-0.0093
(-50.85)	(-18.87)	(69.65)	(-14.48)	(30.27)	(-16.32)
n=61,054		LL=-39,201,715		Pseudo R^2=0.1917	
		2005			
-0.0218	-0.0137	0.0485	-0.0093	0.0089	-0.0125
(-51.85)	(-21.16)	(77.05)	(-22.28)	(25.92)	(-25.16)
n=73,196		LL=-102,187.32		Pseudo R^2=0.1916	

Note: z values in parentheses; control variables are as described in the text.

Table 7: Multinomial logit marginal effects of years of schooling, women aged 15-35

agricultural employment or self-employment	other manual employment	non-manual employment	self-employed outside agriculture	in education	not in work or education
		1993			
-0.0143	-0.0235	0.0497	0.0019	0.0138	-0.0276
(-40.81)	(-39.79)	(73.35)	(5.50)	(36.66)	(-34.27)
n=59,896		LL=-34,645,620		Pseudo R^2=0.2142	
		1999			
-0.0107	-0.0219	0.0477	0.001	0.017	-0.0332
(-34.15)	(-39.92)	(72.20)	(3.27)	(34.17)	(-42.60)
n=65876		LL=-37751791		Pseudo R^2=0.2138	
		2005			
-0.0078	-0.0188	0.0526	-0.0004	0.009	-0.0345
(-28.47)	(-36.40)	(71.56)	(-1.35)	(14.99)	(-47.55)
n=76,216		LL=-44,759,581		Pseudo R^2=0.1917	

Note: z values in parentheses; control variables are as described in the text.

Table 8: Impact of schooling over time – marginal effects on year x schooling interaction in pooled specification, men and women aged 15-35

year	agricultural employment or self-employment	other manual employment	non-manual employment	self-employed outside agriculture	in education	not in work or education
1999	0.0002	0.0012	-0.0038	-0.0004	0.0019	0.0009
	(1.03)	(3.89)	(-16.80)	(-2.40)	(15.62)	(3.23)
2005	0.0024	0.0055	-0.0056	-0.0020	0.0004	-0.0007
	(13.42)	(19.76)	(-26.58)	(-11.24)	(3.48)	(-2.48)
	n=388,686		LL=-2.385e+08		Pseudo R^2=0.2179	

Note: z values in parentheses; control variables are as described in the text.

Table 9: Dynamic discrete choice model: parameter estimates

variable	estimated coefficient	t statistic
α_{10}	5.0937	21120.63
α_{11}	0.0626	526.67
α_{12}	0.3483	704.22
α_{13}	0.1058	134.85
α_{20}	4.2737	1158.51
α_{21}	0.0715	420.11
α_{22}	0.4947	493.27
α_{23}	-0.0195	27.30
$\beta_0/1000$	3.2795	10.54
$\beta_1/1000$	-4.3985	13.67
$\beta_2/1000$	0.0500	2.61
$\gamma_0/1000$	-0.9451	10.48
ρ_{11}	0.0708	2001.34
ρ_{22}	0.1080	1077.96
ρ_{33}	4.8188	27.42
ρ_{44}	2.6870	26.82
Log likelihood	-360923.86	

Note: The ρ terms are the correlations of the error terms such that:
$\varepsilon_{1t} = \rho_{11}\eta_{1t}$
$\varepsilon_{2t} = \rho_{22}\eta_{2t}$
$\varepsilon_{3t} = \rho_{33}\eta_{3t}$
$\varepsilon_{4t} = \rho_{44}\eta_{4t}$
η_{kt} ~N(0,1), k=1,...,4.

APPENDIX

Table A1: Multinomial logit marginal effects, men and women aged 15-35, full results for 2005

