C | E | D | L | A | S

Centro de Estudios
Distributivos, Laborales y Sociales

Maestría en Economía
Facultad de Ciencias Económicas

UNIVERSIDAD NACIONAL DE LA PLATA

Youth Turnover in Brazil: Job and Worker Flows and an Evaluation of a Youth-Targeted Training Program

Carlos Henrique Corseuil, Miguel Foguel, Gustavo Gonzaga y Eduardo Pontual Ribeiro

Documento de Trabajo Nro. 155
Febrero, 2014

ISSN 1853-0168

YOUTH TURNOVER IN BRAZIL:

JOB AND WORKER FLOWS AND

AN EVALUATION OF A YOUTH-TARGETED TRAINING PROGRAM

Carlos Henrique Corseuil (IPEA)
Miguel Foguel (IPEA)
Gustavo Gonzaga (PUC-Rio)
Eduardo Pontual Ribeiro (UFRJ)

WRITTEN FOR THE REGIONAL RESEARCH PROJECT:

LABOR MARKETS FOR INCLUSIVE GROWTH IN LATIN AMERICA

SPONSORED BY

CEDLAS – IDRC

FINAL DRAFT

MARCH 13TH, 2013

Introduction

One of the main distinguishing features of the Brazilian labor market is its impressively high job and worker turnover rates. Even though turnover rates are very high for all workers, the literature has presented evidence that the contribution of some demographic groups, especially young workers, is quite significant to the observation of such high turnover rates in Brazil.

On the other hand, it is well-known that many workers face major obstacles to enter the labor market when they are young. There is ample evidence that unemployment rates for the 16-24 year-old age bracket are much higher than for other age groups, that young workers disproportionately hold informal and/or precarious jobs, such as temporary employment (Betcherman *et al.* 2007), and that they are taking most of the burden in many countries following the 2008 financial crisis (Bell and Blanchflower, 2010; Biavaschi *et al.* 2012).

Hence understanding what drives the attachment of young workers to formal jobs seems to be a promising path to reduce both youth turnover and unemployment rates. In this study we tackle the issue of turnover and labor market attachment of young workers from two perspectives.

We first document statistics on jobs and worker flows in Brazil, illustrating that youth workers display much higher turnover rates than other age groups. Yet in this first perspective we investigate whether this is an intrinsic characteristic related to the low age of these workers or whether this is a spurious relation due to other turnover determinants which may be correlated with age.

The second perspective is to evaluate the impact of a large youth-targeted program, which was substantially increased in 2000, on the formal labor market attachment of its participants. The program is *Lei do Aprendiz* (Apprentice Act), a targeted active labor market program conducted by the Labor Ministry, which concedes payroll subsidies to firms that hire and train young workers under temporary contracts.

Following a suggestion by one of the referees of the first report, this final report is organized as two different papers, each focusing on one of the perspectives described above.

PAPER ONE

YOUTH LABOR MARKET IN BRAZIL THROUGH THE LENS OF THE FLOW APPROACH[♦]

Carlos Henrique Corseuil (IPEA)
Miguel Foguel (IPEA)
Gustavo Gonzaga (PUC-Rio)
Eduardo Pontual Ribeiro (UFRJ)

Abstract

We use matched employer-employee data to study the situation of young workers in the (formal) labor market in Brazil. We employ the flow approach to draw a comparative picture of the patterns of the young and adult movements in the labor market during a period of fifteen years. We also estimate an econometric model that attempts to isolate the contribution of workers' age on employment duration. Our results show that youths experience very high rates of labor market turnover, a phenomenon that comes from elevated rates of hiring and separation from jobs. The estimates from the model show that the age of workers does contribute to decrease employment duration, with or without the inclusion of firm-specific fixed effects. In terms of policy, a decline in the separation rate may be attained through a combination of policies that involves the education system and labor market initiatives that create incentives for workers and firms to invest in each other.

♦ This paper is part of a joint CEDLAS-IDRC project on "Labour Demand and Job Creation: Empirical Evidence from Firms in Latin American". The authors acknowledge not only the financial support provided by the project but also the useful comments received in early stages of the paper by project coordinators and by Gary Fields who contributed to the project as a technical advisor. We also acknowledge the research assistance provided by Katcha Poloponsky and Alessandra Brito.

1. Introduction

One of the most worrisome and widespread stylized facts in Labor Economics is the observation of very low employment rates for young workers, usually resulting in very high unemployment rates. For instance, ILO estimates the youth global unemployment rate at 12.6% in 2011 (ILO 2012). But in some countries the figures are much higher as indicated by the OECD-average youth unemployment rate of 18.5% in the third quarter of 2010 (OECD 2010).

Brazil is no exception for this matter. According to a nation-wide household survey (PNAD/IBGE) the unemployment rate for 15-24 year olds was 16.3% in 2011, while the rates observed for ages 25-49 and 50+, were 5.7% and 2.8%, respectively, in the same year.

The main goal of this paper is to provide a more complete picture of labor market integration of young workers in Brazil. We use the flow approach as advocated by Blanchard and Diamond (1992) as the ideal setting to analyze labor market dynamics. The implementation of this approach is based on worker flow measures such as hiring, separations and turnover computed both for young and adult workers. We use a Brazilian matched employer-employee dataset (RAIS) from 1996 to 2010 to pursue all the empirical analysis.

The collection of results on new dimensions of youth labor market contributes to a more accurate diagnostic of the youth labor market problem. Before describing briefly our results it is worth mentioning two methodological contributions of this paper. The first one is a measurement procedure that identifies how much of young workers separation is due to adult workers crowding-in. The second one is the strategy used here to identify the age effect on employment duration, which is based on a hazard model with establishment fixed effects.

Our first results confirm larger flows for young workers than for adult workers. Perhaps the most striking result is the average turnover rate which amounts to 1.65 for youth workers, twice as large as the adults' figure. Also important is the fact that hiring rates are relatively higher than separation rates for young workers.

In general, low employment rates for young workers can occur either because of a low entry flow into employment or from a high exit flow from employment. Our results are consistent with the latter scenario, where a high exit flow (separations) from employment resulting from large turnover rates is probably the main determinant of high unemployment rates for young workers in Brazil (Flori, 2004, has provided some previous evidence on this).

The pattern of separations for young workers reveals two interesting findings. The first is that most replacements of jobs held by young workers are filled by other youths. Indeed, on average, less than 10% of all replacements of young workers are substitutions for adult workers. Similar results are observed in the other direction, i.e. the replacement of adults by young workers. This is compatible with a view in which young and adult labor enter the aggregate production function in (almost) fixed proportions.

The second interesting finding is the difference between youths and adults as reason for separation from a job. While lay-offs account for a higher share for adults, voluntary quits and the expiration of temporary contracts are relatively more important for younger workers. This is probably due to a combination of a more unstable labor supply behavior of youths and more frequent use of time-limited contracts to hire them.

Regarding the other component of turnover, namely the hiring rate, we investigate whether the high separation rate observed for youths could be attributed to attachment to high turnover jobs. We calculate the relative share of hirings for temporary contracts and for jobs at cooperatives, the latter typically considered quite unstable. Though we confirm that the use of temporary contracts is relatively more important for youths than adults, the difference does not seem to explain the more elevated separation rate for the former group. Jobs at cooperatives represent a negligible fraction of hirings, so it cannot explain the magnitudes observed for the separation rate.

Other job dimensions may be relevant to explain high flow rates for younger workers. In the last part of the paper, we address whether the high flows computed for young workers (in particular the high turnover rate) is an intrinsic characteristic of the lower age of these workers or, rather, whether it is a spurious relation due to other turnover determinants which may also be correlated with age. We take particular attention to establishment characteristics as we present evidence that young workers tend to be allocated to high turnover jobs. We use two complementary methods: a variance decomposition based on firm and worker characteristic and an econometric hazard model. The estimation of a hazard model including firms fixed effect as well as firms' and individuals' observable characteristics suggest that a lower age increases the hazard of separation even taking into account firm and worker controls.

Apart from this introduction, the paper contains six sections. In the second, we present the related literature and some labor market trends for youths based on stock measures. The third section describes the data and set out the basic flow measures used in the paper. In the fourth section, we look at the patterns of hirings in an attempt to check whether the higher separation rate observed for youngsters could be due to an allocation in which they start off from high turnover jobs. Section five contains a deeper look at differences in the pattern of separations between the groups. In section six, we use statistic and econometric models as an attempt to better measure the role played by workers' age in explaining the patterns of the job flow measures analyzed in the previous sections. The last section offers some conclusions.

2. Preliminaries

2.1. Related literature

The bulk of the literature on youth labor market relies on the analysis of stock variables computed from household surveys. Typically, the unemployment rate is the main indicator used in such analyses. We start this section by summarizing the stylized facts unveiled by the analysis of stock variables.

Freeman and Wise (1982) is recognized as an influent piece of work for understanding the underlying forces behind the youth labor market problem. Based on the collection of results in the volume the editors conclude that "Aggregate economic activity was the major determinant of the level of youth jobless in the United States". Another important conclusion was that "severe employment problems were concentrated among a small proportion of youths with distinctive characteristics".

The volume by Blanchflower and Freeman (2000) validates both conclusions for a more recent period (the 1990s) and a broader set of developed countries. The results were further extrapolated by O'Higgins (2003), who analyses the labor market for youths in developing and as well as transition economies. The qualitative results are broadly in agreement with those found for developed countries.[1]

[1] A notable exception is the more prominent role of supply factors in developing countries.

An important stylized fact specific to developing countries is the overrepresentation of young workers in the informal sector. See for instance Saavedra and Chong (1999). This is also studied in Maloney (1999), who associates this pattern to the finding that the informal sector tends to be the entry door for young workers in the labor market.

A somewhat related trend of using non-standard jobs as an entry door for young workers has been documented recently for EU countries with respect to temporary contracts. Evidence on this can be found either in O'Higgins (2012) or in OECD (2012). Both studies mention that the use of such contracts to hire youths increased in the last decade and carried on into the recent economic downturn.

The recent economic downturn also motivated novel contributions claiming that young workers are relatively more sensible to negative economic shocks. See for instance Bell and Blanchflower (2011) and O'Higgins (2012).

Another minor part of the literature proposes a different track to analyze the youth labor market problem. This track set the stage for the flow approach. Leighton and Mincer (1982) can be considered the turning-point contribution that inaugurates this new track. The authors decompose the unemployment rate by age group in unemployment incidence and duration. They show that the difference in unemployment rates between the groups is mostly due to differences in unemployment incidence. Leighton and Mincer (1982) also emphasize labor turnover as the most important dimension for analyzing the relationship between age and unemployment.

Following the same track O'Higgins (2001) focuses on unemployment duration. The author claims that short unemployment spells tend not to be harmful for young workers' prospects in labor market, but long term unemployment is.[2]

As higher unemployment duration can be a consequence of either a burst in separations or a drop in hirings, the flow approach arises as a natural direction to understand the youth labor market problem. This is the direction that we pursue in this paper.