	agricultural employment or self-employment	other manual employment	non-manual employment	self-employed outside agriculture	in education	not in work or education
schooling (years)	-0.0149	-0.0183	0.0517	-0.0059	0.0102	-0.0228
	(0.000250)	(0.000440)	(0.000495)	(0.000261)	(0.000349)	(0.000446)
age	-0.00170	0.0453	0.00810	0.0268	-0.114	0.0356
	(0.00103)	(0.00249)	(0.00261)	(0.00165)	(0.00179)	(0.00238)
age^2	3.89e-08	-0.000746	-2.86e-05	-0.000359	0.00190	-0.000769
	(2.06e-05)	(4.87e-05)	(5.07e-05)	(3.16e-05)	(3.57e-05)	(4.69e-05)
female	-0.0320	-0.191	-0.0207	-0.0804	0.0326	0.291
	(0.00153)	(0.00324)	(0.00319)	(0.00206)	(0.00194)	(0.00316)
Aboriginal Brazilians	-0.00447	-0.0584	0.0123	0.0486	0.0143	-0.0123
	(0.0112)	(0.0276)	(0.0328)	(0.0237)	(0.0197)	(0.0272)
African Brazilians	-0.00419	-0.0466	0.0425	0.00630	0.0154	-0.0134
	(0.00149)	(0.00335)	(0.00325)	(0.00205)	(0.00183)	(0.00318)
Asian Brazilians	-0.0211	0.0284	-0.00173	-0.0170	0.00547	0.00598
	(0.00207)	(0.00602)	(0.00600)	(0.00327)	(0.00354)	(0.00574)
Pardo Brazilians	0.0126	-0.143	0.0501	0.0486	0.0911	-0.0596
	(0.0201)	(0.0178)	(0.0243)	(0.0216)	(0.0211)	(0.0234)
Rondônia	0.177	-0.104	-0.0432	0.0347	-0.000912	-0.0638
	(0.0146)	(0.00798)	(0.0101)	(0.00875)	(0.00668)	(0.00862)
Acre	0.0866	-0.0981	-0.0355	0.0488	0.0460	-0.0478
	(0.0117)	(0.0100)	(0.0127)	(0.0107)	(0.0114)	(0.0109)
Amazonas	0.0167	-0.0861	-0.0409	0.0579	0.0619	-0.00963
	(0.00782)	(0.00723)	(0.00819)	(0.00801)	(0.00789)	(0.00844)
Roraima	0.117	-0.103	-0.0282	0.0145	-0.000685	-0.000146
	(0.0201)	(0.0136)	(0.0158)	(0.0132)	(0.00995)	(0.0168)
Pará	0.0405	-0.0383	-0.0393	0.0622	0.0245	-0.0496
	(0.00663)	(0.00653)	(0.00662)	(0.00655)	(0.00502)	(0.00602)
Amapá	-0.0133	-0.111	0.00383	0.0571	0.0875	-0.0239
	(0.00730)	(0.0113)	(0.0148)	(0.0131)	(0.0142)	(0.0139)
Tocantins	0.190	-0.0757	-0.0443	0.0139	-0.000287	-0.0842
	(0.0139)	(0.00912)	(0.0101)	(0.00837)	(0.00661)	(0.00844)
Maranhão	0.117	-0.131	-0.0474	0.114	0.0155	-0.0672
	(0.0102)	(0.00695)	(0.00956)	(0.0102)	(0.00657)	(0.00803)

Piaui	0.168	-0.142	-0.0672	0.119	0.0256	-0.104
	(0.0131)	(0.00748)	(0.0105)	(0.0116)	(0.00814)	(0.00795)
Ceará	0.113	-0.0684	-0.0770	0.0578	0.00461	-0.0302
	(0.00845)	(0.00573)	(0.00561)	(0.00627)	(0.00404)	(0.00615)
Rio Grande do Norte	0.0561	-0.0906	-0.0482	0.0256	0.0389	0.0183
	(0.00905)	(0.00820)	(0.00968)	(0.00823)	(0.00793)	(0.0101)
Paraiba	0.0866	-0.0928	-0.0489	0.0412	0.0400	-0.0261
	(0.00922)	(0.00766)	(0.00951)	(0.00819)	(0.00750)	(0.00866)
Pernambuco	0.122	-0.121	-0.0621	0.0322	0.0298	-0.000200
	(0.00850)	(0.00497)	(0.00571)	(0.00551)	(0.00469)	(0.00641)
Alagoas	0.158	-0.146	-0.0664	0.00567	0.0503	-0.00183
	(0.0122)	(0.00711)	(0.0105)	(0.00772)	(0.00880)	(0.0102)
Sergipe	0.0712	-0.104	-0.0415	0.0465	0.0534	-0.0261
	(0.0100)	(0.00824)	(0.0104)	(0.00945)	(0.00940)	(0.00991)
Bahia	0.156	-0.124	-0.0615	0.0497	0.0250	-0.0454
	(0.00832)	(0.00446)	(0.00524)	(0.00530)	(0.00410)	(0.00528)
Minas Gerais	0.107	-0.0470	-0.0159	0.0124	-0.0106	-0.0455
	(0.00697)	(0.00527)	(0.00544)	(0.00444)	(0.00308)	(0.00524)
Espirito Santo	0.139	-0.0672	-0.0239	-0.00487	-0.00906	-0.0336
	(0.0123)	(0.00846)	(0.00929)	(0.00714)	(0.00548)	(0.00922)
Rio de Janeiro	-0.0505	-0.0803	0.0101	0.0301	0.0739	0.0166
	(0.00268)	(0.00547)	(0.00625)	(0.00542)	(0.00538)	(0.00665)
Paraná	0.0995	-0.0212	-0.0237	0.0189	-0.0261	-0.0473
	(0.00870)	(0.00670)	(0.00633)	(0.00548)	(0.00333)	(0.00637)
Santa Catarina	0.137	0.0140	-0.0228	0.00489	-0.0364	-0.0972
	(0.0119)	(0.00912)	(0.00807)	(0.00656)	(0.00376)	(0.00704)
Rio Grande do Sul	0.112	-0.0114	-0.0305	0.0169	-0.0153	-0.0715
	(0.00894)	(0.00647)	(0.00592)	(0.00515)	(0.00350)	(0.00563)
Mato Grosso do Sul	0.101	-0.0435	-0.00193	0.00621	0.00156	-0.0630
	(0.0112)	(0.00901)	(0.0104)	(0.00745)	(0.00609)	(0.00844)
Mato Grosso	0.179	-0.0858	-0.0143	-0.0130	-0.00280	-0.0629
	(0.0125)	(0.00758)	(0.00919)	(0.00611)	(0.00555)	(0.00769)
Goiás	0.0668	-0.0458	-0.00253	0.00866	-0.00112	-0.0260
	(0.00756)	(0.00646)	(0.00715)	(0.00538)	(0.00428)	(0.00661)
Distrito Federal	-0.0511	-0.0791	0.0540	-0.0111	0.0647	0.0226

	(0.00309)	(0.00716)	(0.00860)	(0.00588)	(0.00687)	(0.00891)
Number of boys in household	0.0115	0.0108	-0.0354	0.00631	-0.00586	0.0127
	(0.000765)	(0.00197)	(0.00228)	(0.00115)	(0.00121)	(0.00182)
Number of girls in household	0.0112	0.0107	-0.0300	0.00658	-0.00769	0.00920
	(0.000785)	(0.00202)	(0.00230)	(0.00117)	(0.00124)	(0.00185)
Number of working age males in household	0.0162	-0.0194	-0.0248	-0.0137	0.0109	0.0309
	(0.000689)	(0.00191)	(0.00186)	(0.00126)	(0.000921)	(0.00173)
Number of working age females in household	-0.00223	0.0213	0.0296	-0.00742	0.0182	-0.0594
	(0.000948)	(0.00220)	(0.00198)	(0.00150)	(0.00103)	(0.00229)
Number of males older than 60 in household	0.0293	-0.0426	-0.0446	-0.0107	0.0137	0.0549
	(0.00242)	(0.00635)	(0.00598)	(0.00406)	(0.00311)	(0.00592)
Number of females older than 60 in household	-0.00527	-0.0250	-0.00209	-0.0111	0.0272	0.0163
	(0.00267)	(0.00601)	(0.00547)	(0.00372)	(0.00303)	(0.00562)

Notes: Robust standard errors in parentheses; Number of observations: 149,412.

Table A2: Multinomial logit marginal effects, men and women aged 15-35, full results for 1999

	agricultural employment or self-employment	other manual employment	non-manual employment	self-employed outside agriculture	in education	not in work or education
schooling (years)	-0.0181***	-0.0199***	0.0472***	-0.0037***	0.0152***	-0.0205***
	(0.0002)	(0.0004)	(0.0004)	(0.0002)	(0.0003)	(0.0005)
age	-0.0076***	0.0321***	-0.0007	0.0336***	-0.0957***	0.0384***
	(0.0012)	(0.0026)	(0.0026)	(0.0017)	(0.0020)	(0.0026)
age^2	0.0001***	-0.0005***	0.0001**	-0.0004***	0.0015***	-0.0007***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
female	-0.0356***	-0.1664***	-0.0408***	-0.1151***	0.0231***	0.3349***