---

[2] O'Higgins (2001) also argues that analyses based on unemployment are particularly problematic for early ages due to the ambiguous attachment of youths to the education system. Increases in unemployment induced by higher enrolment in schooling could be viewed as a good outcome instead of a bad one.

2.2. Youth Labor Market Trends based on stock measures

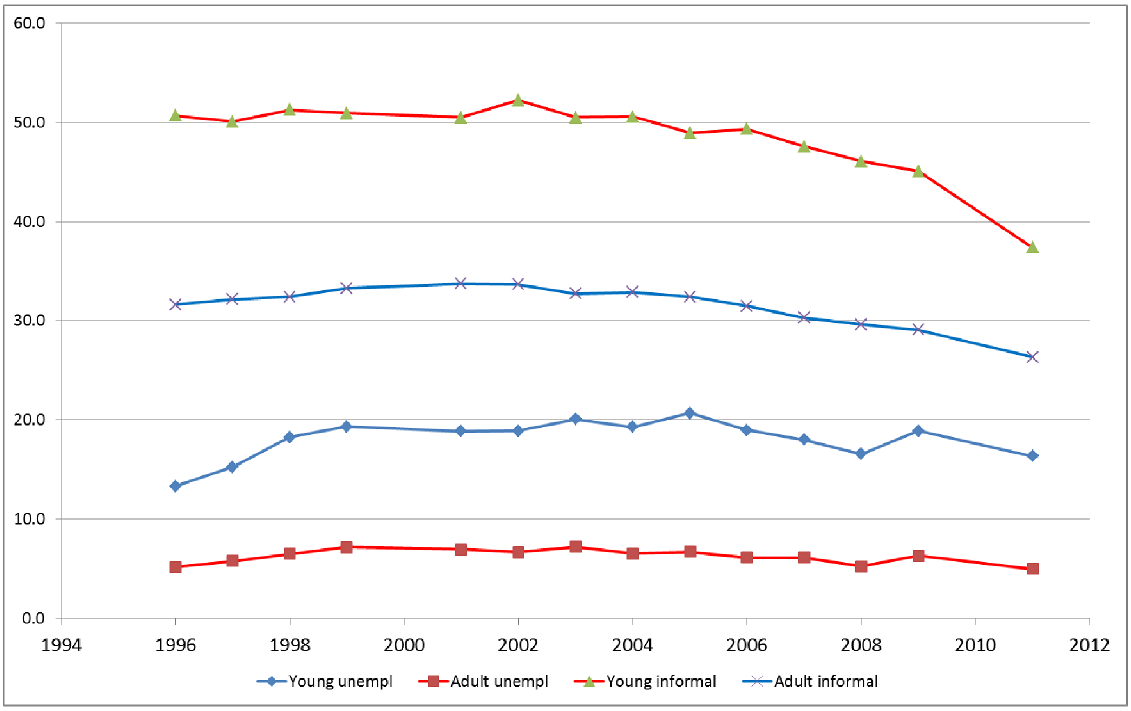
Before presenting our flow analysis based on Brazilian matched employer-employee data, we will briefly report in this section some evidence that the Brazilian labor market does not depart from the general trend summarized above. In particular we want to see if the trend of higher unemployment and higher informality for youths appear in the Brazilian household surveys. Throughout the paper, youths are all workers younger than 24 years old (inclusive) and adults are all workers above that threshold age.

This section only we rely on the main Brazilian household survey (Pesquisa Nacional de Amostra por Domicílios - PNAD[3]) to be able to measure the unemployment and informality rates. The unemployment rate is computed following the standard ILO definition, while informality rate is defined as the share of employed workers in one of the following categories: i) informal salaried worker, ii) self-employed, iii) non-salaried worker. We use data from 1996 to 2011, a period which comprises two distinct phases of the Brazilian labor market. Before 2003 both unemployment and informality showed either upward trends or stagnation at a relatively high level. Later there is a sharp declining trend on both indicators.

Indeed, youth labor market indicators differ sharply from the adult ones regarding informality and unemployment. Figure 1 below shows that: while the unemployment rate for adult workers fluctuated between 5% and 7% over the fifteen years we cover, the unemployment rate for young workers was 2 to 3 times higher, ranging from 13% to 21%. The informality rate is also much higher for youngsters. The share of informal workers peaked in 2002 at 34% for adult workers and at 52% for their younger counterparts. Both age groups benefited from the formality trend of the second part of the 2000's. But the lowest informality rate of 37% over the period for young workers, in 2011, was still well above the 26% observed minimum for adult workers in the same year.

[3] PNAD is a repeated cross-section with annual frequency, has national coverage and is conducted by the IBGE, the Brazilian Census Bureau.

Figure 1 – Unemployment and informality rates for age groups – 1996-2011



Source: Authors´ estimates based on PNAD/IBGE data. The nationally representative survey is not carried out in Census years (2000 and 2010).

In sum, the Brazilian labor market does seem to follow the general trend documented for other developing countries of both higher unemployment and informality rates for youths than for adults.

3. Worker flows: the contrast between young and adult workers

3.1. Basic Measures and Data

Our main data source comes from a Brazilian administrative database (*Relação Anual de Informações Sociais - RAIS*) which is maintained by the Brazilian Ministry of Employment and Labor (*Ministério do Trabalho e Emprego – MTE*). In Brazil, all registered, tax-paying establishments must send information to the Ministry regarding all employees who worked anytime during the reference year.[4]

[4] The absence of tax evaders from the sample prevents us from claiming that the data include information on all Brazilian establishments. Rather, RAIS gathers information on what is typically called the "formal sector".

RAIS provides matched employer-employee longitudinal data similar to those available in developed countries.[5] The data include worker specific information (such as gender, age and schooling), establishments (such as location and industry), and contract information (such as contracted wages, working hours, types of contract, hiring and separation dates, and reasons for separation). In our analysis we make intensive use of the last set of variables.

Our results rely on information on hirings and separations between 1996 and 2010 to compute traditional measures of worker flows, adapted to the context of age specific groups. We calculate the hiring and separation rates for age group a in year t as: $H_{at}= \Sigma_i h_{iat}/X_{at}$ and $S_{at}=\Sigma_i s_{iat}/X_{at}$, respectively, where i represents establishments, h_{iat} is the number of hires of workers of age group a over the course of the year t, s_{iat} is the number of separations for age group a over the year t at establishment i, and X_{at} is the aggregate average (between 31/12/t and 31/12/t-1) employment level of the group of workers under consideration.

These two rates can be combined to provide evidence on turnover. First, we aggregate the overall amount of worker flows using the worker turnover rate: $T_{at}= H_{at} + S_{at}$. The more heterogeneous is the workers' flow profile within firm × age cell, the higher is the distance between turnover and any of its components. Following this insight, another interesting measure is the churning rate, which is defined as $CH_{at}= T_{at} - |NET_{at}|$, where *NET* stands for net employment growth.

The context of age specific groups matter to the way labor flow measures are computed. In the traditional analysis at the firm level, *NET* could be either computed as:[6]

$$\mathrm{NET_t}= \mathrm{H_t} - \mathrm{S_t}, \qquad (1)$$

or as:

$$\mathrm{NET_t}=\Sigma_i \Delta n_{it}/X_t.$$

However when dealing with age specific groups, the two measures differ from each other due to individuals crossing the threshold that divides adjacent age groups while

[5] See Abowd and Kramarz (1999) for a description of the countries where this type of database was available and for information on how labor economics research has benefited from such databases.

[6] Abstracting from inconsistencies in information provided by firms on stocks and flows.

continuously employed in the same business unit *i*. These individuals do contribute to age group specific employment stock variation (Δn), but do not contribute neither to hiring (*H*) nor separation (*S*) rates. Hence we will rely on the first procedure and compute *NET* as in expression (1).

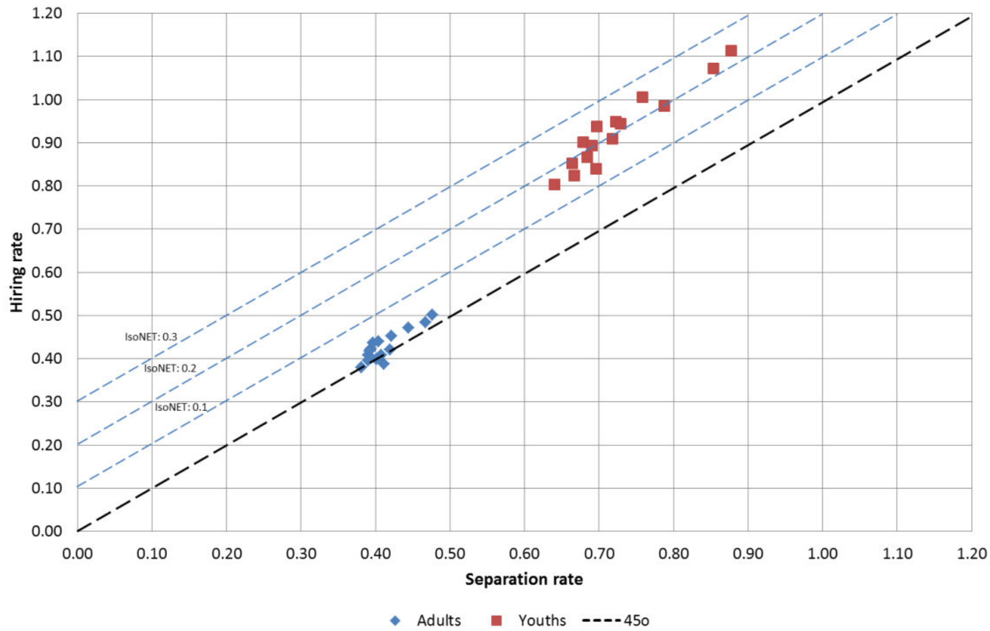
3.2 Youth Labor Market Trends based on flow measures

The striking differences between flows measures of young and adult workers are summarized in a single graph (Figure 2), where we plot *H* (vertical axis) and *S* (horizontal axis), elaborating on Burgess *et al.* (2001). There are large difference in these flow measures between the two groups, with both hiring and separations rates higher for youths. This implies that the turnover rate for young workers overtakes that for adult workers.

Figure 2 also shows a higher net employment rate for young workers. Points along the 45^{o} line corresponds to *NET=0*, as *H=S*. The scatter points for adult workers are either around or a little above the 45^{o} line, a pattern which evinces a small, positive average net employment growth for this group in the period of analysis. Net employment growth increases northwest with respect to the 45^{o} line. 'IsoNET' lines are also plotted in Figure 2, indicating the different combinations of *H* and *S* that yield 10, 20, and 30% net employment growth rates. Differently from the adults' pattern, the net growth rates for young workers tend be spread along the 20% IsoNet line. This shows that, on average, youths experience a much higher employment growth rate than adults in the formal labor market in Brazil. Figure 3 confirms that and shows that the growth rates have exhibited a slight increasing trend for both groups over the period of analysis.[7]

[7] It is interesting to note that this pattern of formal-sector employment growth based on RAIS is consistent with the patterns for the unemployment rate based on PNAD and described in section 2.2.