	(0.0018)	(0.0034)	(0.0032)	(0.0023)	(0.0019)	(0.0034)
Aboriginal Brazilians	-0.0203*	0.0116	0.0046	0.0560*	-0.0205	-0.0315
	(0.0116)	(0.0413)	(0.0400)	(0.0293)	(0.0169)	(0.0333)
African Brazilians	-0.0210***	0.0863***	-0.0339***	-0.0273***	-0.0081**	0.0042
	(0.0028)	(0.008)	(0.0064)	(0.0036)	(0.0035)	(0.0072)
Asian Brazilians	0.0107	-0.0877***	-0.0135	0.0536**	0.0595***	-0.0226
	(0.0276)	(0.0306)	(0.0233)	(0.0270)	(0.0207)	(0.0321)
Pardo Brazilians	-0.0056***	0.0421***	-0.0268***	-0.0138***	-0.0070***	0.0112***
	(0.0017)	(0.0037)	(0.0033)	(0.0022)	(0.0018)	(0.0036)
Rondônia	-0.0313***	-0.0506***	0.0374**	0.0402***	0.0060	-0.0016
	(0.0093)	(0.0146)	(0.0166)	(0.0149)	(0.0094)	(0.0165)
Acre	-0.0293**	-0.0907***	0.0597**	0.0259	0.0713***	-0.0369*
	(0.0131)	(0.0197)	(0.0250)	(0.0204)	(0.0201)	(0.0220)
Amazonas	-0.0358***	-0.1261***	-0.0416***	0.0609***	0.0901***	0.0524***
	(0.0060)	(0.0089)	(0.0103)	(0.0116)	(0.0108)	(0.0138)
Roraima	-0.0116	-0.0985***	0.0756**	0.0525*	0.0296	-0.0477
	(0.0229)	(0.0257)	(0.0329)	(0.0275)	(0.0215)	(0.0303)
Pará	-0.0161***	-0.0791***	0.0140	0.1002***	0.0429***	-0.0619***
	(0.0057)	(0.0075)	(0.0093)	(0.0099)	(0.0065)	(0.0079)
Amapá	-0.0275**	-0.0963***	-0.0134	0.0060	0.0823***	0.0488*
	(0.0134)	(0.0195)	(0.0230)	(0.0196)	(0.0211)	(0.0274)
Tocantins	0.1398***	-0.0642***	0.0175	0.0304***	-0.0042	-0.1192***
	(0.0147)	(0.0112)	(0.0137)	(0.0118)	(0.0077)	(0.0095)
Maranhão	0.0984***	-0.1569***	-0.0585***	0.2644***	-0.0173***	-0.1301***
	(0.0103)	(0.0069)	(0.0099)	(0.0141)	(0.0051)	(0.008)
Piaui	0.1586***	-0.1377***	-0.0269**	0.1548***	-0.0024	-0.1462***
	(0.0130)	(0.0077)	(0.0122)	(0.0134)	(0.0065)	(0.0076)
Ceará	0.1020***	-0.0915***	-0.0220***	0.0881***	0.0098**	-0.0864***
	(0.0084)	(0.0057)	(0.0068)	(0.0077)	(0.0043)	(0.0060)
Rio Grande do Norte	0.0127	-0.0679***	0.0079	0.0361***	0.0155**	-0.0044
	(0.0082)	(0.0097)	(0.0116)	(0.0104)	(0.0073)	(0.0114)
Paraiba	0.0929***	-0.0981***	-0.0280***	0.0359***	0.0354***	-0.0381***
	(0.0105)	(0.0083)	(0.0103)	(0.0095)	(0.0075)	(0.0099)
Pernambuco	0.0588***	-0.0909***	-0.0240***	0.0689***	0.0141***	-0.0268***
	(0.007)	(0.0056)	(0.0063)	(0.0069)	(0.0042)	(0.0067)
Alagoas	0.0819***	-0.1510***	-0.0131	0.0136	0.0602***	0.0083
	(0.0110)	(0.0077)	(0.0129)	(0.0099)	(0.0099)	(0.0124)
Sergipe	0.1097***	-0.1129***	-0.0265**	0.0711***	0.0371***	-0.0785***
	(0.0117)	(0.0083)	(0.0111)	(0.0115)	(0.0084)	(0.0096)

Bahia	0.1211***	-0.1259***	-0.0167***	0.0729***	0.0338***	-0.0853***
	(0.0076)	(0.0047)	(0.0061)	(0.0064)	(0.0044)	(0.0055)
Minas Gerais	0.1240***	-0.0422***	-0.0213***	0.0234***	-0.0142***	-0.0697***
	(0.0071)	(0.0053)	(0.0052)	(0.0048)	(0.0028)	(0.0052)
Espirito Santo	0.2091***	-0.0673***	-0.0427***	0.0079	-0.0216***	-0.0853***
	(0.0146)	(0.0093)	(0.0091)	(0.0084)	(0.0050)	(0.0092)
Rio de Janeiro	-0.0525***	-0.0611***	0.0060	0.0471***	0.0445***	0.0160**
	(0.0032)	(0.0058)	(0.0060)	(0.0060)	(0.0044)	(0.0069)
Paraná	0.1363***	-0.0289***	-0.0478***	0.0300***	-0.0323***	-0.0573***
	(0.0095)	(0.0068)	(0.0057)	(0.0059)	(0.0028)	(0.0064)
Santa Catarina	0.1333***	0.0420***	-0.0508***	0.0186**	-0.0435***	-0.0996***
	(0.0117)	(0.0098)	(0.0076)	(0.0073)	(0.0032)	(0.0076)
Rio Grande do Sul	0.1202***	0.0049	-0.0433***	0.0268***	-0.0201***	-0.0885***
	(0.0093)	(0.0069)	(0.0056)	(0.0056)	(0.0032)	(0.0058)
Mato Grosso do Sul	0.2021***	-0.0767***	-0.0331***	0.0116	-0.0219***	-0.0820***
	(0.0144)	(0.0090)	(0.0097)	(0.0085)	(0.0048)	(0.0090)
Mato Grosso	0.1754***	-0.0767***	0.0055	0.0258***	-0.0202***	-0.1099***
	(0.0128)	(0.0082)	(0.0100)	(0.0083)	(0.0047)	(0.0075)
Goiás	0.0965***	-0.0530***	-0.0003	0.0227***	0.0007	-0.0666***
	(0.0086)	(0.0067)	(0.0073)	(0.0063)	(0.0044)	(0.0066)
Distrito Federal	-0.0512***	-0.0421***	0.0545***	0.0060	0.0429***	-0.0101
	(0.0041)	(0.0082)	(0.0088)	(0.0070)	(0.0061)	(0.0090)
Number of boys in household	0.0135***	0.0041**	-0.0262***	0.0073***	-0.0079***	0.0091***
	(0.0008)	(0.0020)	(0.0021)	(0.0012)	(0.0011)	(0.0019)
Number of girls in household	0.0115***	0.0045**	-0.0243***	0.0056***	-0.0055***	0.0082***
	(0.0008)	(0.0020)	(0.0021)	(0.0012)	(0.0011)	(0.0019)
Number of working age males in household	0.0148***	-0.0214***	-0.0224***	-0.0139***	0.0084***	0.0345***
	(0.0008)	(0.0020)	(0.0018)	(0.0013)	(0.0009)	(0.0018)
Number of working age females in household	-0.0010	0.0314***	0.0312***	-0.0084***	0.0191***	-0.0723***

	(0.0011)	(0.0021)	(0.0019)	(0.0015)	(0.001)	(0.0025)
Number of males older than 60 in household	0.0429***	-0.0518***	-0.0419***	-0.0133***	0.0173***	0.0468***
	(0.0027)	(0.0067)	(0.0060)	(0.0043)	(0.0031)	(0.0068)
Number of females older than 60 in household	-0.0045	-0.0043	-0.0055	-0.0135***	0.0258***	0.0021
	(0.0030)	(0.0063)	(0.0056)	(0.004)	(0.0030)	(0.0065)
Observations	126930	Pseudo R^2	0.2205	Log pseudolikelihood	-77647869	

Notes: Robust standard errors in parentheses.