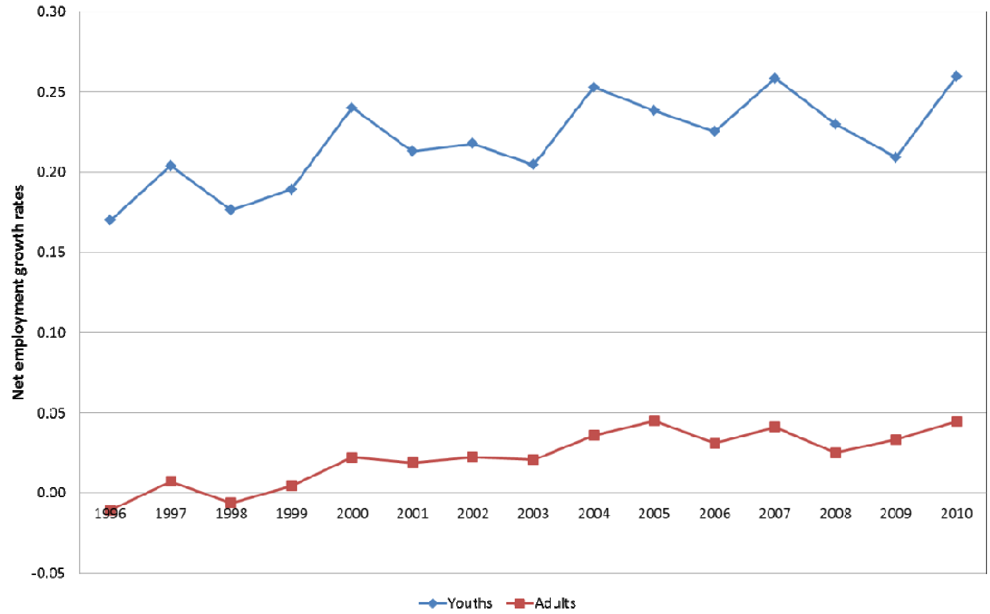
Figure 2: Hiring and separation rates by age group, 1996-2010



Source: Authors' calculations based on RAIS/MTE data.

Table 1 summarizes the time series of each group flow indicators. The Table initially reports the average value for the hiring rate (*H*), the separation rate (*S*), and the turnover rates (*T* and *CH*) for both age groups for the period 1996-2010. From Figure 2 we knew that both the hiring and separation rates were higher for younger workers. Table 1 brings the magnitude of such difference. The first striking result is that the average hiring rate for youths (92.6%) is more than two times higher than the average hiring rate for adults (42.8%). This indicates that Brazilian youths do not seem to face problems to get jobs in the formal labor market in the country. The difference in separation is a little less pronounced but still of considerable magnitude. Indeed, the figure for youths is as high as 72.4%, while it amounts to 41.3% for the older group (1.8 ratio).

Figure 3: Net employment growth rate by age group, 1996-2010



Source: Authors' calculations based on RAIS/MTE data.

The comparison of turnover rates adds these two differences and provides the second striking result. Turnover rates reach the impressive value of 165.1% for young workers and 84.1% for adults. The rate of 1.65 for youths means that there are more than eight younger worker transitions into and from formal employment for each five employed young workers, on average, each year. Even the adult turnover rate is very high for international standards (see, e.g., Davis and Haltiwanger, 1999[8] and Corseuil and Santos, 2006). Given the behavior of net employment growth rates mentioned above, the churning rate for youngsters decreases more than that for adults (with respect to the turnover rates), but the ratio between the two churning rates is still around 2.

From Figure 2 we can see that the hiring and separation rates for young workers are more disperse over time than the corresponding ones for adults. In order to take that into account, we calculated the coefficients of variation (CV), which are also presented in Table 1. The interesting finding brought by the CV calculation is the reversal of the order of the comparison of the hiring and separation rates between the two groups.

[8] Table 9 in Davis and Haltiwanger (1999) reports hiring and separation rates between 20% and 30% for developed countries.

Indeed, when the dispersion in the rates is incorporated, the differences between the groups become higher for separations (CV of 0.095 for youths and 0.069 for adults, ratio of 1.4) than for hirings (CV of 0.096 for youths and 0.084 for adults, ratio 1.1). One possible interpretation for this is that young workers flows, particularly hiring rates, are relatively more affected by the business cycle.

Table 1: Summary of flow indicators by age group, 1996-2010

	H	S	T	CH
Youths				
Average	0.926	0.724	1.651	1.448
CV	0.096	0.095	0.094	0.137
CORR GDP	0.708	0.563	0.652	0.563
Adults				
Average	0.428	0.413	0.841	0.823
CV	0.084	0.069	0.074	0.059
CORR GDP	0.555	0.405	0.505	0.429
Ratio Youth/Adults				
Average	2.2	1.8	2.0	1.8
CV	1.1	1.4	1.3	2.3
CORR GDP	1.3	1.4	1.3	1.3

Source: Authors' calculations based on RAIS/MTE data. H: Hiring rate; S: Separation rate; T = Turnover rate (T=H+S); CH= T - |H-S|. CV is the coefficient of variation. Corr GDP represents the simple correlation of each variable with GDP.

Table 1 allows us to check how close is the association between the business cycle and the age-specific labor flows. The rows Corr GDP in the Table display the correlation coefficient between each flow measure and the GDP. The correlations confirm that the flow measures for young workers are more sensitive to business cycle than the ones computed for adult workers (correlations for the former group are around 30 to 40% higher than that of the latter group).

The results presented in this section indicate that young workers do not seem to face strong barriers to formal labor market entry in Brazil. The lowest value of the hiring rate was 80%, its average amounted to more than 90%, and in three years (2007, 2008, and 2010) it surpassed 100%. But, though jobs are relatively easy to get, they are also riskier to lose. Indeed, the figures show that separations rates are also very high for youths:

minimum of 65%, average of more than 70%, and in two years (2008 and 2010) it was above 85%. As a result, young workers end up experiencing very high levels of labor market turnover. On the one hand transiting across many different jobs may enhance better matching with firms. On the other hand, entering and leaving jobs very easily tend to depress the acquisition of general and firm-specific labor experience. Since the accumulation of this type of human capital is important, the elevated turnover experienced by youths in Brazil is a factor that hinders the increase in their (future) productivity and wages.

4. A closer look at youth hiring rates: the role of unstable jobs

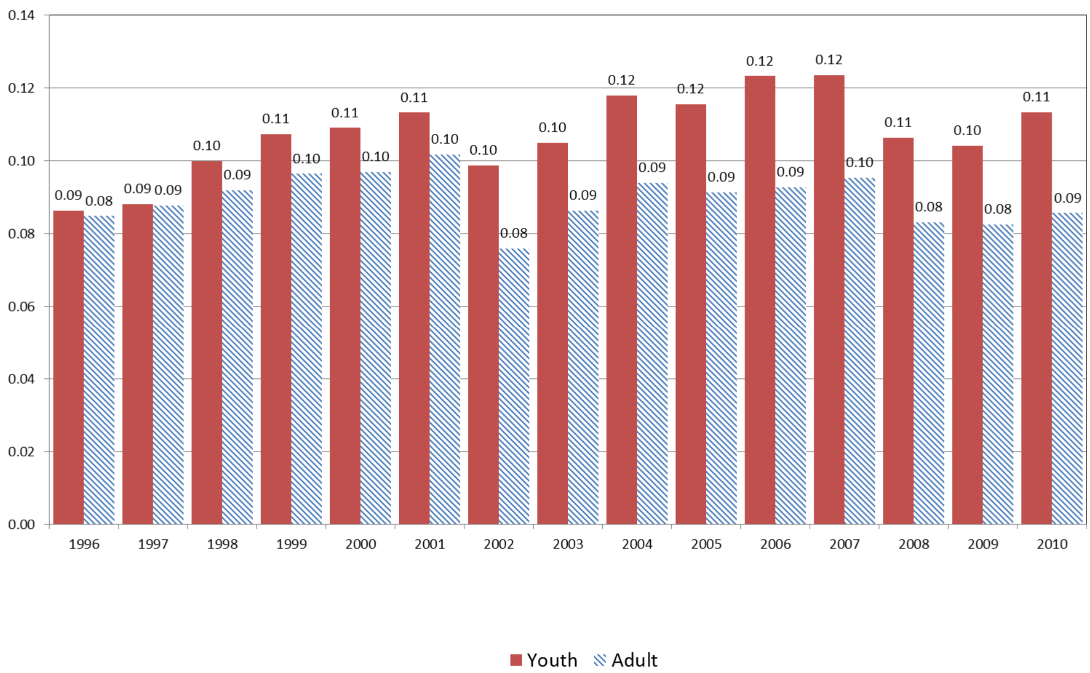
Hirings and separations from jobs are not necessarily independent events. For instance, in a developing country like Brazil, there is a large share of jobs of inferior quality (low wages, temporary contracts, unsatisfactory working conditions etc.) which are easily filled by the large share of less qualified workers available in the country. As they do not retain workers for long periods, high levels of hirings and separations are a common feature of this kind of jobs. If youths' hirings are overrepresented in this type of jobs, then at least part of the high levels of separations we observe for them comes from the high levels of hiring to unstable jobs. In order words, high separation rates could be induced by entrance "through the wrong door".

In order to investigate this possibility, we explore some features available in our data to check whether or not younger workers are overrepresented in some types of jobs that tend to have higher degree of instability. Specifically, we will look at the proportion of temporary jobs or jobs at cooperatives in hiring episodes involving young workers.

Figure 4 shows the share of hirings in temporary contracts for young and adult workers. Two points stand out from this Figure. First, the share of temporary contracts is much lower in Brazil than what is reported for EU countries: while the share for Brazilian young workers never reached 13% between 1996 and 2010, the figure is close to 43% in the case of EU countries (O'Higgins, 2012). Second, despite the similar values at the beginning of the period, the share of youths hired for temporary jobs rose, while the corresponding share for adults did not change much. In part, the rise observed for youths can be attributed to the increased use of the apprentice contract, which was launched by the government in 2000.[9]

Another form of unstable jobs often pointed as partly responsible for downgrading labor relations in the country are jobs offered by cooperatives ("coops").[10] Figure 5 shows the share of youths and adults that were hired by cooperatives between 1996 and 2010. The main point to notice from this Figure is that the fraction of the age groups hired by coops during this interval was less than 1%, that is, almost a negligible fraction. Coops do not seem to contribute to inflate neither the hiring nor the separation rates of young workers.

Figure 4: Share of hirings in temporary contracts by age group, 1996-2010



Source: Authors' calculations based on RAIS/MTE data.