Table A3: Multinomial logit marginal effects, men and women aged 15-35, full results for 1993

	agricultural employment or self-employment	other manual employment	non-manual employment	self-employed outside agriculture	in education	not in work or education
schooling (years)	-0.0239***	-0.0231***	0.0531***	-0.0019***	0.0110***	-0.0151***
	(0.0003)	(0.0005)	(0.0005)	(0.0003)	(0.0002)	(0.0005)
age	-0.0043***	0.0053**	-0.0156***	0.0364***	-0.0437***	0.0218***
	(0.0014)	(0.0027)	(0.0026)	(0.0019)	(0.0015)	(0.0027)
age^2	0.0000	-0.0000	0.0003***	-0.0005***	0.0006***	-0.0004***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
female	-0.0436***	-0.1803***	-0.0608***	-0.1254***	0.0124***	0.3978***
	(0.0021)	(0.0036)	(0.0033)	(0.0025)	(0.0010)	(0.0033)
Aboriginal Brazilians	0.0340	0.0629	0.0385	0.0323	-0.0130	-0.1549***
	(0.0278)	(0.0602)	(0.0703)	(0.0369)	(0.0112)	(0.0296)
African Brazilians	-0.0284***	0.1314***	-0.0506***	-0.0259***	-0.0017	-0.0246***
	(0.0034)	(0.0086)	(0.0067)	(0.0042)	(0.0021)	(0.0072)
Asian Brazilians	0.0297	-0.0621**	-0.0269	0.0271	0.0251**	0.0071
	(0.0339)	(0.0312)	(0.0220)	(0.0259)	(0.0105)	(0.0336)
Pardo Brazilians	-0.0116***	0.0533***	-0.0182***	-0.0155***	-0.0030***	-0.0048
	(0.0021)	(0.0040)	(0.0036)	(0.0024)	(0.0010)	(0.0037)
Rondônia	-0.0258**	-0.0374**	0.0455**	0.0217	0.0146**	-0.0186
	(0.011)	(0.0161)	(0.0180)	(0.0152)	(0.0068)	(0.0165)
Acre	-0.0732***	-0.1005***	0.1170***	0.0617**	0.0762***	-0.0812***

	(0.0098)	(0.0228)	(0.0296)	(0.0289)	(0.0180)	(0.0232)
Amazonas	-0.0499***	-0.0753***	-0.0075	0.0682***	0.0500***	0.0145
	(0.0061)	(0.0110)	(0.0115)	(0.0125)	(0.0069)	(0.0130)
Roraima	-0.0491**	-0.0879***	0.0244	0.0629*	0.0242	0.0255
	(0.0226)	(0.0329)	(0.0354)	(0.035)	(0.0156)	(0.0422)
Pará	-0.0388***	-0.0886***	0.0016	0.1183***	0.0356***	-0.0281***
	(0.0054)	(0.0079)	(0.0093)	(0.0108)	(0.0044)	(0.0088)
Amapá	-0.0717***	-0.0984***	0.0797**	-0.0013	0.0570***	0.0349
	(0.0102)	(0.0239)	(0.0325)	(0.0240)	(0.0169)	(0.0324)
Tocantins	0.1064***	-0.0848***	-0.0074	0.0496***	0.0036	-0.0674***
	(0.0146)	(0.0122)	(0.0148)	(0.0146)	(0.0052)	(0.0123)
Maranhão	0.0272***	-0.1446***	-0.0666***	0.2931***	-0.0041	-0.1049***
	(0.0082)	(0.0080)	(0.0101)	(0.0153)	(0.0033)	(0.0092)
Piaui	0.0734***	-0.1475***	-0.0060	0.1551***	0.0209***	-0.0959***
	(0.0105)	(0.0085)	(0.0129)	(0.0142)	(0.0054)	(0.0100)
Ceará	0.0470***	-0.1154***	-0.0032	0.0843***	0.0157***	-0.0284***
	(0.0070)	(0.0060)	(0.0081)	(0.0085)	(0.0031)	(0.0076)
Rio Grande do Norte	0.0033	-0.0827***	-0.0182	0.0535***	0.0237***	0.0204
	(0.0084)	(0.0106)	(0.0120)	(0.0121)	(0.0053)	(0.0130)
Paraiba	0.0785***	-0.1354***	-0.0510***	0.0896***	0.0224***	-0.0042
	(0.0104)	(0.0085)	(0.0103)	(0.0123)	(0.0050)	(0.0118)
Pernambuco	0.0333***	-0.0904***	-0.0479***	0.0699***	0.0130***	0.0220***
	(0.0064)	(0.0062)	(0.0062)	(0.0076)	(0.0026)	(0.0078)
Alagoas	0.0552***	-0.1551***	-0.0242*	0.0388***	0.0277***	0.0575***
	(0.0105)	(0.0086)	(0.0127)	(0.0119)	(0.0057)	(0.0144)
Sergipe	0.0870***	-0.1162***	-0.0158	0.0327***	0.0391***	-0.0268**
	(0.0116)	(0.0096)	(0.0126)	(0.0116)	(0.0066)	(0.0123)
Bahia	0.0997***	-0.1477***	-0.0257***	0.0826***	0.0278***	-0.0367***
	(0.0072)	(0.0050)	(0.0065)	(0.0070)	(0.0030)	(0.0065)
Minas Gerais	0.1023***	-0.0625***	-0.0316***	0.0329***	-0.0032*	-0.0378***
	(0.0064)	(0.0054)	(0.0053)	(0.0052)	(0.0016)	(0.0057)
Espirito Santo	0.1829***	-0.0691***	-0.0660***	0.0263***	-0.0052*	-0.0689***
	(0.0138)	(0.0101)	(0.0087)	(0.0100)	(0.0029)	(0.0101)
Rio de Janeiro	-0.0681***	-0.0453***	0.0031	0.0371***	0.0237***	0.0493***
	(0.0032)	(0.0064)	(0.0061)	(0.0060)	(0.0026)	(0.0073)
Paraná	0.1392***	-0.0457***	-0.0478***	0.0281***	-0.0145***	-0.0592***
	(0.0088)	(0.0069)	(0.006)	(0.0062)	(0.0016)	(0.0064)
Santa Catarina	0.1517***	-0.0028	-0.0656***	0.0313***	-0.0192***	-0.0954***
	(0.0114)	(0.0096)	(0.0074)	(0.0081)	(0.0018)	(0.0078)