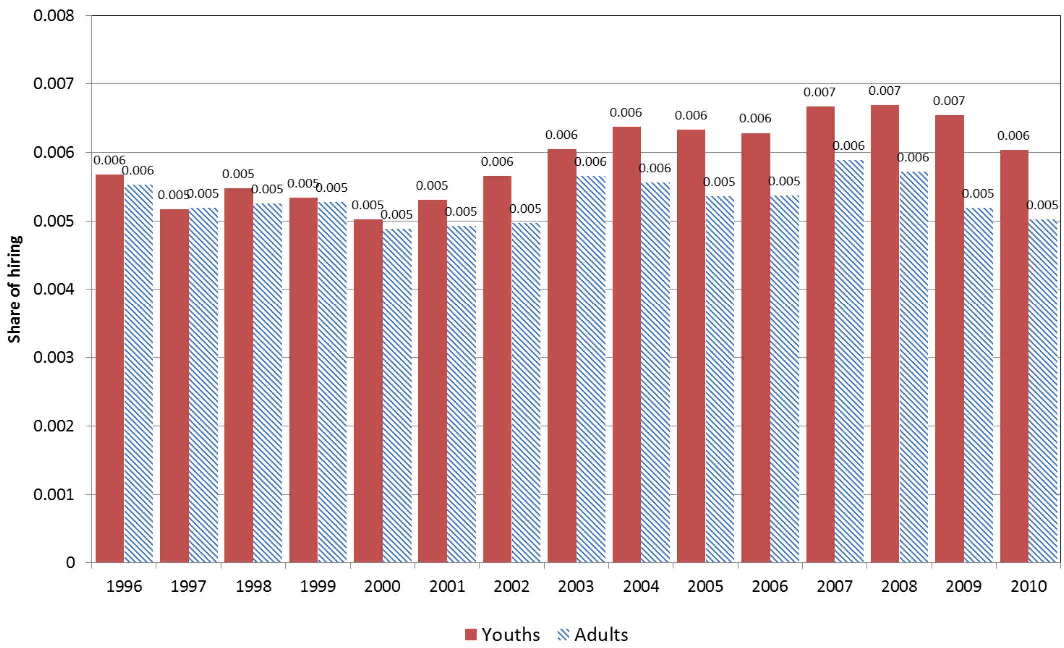
---

[9] See Corseuil *et al.* (2013) for an evaluation of the effects of the apprenticeship program on youth labor market outcomes.

[10] Since 1994, cooperatives have been exempted from paying many types of labor taxes. Many argue that this exemption has incentivized the creation of several "fake" cooperatives, i.e. coops that are hired by other firms only because they can "formally" hire workers without paying taxes. While we can identify cooperatives in the data, we cannot distinguish whether or not they behave in this counterfeit way.

In summary, the evidence presented in this section shows that the pattern of hirings of young workers either in temporary contracts or cooperatives does not seem to be responsible for the relative higher separation rate observed for this group. Further investigation seems necessary to check whether this connection from higher hiring to higher separation in fact exists and, if so, how it operates, especially for young workers. We focus on separation rates themselves in the next section.

Figure 5: Share of hirings by cooperatives by age group, 1996-2010



Source: Authors' calculations based on RAIS/MTE data.

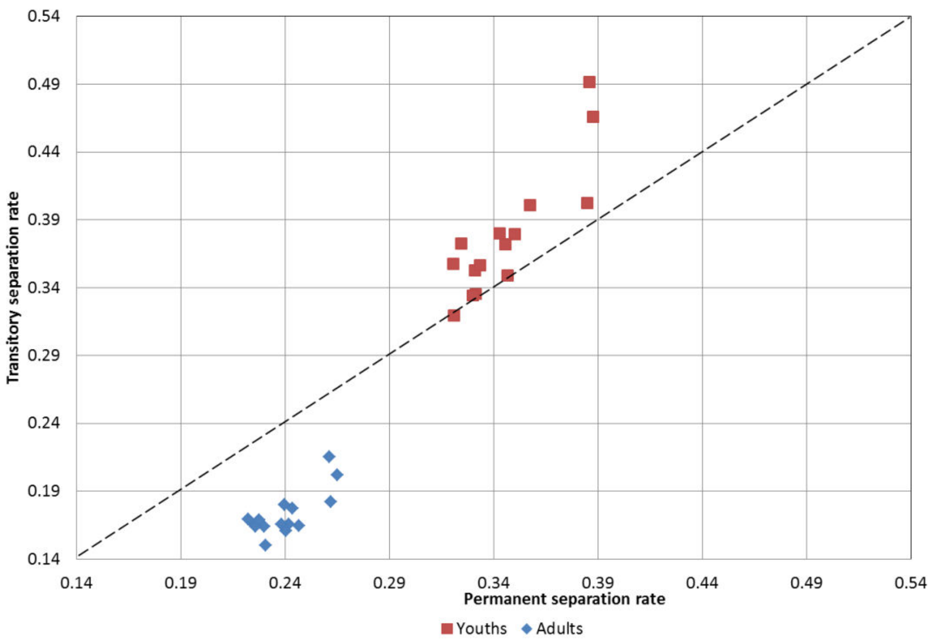
5. Looking for the determinants of higher separation rates for youths

In this section, we look deeper at the separation rate for young workers. Our data allow us to investigate three important dimensions of separations. First, we distinguish between what we call permanent and transitory separations. Second, we see whether separations are relatively more motivated by one of three reasons: the voluntary decision of the worker to quit the job, the decision of employers to lay-off the worker, or the simple expiration of a temporary contract. Third, we try to look at whether separations of young workers from their jobs result in a job destruction or the substitution of the young worker for an adult or another young worker.

5.1. Permanent versus transitory separations

Our first distinction is between permanent and transitory separations. It is argued that the Brazilian labor legislation induces fake lay-offs.[11] When this happens, data register a hiring and a separation of the same worker by the same firm within a certain period of time. In order to minimize the effect of this "double" counting, we define the permanent separation rate for age group *a* as: $S^{p}_{at}=\Sigma_i s^{p}_{it}/X_{at}$, where s^{p}_{it} is the number of separations that were not reverted at firm *i* during year *t*. The transitory separation rate can then be defined as: $S^{t}_{at}=S_{at}- S^{p}_{at}$, where S_{at} is the (gross) separation rate for age group *a* in year *t*.

Figure 6: Permanent and temporary separation rates by age group, 1996-2010.



Source: Authors' estimates based on RAIS/MTE data.

Figure 6 displays the figures for the two types of separations for youths and adults. As it can be seen, both rates are higher for the younger group but the permanent rate is relatively higher for adults (it represents on average almost 3/5 of all separations) than for youths (account for around half of all separations on average). While transitory

[11] One cited example is the situation where the employee negotiates a fake dismissal with the employer so that the worker can receive the unemployment benefit for some months and access the accumulated amount in his/her individual severance payment account (FGTS). See Camargo (1996) for a discussion of the (negative) incentives embodied in the Brazilian labor legislation.

separations do increase turnover rates, the sheer magnitude of permanent separations rates of about 34% (always larger than 30%) confirm the volatile attachment of younger workers, compared to older workers. Permanent separations for adult workers are lower, at about 24% and never higher than 28%.

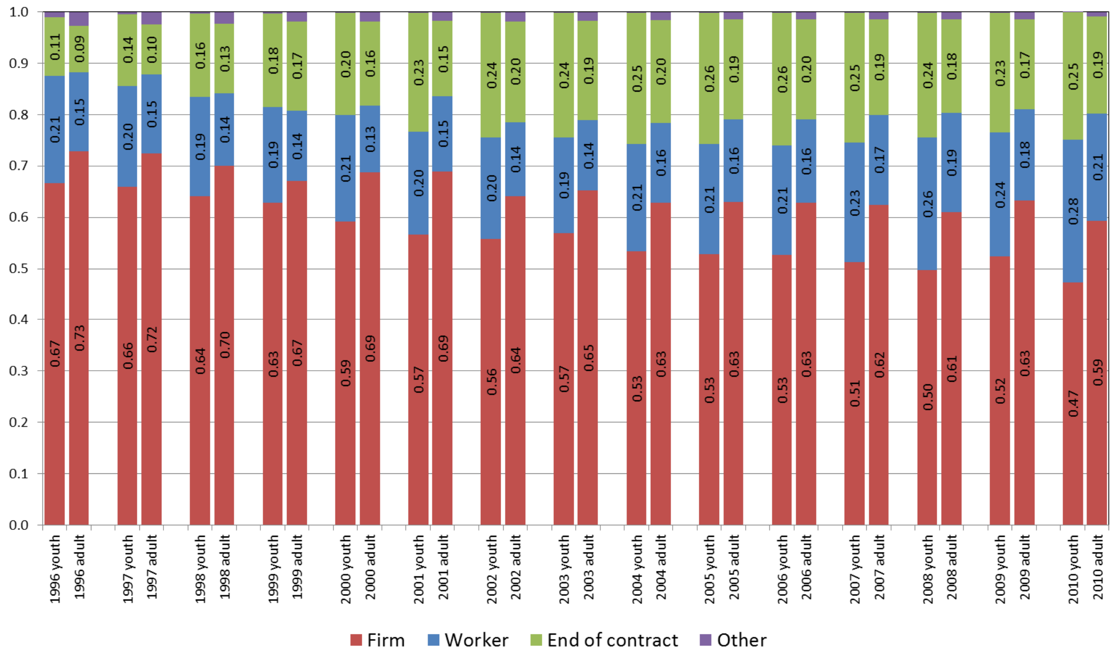
In order to analyze in a more structural fashion the differences in separations between our groups of interest, the rest of this section will be based on the measure of permanent separation.

5.2. Quits versus layoffs

Figure 7 shows that lay-offs are the most important reason for separations for both age groups. The Figure also shows that, though they have lost some importance over time for both groups, lay-offs are more relevant as a cause of separation for adults than for young workers. In fact, this difference has doubled over the years, rising from 6 p.p. in 1996 to 12 p.p. in 2010. The decline in the share of lay-offs was initially compensated by an increase in the contribution of expiration of temporary contracts but towards the end of the period there was a rise in the share of voluntary quits. It is worth observing that the termination of temporary contracts and quits are more relevant for youths than for adults and these differences have increased over time between the groups.

A set of factors can explain what we observe in Figure 7. First, as temporary contracts are relatively more relevant as a means to hire youth labor (see Figure 4), it is not surprising that separations due to the expiration of this type of contract are relatively more important for this group. Also, in the last two decades, labor legislation encouraged the use of more flexible forms of contracts (e.g. temporary jobs, part-time jobs, and temporary lay-offs) in many countries. This has not been different in Brazil, a fact that can explain the increase in this form of separation at least in the first part of our period of analysis.

Figure 7: Share of Separations by Proposer, 1996-2010



Source: Authors' estimates based on RAIS/MTE data.

Another factor is associated with the labor supply behavior of youths, who tend to "shop" jobs around more than adults. This can explain why quitting is more prevalent among the former group. In addition, as workers respond to the prevailing economic conditions, the supply side can also explain why we observe increases in the contribution of voluntary leaving for both types of workers in years of economic expansion. Firms may also have different sensitivities to dismissing adult and young workers over the economic cycle. We saw evidence of that in section 3, so at least part of the rise in the difference between the groups in the contribution of lay-offs may be attributed to the distinct response of firms to the last economic cycle in Brazil.[12]

5.3. Job destruction versus worker substitution

This subsection is based on a decomposition of separation rates. First, when a separation occurs it can ensue what we term *effective job destruction*, when the firm terminates

[12] Looking at the correlation coefficient between the share of lay-offs for each group and the GDP growth rate between 1996 and 2010, the estimate for youths is -0.57 and -0.52 for adults. The test of equality of these two estimates cannot be statistically rejected, though.

position. Second, when a substitution does take place, the worker can be replaced by another worker of same age (*within* substitution) or by a worker from a different age group (*between* substitution). We will decompose the separation rate in these three categories.

Let $JD_{at} = \Sigma_i \Delta n_{iat} I(\Delta n_{iat} < 0)/X_{at}$ be the job destruction rate for age group a, where i represents firms, t the year, and $I(.)$ is the indicator function that assumes value one when its argument is true and zero otherwise. Similarly, let $JC_{at} = \Sigma_i \Delta n_{iat} I(\Delta n_{iat} > 0)/X_{at}$ be the job creation rate for age group a in year t.

We define the *within age-group substitution* rate as the difference between the permanent separation rate and the job destruction rate for age group a: $W_{at} = S^p_{at} - JD_{at}$. The *between age-group substitution* rate is defined as: $B_{at} = min\{JC_{a't} * x_{a',a}; JD_{at}\}$, where a' denotes a different age group from a and $x_{a',a} = X_{a't}/ X_{at}$.[13] Finally, we can define what we call the *effective job destruction* rate for group a at time t: $EJD_{at} = JD_{at} - B_{at}$. The interpretation of these concepts is that from all separations that occurred for age group a in the economy, part resulted in the substitution of workers from the same age group, part in the substitution of workers from another age group and the rest is attributed to what would be the effective destruction of the job occupied by workers of group a. [14]

Perhaps the most interesting result revealed by Figure 8 is the low degree of substitution between youths and adults. Indeed, the share of substitution of one type of worker for the other is on average 4% and never surpasses the 5% level over the entire period of analysis. Though a more in-depth analysis would be needed, these low figures give an indication that young and adult labor enter the aggregate production function almost in a fixed proportion fashion. Figure 8 also reveals that substitution within the same age category is more common for youths than for adults, with a difference in shares of around 6 p.p. for the former group. It is also noticeable that replacement within the same age group became more important across the years for both groups. Indeed, there was a rise of more than 10 p.p. for youths and adults when we compare the share of within substitution in last half of the 1990's with the last half of the 2000's. The opposite

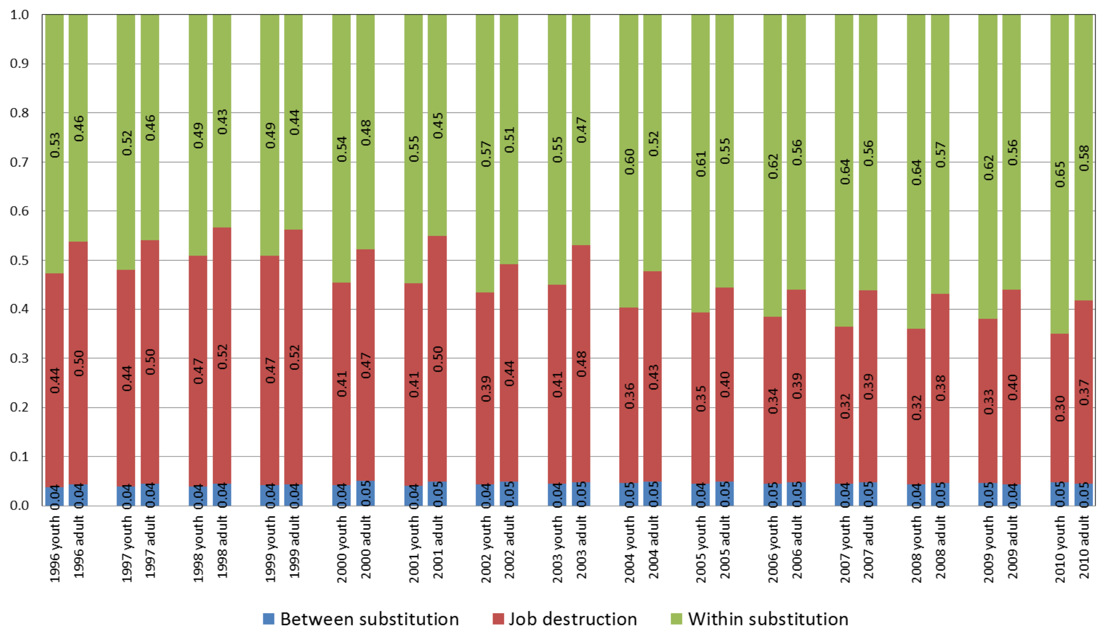
---

[13] The inclusion of the ratio $x_{a',a}$ is for compatibility with the denominator of the other rates.

[14] It may be easily seen that $S_{at}=W_{at}+B_{at}+EJD_{at}$.

movement happened with the share of separations due to job destruction. Again, part of this may be explained by the response of workers and firms to the economic cycle.

Figure 8: Share of Separations by Type: Job Destruction and Substitution Within or Between Age Groups, 1996-2010



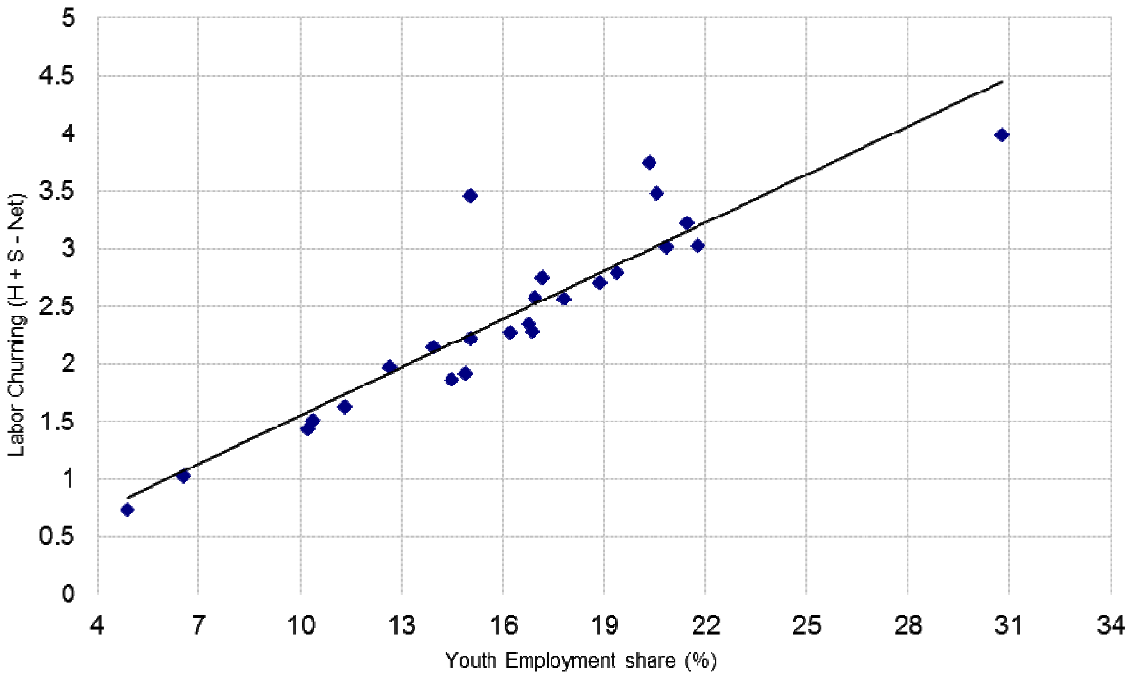
Source: Authors' estimates based on RAIS/MTE data.

In sum, the evidence presented in this section shows that: permanent separations are an important fraction of separations for youths, voluntary quits and the termination of time-limited contracts are relatively more relevant for youths than for adults, and the main form of replacement of youths is not a substitution by an adult but a substitution by another worker of the same age group. Another important result is that job separations (particularly lay-offs and job destruction) for both age groups seem to be affected by the business cycle. Although a deeper investigation would be needed to reach a clearer understanding of these findings, it is quite likely that demand, supply, and institutional factors, played their role in explaining the turnover patters in the last 15 years.

6. The effects of workers' age on turnover

One conclusion that emerges from the previous section is that young and adult workers may enter the production function in a fixed proportion fashion. That is to say that some jobs may be allocated only for young workers. Assume that the turnover rates for these jobs were intrinsically higher. If so, the higher turnover rates observed for young workers may be a consequence of their allocation to high turnover jobs. For instance, young workers may be allocated to high turnover industries, like construction or retail trade. Figure 9 below confirms that there are sharp differences in turnover or churning rates across industries in our data.

Figure 9 - Labor churning and youth employment share by industry



Source: Authors' estimates based on RAIS/MTE data.

In Figure 9 each point corresponds to an industry. The line represents the linear correlation between the average turnover (measured by labor churning) and the average share of young workers across industries between 1998 and 2010. The Figure clearly shows that establishments in high turnover sectors tend to employ a higher share of young workers. Therefore the high turnover computed for young workers may be, at least in part, due to the allocation of this group across industries. The main goal of this section is to go one step further towards the identification of a direct effect of age on turnover.

6.1. Variance Decomposition

The argument developed above may be generalized to other establishments' observable characteristics. In order to expand the analysis to a multi-dimensional approach, we first perform a traditional within-between variance decomposition. Worker and job flow metrics are calculated for cells defined by a combination of the following characteristics: industry, workers' age, establishments' age, establishment size and year.[15] These measures are then regressed against a series of dummies for the characteristics according to the following model.

$$Y_{a,j,k,m,t} = \alpha_a + \beta_j + \delta_k + \theta_m + \lambda_t + \varepsilon_{a,j,k,m,t},$$

where $Y_{a,j,k,m,t}$ represents a measure of either job or worker flow computed for the cell defined for worker of age "a", industry "j", plant age "k", size "m", and year "t". In the right hand side we have the terms capturing the effects of each of these variables, plus a cell idiosyncratic non-observable component.

Table 2 below presents the variance decomposition results for each job and worker flow measure. The Table reports the explanatory power of each characteristic, measured as the characteristic mean square divided by the regression mean square. We use mean squares instead of just squares in order to control for the higher explanatory power of characteristics with more degrees of freedom. Characteristics that explain twice the regression mean square are highlighted in gray.