Rio Grande do Sul	0.1357***	0.0245***	-0.0693***	0.0342***	-0.0146***	-0.1105***
	(0.009)	(0.0073)	(0.0051)	(0.0061)	(0.0015)	(0.0055)
Mato Grosso do Sul	0.1589***	-0.0771***	-0.0362***	0.0216**	-0.0108***	-0.0563***
	(0.0135)	(0.0099)	(0.0102)	(0.0095)	(0.0028)	(0.0099)
Mato Grosso	0.1269***	-0.0879***	-0.0039	0.0316***	-0.0047	-0.0620***
	(0.0116)	(0.0088)	(0.0100)	(0.0093)	(0.0030)	(0.0090)
Goiás	0.1204***	-0.0707***	-0.0254***	0.0318***	-0.0038*	-0.0521***
	(0.0091)	(0.0071)	(0.0073)	(0.0071)	(0.0022)	(0.0072)
Distrito Federal	-0.0511***	-0.0429***	0.0244***	-0.0041	0.0314***	0.0423***
	(0.0056)	(0.0094)	(0.0090)	(0.0079)	(0.0041)	(0.0109)
Number of boys in household	0.0145***	0.0023	-0.0224***	0.0035***	-0.0030***	0.0050***
	(0.0009)	(0.0019)	(0.0020)	(0.0012)	(0.0005)	(0.0018)
Number of girls in household	0.0135***	0.0066***	-0.0192***	0.0045***	-0.0039***	-0.0016
	(0.0009)	(0.0019)	(0.0020)	(0.0012)	(0.0005)	(0.0018)
Number of working age males in household	0.0182***	-0.0284***	-0.0202***	-0.0148***	0.0042***	0.0409***
	(0.0009)	(0.0021)	(0.0019)	(0.0014)	(0.0004)	(0.0019)
Number of working age females in household	-0.0047***	0.0489***	0.0379***	-0.0033**	0.0096***	-0.0884***
	(0.0013)	(0.0021)	(0.0019)	(0.0016)	(0.0005)	(0.0025)
Number of males older than 60 in household	0.0430***	-0.0382***	-0.0205***	-0.0103**	0.0033**	0.0228***
	(0.0033)	(0.007)	(0.0061)	(0.0046)	(0.0017)	(0.0070)
Number of females older than 60 in household	-0.0056	-0.0035	-0.0013	-0.0069	0.0117***	0.0058
	(0.0038)	(0.0068)	(0.0059)	(0.0044)	(0.0016)	(0.0067)
Observations	115253	Pseudo R^2	0.2248	Log pseudolikelihood		-70826214

Notes: Robust standard errors in parentheses.

Figure A.1: Brazilian states