Table 2 – Worker flow metrics within- and between-characteristics variance decomposition

	Year	Worker age	Firm size	Firm age	Sector
NET	0,48	10,32	0,09	10,38	0,03
H	0,58	8,34	0,18	9,84	0,17
S	0,18	7,78	0,13	2,21	0,88
T	0,61	5,49	0,29	7,01	0,56
CH	0,10	10,64	0,23	0,59	0,68
d_firma	0,86	6,93	0,57	0,23	0,54
df	*14*	*3*	*7*	*1*	*25*

Source: Authors calculations based on RAIS/MTE data. H = Hiring rate; S = Separation rate; T = Turnover rate (T=H+S); NET=Net Employment Growth (NET=H-S); Churning rate (CH=T – |NET|)). Df – characteristic degrees of freedom (number of categories minus one). Cells measure the ratio between the characteristic mean square and the explained mean squares. Characteristics mean squares are measured comparing residual sum of squares with and without the variable group (Anova decomposition). Gray cells indicate characteristics with twice the regression mean square.

[15] We use three age categories (14-17; 18-23 and 24-60), seven firm size categories (0-4; 5-9; 10-19; 20-49; 50-99; 100-249; 250-499; and 500 and more workers), two firm age categories (up to 4 years of age; five years or more) and twenty five industry classifications (roughly following national accounts, *subsetor ibge*). The industry classification is coarse due to the many classification changes over the period.

The results indicate that the workers age generate the sharpest difference across cell flow measures. Other characteristics are relevant (and significant), but with smaller explanatory power. Business unit age seems more important than industry and firm size. Business cycle variation has relatively small explanatory power, a result that contrast with one of the stylized facts in the literature, namely the dominant role of business cycle to explain youth unemployment. Therefore, there seems to be an important role for cell idiosyncratic non-observable characteristics in explaining job and worker flow. This is further explored in the next section.

6.2. Estimation of hazard models

The decomposition above can be considered an illustrative first step in trying to isolate the intrinsic contribution of age to turnover. But the challenging identification problem remains as the results on the previous section cannot be used "*prima facie*" to address this question. The reason is that age may be associated either to individual characteristics or unobservable job characteristics. For instance, it has been previously shown in the literature that young workers accumulate less human capital[16], a finding that can be associated with high turnover.

An ideal setting to tackle this issue is to analyze longitudinal worker level data that carries information on the establishment that hosts his job. At this level, we can analyze the determinants of employment duration considering both individual and job characteristics.

The standard econometric procedure to study duration events is the estimation of hazard models. Our previous discussion suggests that such estimation should be able to take into account idiosyncratic characteristics on top of observable characteristics. We chose to use the following proportional hazard model specification with fixed effects:

$$h_{ij}(t) = \alpha.D_{ij} + \beta.X_{ij} + \gamma_j (t) , \qquad (2)$$

where $h_{ij}(t)$ denotes the logarithm of the hazard rate of worker *i* completing her employment spell at establishment *j* at a length t. D_{ij} are workers age-group dummies, and X_{ij} are observable characteristics of workers and establishments (to be detailed

[16] See, for instance, Farber (1998) and references therein.

below). These variables are measured at the start of the corresponding employment spell. The term $\gamma_j(.)$ is the baseline hazard function of the employment spells. As we use a Cox version of the proportional likelihood estimator, there is no need to specify any parametric form for the baseline hazard.

The key departure of our specification from conventional proportional hazard models is to allow variations of the baseline hazard across establishments.[17] This takes into account any non-observable specificity at the firm level that may affect the hazards of its employees, even if such specificity is also correlated with any other observable characteristic.[18] This strategy enhances the credibility of our identification strategy as does the inclusion of a fixed effect term in conventional regression models with panel data.[19] As mentioned before, workers with similar observable characteristics could have different separation rates because of heterogeneity across firms in unobservable characteristics (like manager tolerance with either worker performance or behavior in the workplace). Allowing establishment idiosyncratic effects as specific baseline hazard rates allows us to compare workers that are in the same establishment (and therefore subject to the same idiosyncratic factors). This information is delivered by the parameter α, which informs how the hazard rates vary among similar workers in the same firm by age group.

As pointed out by Chamberlain (1985) we can use partial likelihood (PL) methods to get rid of $\gamma_j(.)$ and estimate the model without further complications. Allison (1996) shows that such estimator, which he refers as fixed-effect partial likelihood, performs very well with simulated data despite Chamberlain's concerns with the validity of one assumption for PL methods in the context of duration models.[20]

---

[17] Another important departure from duration models with longitudinal data is that we do not have multiple spells of the same individuals, but rather multiple workers of the same establishment, where each worker contributes with a different spell within the same establishment.

[18] Allowing such possible correlation is not a standard procedure in the economics literature using hazard models. These non-observable specificities, when incorporated, are usually treated as independent from all observable covariates.

[19] As a matter of fact, the following model specification with an additive fixed-effect component, analogous to conventional regression models using panel data, is a special case of our model:
$h_{ij}(t) = \alpha_j + \delta.D_{ij} + \beta.X_{ij} + \gamma(t)$.

[20] Specifically, Chamberlain pointed that in the context of multiple spells for each individual the censoring time for the last spell depends on the lengths of the preceding spells. This would violate a necessary condition for the implementation partial likelihood methods. In addition to Allison's

For the analysis of hazard rates, we use all episodes of hirings that took place in the period from 1996 to 1998. We measure the employment spell following the worker-establishment match until one of three restrictions occurs: i) the match is broken and the establishment keeps employing other workers, ii) the establishment leaves the market (or at least disappears from RAIS), and iii) the match survives until the last year of our data (2010). If the match faces one of the two last restrictions we classified the employment spell as a censored one.

As in other applications using RAIS data, we apply some filters. First, we eliminate separation episodes that resulted from individual death or retirement. We also exclude from the analysis employment spells that satisfy, at the initial point, at least one of the following conditions: worker aged 55 or older, in agriculture, in the public sector, or under a temporary contract. These procedures leave our sample with 27,162,416 employment spells. For each employment spell we collected information on workers age, gender, schooling level, and contractual number of hours.

In order to get more intuition from the results, we present two alternative specifications for the hazard model: the first includes plant fixed effects and the other does not. Table 3 below presents the results. The first three rows report the results for the effect of distinct age categories on the hazard rate relative to the base age category of over 30 years old. The first thing to notice is that the dummies for young ages (14 to 17 and 18 to 23 years old) are associated with positive and significant coefficients, in both specifications. This confirms that hazard rates are higher for young workers. However it is interesting to point the non-monotonic effect of workers age on the hazard rate. Late young workers (18 to 23) are associated with the highest hazard, irrespective of the model specification.

downgrade in this issue, in our case there is no particular reason to believe that a worker's employment spell will be influenced by the ones of his colleagues.

Table 3: Hazard estimations for the separation of a worker from the current employer

	plant fixed-effect		no plant fixed-effect	
Parameter	**Coefficients**	**Std Errors**	**Coefficients**	**Std Errors**
[14-17]	0.06839	0.00190	0.13461	0.00154
[18-23]	0.09718	0.0008832	0.16531	0.0007722
[24-30]	0.04727	0.0008430	0.06041	0.0007701
very low education	0.16722	0.00155	0.69838	0.0009425
low education	0.10209	0.00121	0.27499	0.0008598
medium education	0.06648	0.00116	0.09331	0.0008985
man	0.01825	0.0009346	0.01230	0.0006907
part-time	0.06070	0.00239	-0.13722	0.00158
1996	-0.04523	0.00141	-0.06877	0.00122
1997	-0.03690	0.0007298	-0.03153	0.0006214

Note:

Very low education: First half of primary education
Low education: Second half of primary education (but not complete)
Medium education: Completed primary education or incomplete secondary education
Part-time: Work 30 hours or less per week
Basel categories: women older than 30 years, highly educated that was hired in 1998.

Also interesting is the comparison across specifications of the estimated values of the coefficients for the first two age categories. The introduction of establishment fixed effects has similar impacts on the estimated coefficients of these two age categories. The effect for the 14 to 17 years of age group (relative to the base category of older than 30) decreases from 0.135 to 0.068 as we add establishment fixed effects. This 0.07 difference between the two specifications, almost doubling the coefficient, represent around half the initial estimate and it is also observed for the 18 to 23 category, for which it represents around 45% of the initial estimate. This common pattern suggests that both "teen" (14 to 17) and late young workers (18 to 23) tend to be allocated in high turnover establishments, relative to older workers.

The remaining rows in Table 3 report the estimated coefficients of the control variables in each one of the two specifications. An analysis of the estimated values for these coefficients can be grouped into three categories according to how the effect changes across model specifications. Firstly, the effect of education is also reduced once we add establishment fixed effects. In fact, the magnitude of this reduction is even higher than the one registered for age. We should point out however that, despite this reduction, the magnitude of the effect of education is still very high. Moreover, as it can be seen, the lower the education level of the worker, the higher the impact on turnover. Secondly, the effect of gender is stable across model specifications. Finally, the effect of working under a part-time contract not only increases once we add establishment fixed effects, as it flips sign, becoming positive.

7. Concluding comments

Using a rich employer-employee dataset we were able to draw an overall picture of how youths have performed in the formal labor market in Brazil in a recent period of 15 years. Based on the flow approach, we show that both the hiring and the separation rate for this group are quite high both in absolute and relative terms. The average figures for the hiring and separation rates for youths are over 90% and 70%, respectively, leading to an impressive turnover rate of more than 160%, twice the value observed for adults. Though it may induce better matching with firms, such a high level of turnover tends to hamper the accumulation of firm-specific experience, which can be an important form of human capital. A lower level of productivity can result, producing negative impacts at both the individual and the aggregate levels.

We look deeper at each component of the turnover rate. Potentially, an elevated hiring rate has both a positive and a negative side. On the one hand, it makes it easier for youths to get a job but, on the other, it generates less incentive to keep them. This last force induces job separations and therefore diminishes the duration of employment. In addition, if youth hirings are concentrated in unstable jobs, even higher levels of separations are expected. Our initial empirical investigation of the connection from hirings to separations was able to find some evidence that youths do not seem to be particularly allocated to more unstable jobs (temporary contracts or cooperative jobs).

Looking at job separations patterns, we found that quits are more prevalent among young workers than among adults. As mentioned, this can be associated with the high hiring rates of the former group. But it can also be associated with the supply behavior of youths, which typically involves more “shopping” across jobs in the labor market. We also found evidence that separations due to the expiration of temporary contracts are relatively more important for youths than for adults. The results also show that this cause of separations increased for both groups during the 2000’s, a phenomenon that may have to do with the introduction of incentives to use more flexible forms of labor contracts. Finally, we also found that a considerable fraction of separations do not end up in job destruction but rather in the replacement of one worker for another. In particular, the results evince that the more prevalent form of substitution is not across workers of different age groups but between workers of the same group.

Going one step further, we investigated to which extent one can say that the high turnover measures for youths can be attributed to their younger age. In other words, we conducted some exercises to isolate the contribution of the workers' age from that of other factors. This was carried out through two exercises. The first was a statistical model that tried to separate out the relative importance of age to explain the variation observed in various flow measures we used throughout the paper. The second was an econometric model of duration that tried to isolate the contribution of age on the duration of employment. Establishments' unobserved characteristics were incorporated in this last model. The results from both types of models show that to some extent the age of the worker contributes to explain the higher turnover rates and the lower employment duration observed for younger workers.

Condensing these results for policy purposes, the main empirical result is that young workers experience very high rates of turnover in Brazil due to both hiring and separation rates. In order to make the turnover rate decline, the main margin of policy attention should be the separation rate. Indeed, though hirings and separations are interrelated, tackling the problem of high levels of separations looks more efficient in the sense it directly attempts to keep workers longer in their jobs. The high hiring rates does not credence a lack of jobs for youth. Rather, the high separation rates imply short lived unstable jobs.

One must firstly recognize that other factors apart from the age of individuals operate. In particular, as the results of section 6 show, the education of workers seems to be an important factor to decrease turnover. In this sense, the more the education policy accelerates the increase in the schooling level of the new cohorts of workers, the lower should be the turnover expected for them.

Labor market policies should also be part of the strategy to lower the separation rate. Probably, job search assistance initiatives cannot do much, unless they are capable of generating worker-firm matchings that produce longer employment durations. Providing wage or tax subsidies for firms to extend the tenure of youths should be thought very carefully in particular because its costs can become very high. One could also devise a system that creates incentives for young workers and firms to increase the value of longer job relationships. Finally, training programs partially funded by the worker and the firm may create incentives for both parties to invest in each other in the longer term.

References

Abowd, J. M. and Kramarz, F., (1999). "The analysis of labor markets using matched employer-employee data" in: O. Ashenfelter & D. Card (ed.), Handbook of Labor Economics, volume 3, chapter 40, pages 2629-2710 Elsevier

Allison, P. (1996) "Fixed-Effects Partial Likelihood for Repeated Events" *Sociological Methods & Research*, 25(2): 207-222.

Blanchard, O.; Diamond, P. The flow approach to labor market. *American Economic Review*, Papers and Proceedings, v. 82, p. 354-359, 1992.

Blanchflower, D. and Freeman, R. (2000) Youth Employment and Joblessness in Advanced Countries, Chicago: NBER and University of Chicago Press.

Bell, D. and Blanchflower, D. (2011) "Youth Unemployment in Europe and the United States", IZA Discussion Paper n. 5673.

Burguess, S., Lane and Stevens (2001), "Churning dynamics: an analysis of hires and separations at the employer level", *Labour Economics*, 8(1): 1-14.

Camargo, J.M. (1996): *Flexibilidade no Mercado de Trabalho no Brasil*, FGV, Rio de Janeiro.

Chamberlain, G. (1985) "Heterogeneity, Ommited Variable Bias, and Duration Dependence" pp.3-38 in Longitudinal Analysis of Labor Market Data, ed. J. Heckman and B. Singer. Cambridge Univ. Press.

Corseuil, C.H. and L.S. Santos (2006) *Criação, Destruição e Realocação de Empregos no Brasil*. Rio de Janeiro: IPEA.

Corseuil, C. H., Foguel, M. and Gonzaga, G. (2013) "The effects of an apprenticeship program on labor market outcomes of youths in Brazil" REAP working paper 042

Davis, S. and Haltiwanger, J.. *Gross job flows*, volume 3B of *Handbook of Labor Economics*. North Holland, 1999. editors Orley Ashenfelter and David Card.

Farber, H. (1998), "Mobility and Stability: The Dynamics of Job Change in Labor Markets", Working Paper, Industrial Relations Section 400, Princeton.

Flori, P. (2004), "Desemprego de Jovens no Brasil", mimeo.

Freeman, R. and Wise, D. (1982) The Youth Labor Market Problem: Its nature, causes and consequences. Chicago: NBER and University of Chicago Press.

ILO (2012) Global Employment Trends for Youth 2012. ILO, Geneva.

Leighton, L. and Mincer, J. (1982) "Labor Turnover and Youth Unemployment", in Freeman and Wise (editors) The Youth Labor Market Problem: Its nature, causes and consequences. Chicago: University of Chicago Press

Maloney, W. (1999) "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico", The World Bank Economic Review 13(2): 275-302.

OECD (2010) "Off to a good start? Jobs for the Youth" available at: **http://www.oecd.org/els/emp/46717876.pdf**

O'Higgins (2001) Youth Unemployment and Employment Policies: a global perspective. ILO, Geneva.

O'Higgins (2003) "Trends in the youth labor market in Developing and Transition Countries", World Bank Social Protection Discussion Paper n.0321.

O'Higgins (2012) "This time it's different? Youth labor markets during 'The Great' Recession". IZA Discussion Paper n. 6434.

Saavedra, J. and Chong (1999) "Structural Reform, Institutions and Earnings: Evidence from the Formal and Informal Sectors in Urban Peru", The Journal of Development Studies 35(4): 95-116

PAPER TWO

THE EFFECTS OF AN APPRENTICESHIP PROGRAM ON WAGES AND EMPLOYABILITY OF YOUTHS IN BRAZIL[1]

Carlos Henrique Corseuil (IPEA)
Miguel Foguel (IPEA)
Gustavo Gonzaga (PUC-Rio)

JEL CODE: J13, J63
KEY WORDS: EVALUATION; APPRENTICESHIP PROGRAM; YOUTH-TARGETED ALMP

Abstract

The objective of this paper is to evaluate the Apprenticeship program (*Lei do Aprendiz*) that has been adopted in a large scale since 2000 in Brazil. This is a youth-targeted ALMP, which concedes payroll subsidies to firms that hire and train young workers under special temporary contracts aiming to help them to successfully complete the transition from school to work. We make use of a longitudinal matched employee-employer dataset covering the universe of formally employed workers, including apprentices. Our identification strategy exploits a discontinuity by age in the eligibility to the Apprenticeship program. We examine the program impacts in terms of wage growth and attachment to the formal labor market using other temporary workers as a control group. We find that the program increases the chances of getting relatively better paid and more stable jobs, especially in the medium run – four and five years after the program.

[1] The authors would like to thank Cecilia Machado, Claudio Ferraz, Marcos Rangel, Ajax Moreira, and Danilo Coelho for their comments; Katleem Marla Pires de Lima for helpful guidance in interpreting the Brazilian apprenticeship law; and Katcha Poloponsky for providing superb research assistance with data processing. The usual disclaimer applies.

1. Introduction

It is well-known that many workers face major obstacles to enter the labor market when they are young. There is ample evidence that unemployment rates for the 16-24 year-old age bracket are much higher than for other age groups,[2] that young workers disproportionately hold informal and/or precarious jobs, such as temporary employment (Betcherman *et al*. 2007), and that they are taking most of the burden in many countries following the 2008 financial crisis (Bell and Blanchflower, 2010; Biavaschi *et al.* 2012).[3] These facts brought youth employment to the forefront of recent policy debate with an increasing number of countries adopting youth-targeted active labor market programs (ALMPs) with a predominant focus in training (OECD, 2010).[4]

The general message from the literature is that there is substantial variation across countries and regions in the effectiveness of youth-employment policies, reflecting the large diversity of labor regulations, institutional arrangements for educational and training systems, and the relative importance of the informal sector.[5] The main challenges for the youth to successfully complete the transition from school to work are mostly influenced by institutional factors, such as the format of general education and vocational training systems, as well as regulations and existing ALMPs (Biavaschi *et al.* 2012). The role of the informal sector in providing training for young workers has also been recognized as an essential characteristic of labor markets in developing countries (Betcherman *et al*. 2007).

---

[2] In 2009, the ratio of young workers' unemployment rates (15-24 year old workers) over aggregate unemployment rates for a large sample of countries was around 2. The following ratios were observed in 2009: 2.04, on average, for the OECD countries; 2.14, on average, for UE-15; 1.9 in the U.S.; 2.09 in Brazil; 1.93 in Mexico; 2.3 in Chile; 1.8 in Colombia (OECD, 2010).

[3] In Brazil, according to the national household survey (PNAD), the unemployment rate for 15-24 year olds reached 18.9% in 2009, while the rates observed for ages 25-49 and 50+, were respectively, 7.1% and 3.7%. These rates are similar if one considers a longer period: the respective averages for the three age groups for the 1992-2009 period were 17.1%, 6.6%, and 3.6%.

[4] The main justification for having ALMP's targeted to young workers is based on evidence of "scarring" effects of early unemployment and job-loss experiences (see Eliason and Storrie, 2006; Skans, 2011; Nilsen and Reiso, 2011; and Cruces *et al.*, 2013 for recent studies on this issue). The literature has increasingly stressed the importance of the early years in a worker's career, a period in which workers make important human capital accumulation decisions that may be affected by the fact that young workers are probably the most vulnerable group to economic fluctuations (Adda *et al*., 2013). On the other hand, some youth-targeted ALMP's may not be optimal, especially those that aim reducing turnover through increasing rigidities, given a tendency of young workers to experiment new job matches which is not only natural but also desirable from an efficiency point of view.

[5] Adda *et al.* (2013), for instance, provide evidence that young workers were less hit by the Great Recession in countries with better designed vocational training institutions, like Germany, Austria and Switzerland.

By contrast, the number of reliable evaluations of youth-targeted ALMPs is still relatively small, especially in low and middle-income countries. Despite a recent increase in the number of studies using randomized experiments and other methods to deal with non-random selection of participation, more efforts devoted to evaluating the effectiveness of youth-targeted programs are clearly needed as well as the additional task of identifying which components of each program are important in each context.

The main goal of the paper is to evaluate a very large youth program, the Apprenticeship program (*Lei do Aprendiz*) that has been adopted in a large scale since 2000 in Brazil. This is a youth-targeted ALMP conducted by the Brazilian Labor Ministry, which concedes payroll subsidies to firms that hire and train young workers (aged 14-17 years from 2000 to 2005; 14-24 years since 2005) under special temporary contracts. The program intends to provide professional skills to young workers and help them to successfully complete the transition from school to work. Its main objective is to place participants in formal first jobs with adequate specialized training that increases their employability at the beginning of their labor market careers (*Ministério do Trabalho e Emprego*, 2009).

We use a very large restricted-access administrative dataset that has information on the whole history of formal jobs for millions of Brazilian workers: the *Relatório Anual de Informações Sociais (RAIS)*, collected by the Labor Ministry. RAIS is a longitudinal matched employee-employer dataset covering by law the universe of formally employed workers, including apprentices hired under the Apprenticeship program. The use of RAIS provides a rare opportunity to observe careers of young workers from the starting point in a developing country.[6]

A crucial issue for evaluation is to define what a successful youth-targeted training program is. The aim of this type of program is usually associated with achieving a better labor market integration of the young labor force (Biavaschi *et al.* 2012). The choice of the appropriate counterfactual depends, however, on the context of each program. In developing countries, for instance, informal and/or temporary jobs are a common first

---

[6] A recent review of the literature stresses that "…research into the effects of vocational training and related ALMPs would benefit enormously from the availability of better data and a suitable program design allowing for the proper evaluation of policy initiatives. Regarding data, the generation of representative survey data, in particular longitudinal data with a full set of individual characteristics, is essential" (Biavaschi *et al.* 2012). An example for a developed country is Adda *et al.* (2013) who uses an exceptionally rich data set based on social security records from 1975 until 2004 for male workers in the former West Germany.

step into the labor market. Given the low quality of schools, productivity signals of young people with low education are very imprecise and access to good formal jobs is often restricted to more educated workers. Most employers are thus reluctant to formalize contracts with young workers without previous experience and referrals from former employers.[7] This creates a vicious circle for these workers who do not get formal job offers because of no previous experience, which is hard to attain because of no job offers.

It has been argued that an informal first job in this context may help break this cycle by providing training and productivity signals to formal employers.[8] Cunningham and Salvagno (2011) show that young people in Argentina, Brazil and Mexico typically spend a short time in the informal sector after school before moving to longer spells in formal jobs. That is, having an informal first job seems to not preclude a later long-term attachment to the formal labor market. The same could also be expected for flexible arrangements such as fixed-term contracts that are common in countries with rigid labor legislation, especially for low-education workers. The crucial question is whether these types of entrance jobs provide young workers with the skills needed to make progress in their future careers.

Brazil is known for its rigid and imperfectly enforced labor legislation which results in a large number of informal jobs (Almeida and Carneiro, 2012; Gonzaga, 2003), as well as for the bad quality of its primary schools and the low level of education of most of its young workers. Therefore a youth-targeted program that focuses on providing professional skills in subsidized entry-level formal jobs seems to be an adequate ALMP in this context. If well designed, it has the potential to be an attractive substitute for other entry-level alternatives such as temporary or informal jobs.

The discussion above suggests defining a successful program as one that eventually leads to better paid and more stable jobs, compared to other temporary or informal jobs. Biavaschi *et al.* (2012) argue that "compared to fixed-term contracts, without training,

---

[7] Another evidence of this high uncertainty in hiring young workers in developing countries is the enormous level of turnover for this age group observed in Brazil (Corseuil *et al.*, 2013).

[8] This is a well-known phenomenon, but in a country like Brazil it usually results in equilibrium with large rates of school dropouts and very low formal employment rates. The low levels of education tend to perpetuate this problem, acting as a barrier to investment in training by eventual employers given its very low expected returns. The result is a labor market with a large number of workers trapped in low-paying jobs, mostly informal, with a very little chance of promotion and future real wage increases.

apprenticeships are better temporary contracts as they include systematic training and favourable prospects for subsequent job promotion, wages and employment stability".

We evaluate the Apprenticeship program by estimating its impact on some labor market outcomes of young workers. The treatment group is composed of young workers that started their careers in the formal sector as apprentices. Following the line of reasoning above we use as a control group workers of the same age that had other formal temporary contracts as first jobs in the same periods. Note that our dataset does not allow us to use workers with informal first jobs as a control group since we only observe workers with formal contracts.

We examine how the Apprenticeship Act program affects the career prospects of these young workers, in terms of degree of attachment to the formal labor market and wage growth. More specifically, we estimate the impacts of the program on: i) formal employment probability (overall and for non-temporary jobs); ii) measures of experience in the formal labor market (accumulated number of hours and months in formal sector jobs; probability of staying in the same firm or occupation); iii) measures of turnover (accumulated number of admissions and dismissals; probability of quits); and iv) real hourly wages. The analysis is carried on both for the short (2-3 years after the program) and the medium run (4-5 years after the program).

Since participation in the program is endogenous, the challenge is to deal with non-random selection based on unobservables as in most papers of the literature. Our identification strategy exploits a discontinuity by age in the eligibility to the Apprenticeship program. From 2000 to 2005, only individuals aged 14 to 17 years old could participate in the program. Individuals aged 18 years old or more were not eligible. This corresponds to the partially-fuzzy regression discontinuity setting discussed in Battistin and Rettore (2008), in which workers aged above a cutoff value cannot, and in fact do not, participate in the program, yet there is imperfect compliance for those below the cutoff.

We use three different estimators in the literature that exploit this design: the adjusted matching estimator proposed by Dias *et al.* (2013); the semi-parametric IV estimator introduced by Battistin and Rettore (2008); and a standard parametric IV (2SLS) estimator. All three estimators rely on strategies that allow the identification of a local parameter. In our context we identify the average impact of the program on the 17-years-old youths that choose (or are chosen) to enter the labor market as apprentices.

Identification is achieved combining information on those who enter the labor market aged 18 years old with those 17-years-old youths that choose (or are chosen) to enter the labor market in other temporary jobs.

We find that the program increases the chances of apprentices to get relatively better paid and more stable jobs, especially in the medium run – four and five years after the program. In particular, we find a very large impact on real wages that increase over time. We also find that the program is effective in increasing the probability of employment in a non-temporary job in the formal sector, especially in the medium run. The impact of the program on turnover is negative both in the short and in the medium run. On the other hand, we find a negative effect on accumulated formal labor market experience in the short run, which tends to vanish after 4-5 years.

The paper is organized as follows. In Section 2, we provide a brief literature review on youth-targeted programs. In Section 3, we provide some institutional background, and describe the Apprenticeship program and the data set used in the study. Section 4 discusses the estimation methods. Section 5 presents the main results. Section 6 concludes.

2. Literature Review on Youth-Targeted ALMP's

Our paper is more closely related to a strand of the literature which studies whether youth-targeted programs affect the career prospects of young workers, in terms of either wage growth or a higher degree of labor market attachment. There is a wide variation in evaluation methods in the literature with just a few experimental studies adopted in some countries. Card *et al.* (2010) and Kluve (2010) summarize the findings of the evaluation of several ALMP's in a large list of countries based on a meta-analytical framework. Both studies conclude that youth-targeted programs are less successful than other types of ALMP's.

Another important finding in the meta-analysis of Kluve (2010) is that ALMP's have different impacts depending on the time horizon studied. Many training programs, for instance, have positive effects only two or three years after implementation. This underlines the advantage of using data that allows one to follow workers for a long period after the intervention as we do in this paper.[9]

In another recent review of ALMP's, Urzua and Puentes (2010) present evidence that youth-targeted programs tend to have better results in Latin American countries than in developed countries.[10] This is consistent with a view that training programs should have more potential in low and middle-income countries since returns to skills are larger where skills are scarce.[11]

In fact, three recent studies that exploit randomized experiments in two Latin American countries (Colombia and Dominican Republic) find positive impacts of the programs on some youth labor market outcomes, although results for the first wave of the Dominican Republic program are either not significant or modest when significant (Attanasio *et al,* 2011; Card *et al.*, 2011; Ibarrarán *et al.* 2012).[12]

Attanasio *et al.* (2011) evaluate the youth-targeted training program *Jóvenes en Acción* which was introduced in Colombia between 2001 and 2005. They find sizable and significant impacts of the program for women on wages, formal wages, probability of employment, probability of formal employment, and hours worked. For men, the program only significantly affected the probability of formal employment and formal wages. They find very large effects on formal wages: 23% for men and 33% for women.

Card *et al.* (2011) also exploits a randomized experiment to evaluate a youth-targeted training program introduced in the Dominican Republic in the early 00's: *Juventud y Empleo* (JE). They find no significant impact of the program on employment and only modest impacts on wages and formality for men. They stress that the results could have been compromised by some flaws in the experiment design. Follow-up problems were observed with imperfect compliance and some crossover from control to treatment groups.

---

[9] Biavaschi *et al.* (2012) also stress the importance of using better data for evaluating youth training programs.

[10] Attanasio *et al.* (2011) argue that the introduction of youth-training programs in middle and low-income countries might have been discouraged by the mixed findings in the early literature for developed countries.

[11] Biavaschi *et al.* (2012) show that training provision is the primary form of ALMP in Latin America.

[12] Betcherman *et al.* (2007) review several impact evaluations of youth-targeted training programs in developing and developed countries. They also find that youth-training programs have on average more positive impacts in Latin America than in developed countries.

Ibarrarán *et al.* (2012) evaluate a modified version of the JE program in the Dominican Republic. The second wave of the program had a larger sample and improved on the follow-up design. The authors do not find a significant impact of JE on employment but estimate a 7% impact on wages. They also find positive impacts of the program on non-cognitive skills, such as leadership, conflict resolution, self-organization and persistency of effort; and show evidence that JE significantly reduced pregnancy rates.

An important trend to be noticed in the recent literature on evaluation of youth-targeted programs is the use of better micro data and modern microeconometric methods for program evaluation. We review below some studies to provide a flavor of what has been recently done in the literature.

Larsson (2003) uses propensity score matching methods to evaluate two youth programs (practice and training) in Sweden. He finds negative effects on earnings and employment one year after the intervention, with most coefficients becoming insignificant two years after the program.

De Giorgi (2005) uses a regression discontinuity design exploiting an eligibility rule to evaluate the New Deal for Young People (NDYP), a major youth-targeted intervention in the UK that combines different aspects of ALMP's (training, subsidized-employment, and job-search assistance). He finds that the program significantly increased employability of participants. Dorsett (2006) evaluates which of the different aspects of the NDYP program was most effective in reducing unemployment of participants. He finds that subsidized employment was the most effective means of exiting unemployment compared to the other options of NYDP.[13]

Centeno *et al.* (2008) evaluate a youth-targeted training program implemented in Portugal (*InserJovem*) in the late 1990s. They use a difference-in-difference estimation, exploiting the fact that the program was introduced only in some regions of the country, apparently for exogenous reasons. They find a negative impact: a very small effect of the program in *increasing* unemployment duration.

Finally, Caliendo *et al.* (2011) investigate the effectiveness of several youth-targeted programs implemented in Germany, based on a matching method (inverse probability weighting) applied to administrative data from 2002 to 2008. In general, they find

[13] See also Blundell *et al.* (2003) for an early evaluation of the NDYP program, in which a positive effect of the program on reducing unemployment in pilot areas was found.

positive effects of most of the programs evaluated on the employment probabilities of participants. Wage subsidies are found to have the largest effects in the long run.

As this brief literature review shows, despite the recent increase of studies using better data and more rigorous nonexperimental evaluations, evidence on the impacts of youth-targeted training programs is still limited, especially in developing countries.

3. Institutional Background and Data

3.1. Training Programs in Latin America

Vocational training has a long tradition in Latin America (Betcherman *et al.*, 2007; Biavaschi *et al.*, 2012). The first wave of training programs in the region started in the 1940s and was inspired by the German apprenticeship model. In fact, the first vocational training program in Latin America was implemented in 1942 with the creation of *Senai (Serviço Nacional de Aprendizagem Industrial*) in Brazil. In the following years, vocational training institutions (VTIs) were created in several Latin American countries. They usually had the explicit objective of providing skills in short supply to help the industrial sector to face the needs of an import-substitution development strategy. As Biavaschi *et al.* (2012) describe, these VTIs were "primarily supply-driven, state managed, financed through payroll taxes, independent from academic schools and from the Ministry of Education and usually quite close to the needs of the industry" (see also Moura Castro and Verdisco,1998).

The incompatibilities of these institutions to adapt to the economic structural changes Latin American countries faced in the 1980s and early 1990s resulted in a second phase of vocational training policies in the region. Pioneered by a program implemented in Chile in the early 1990s, many new training programs targeted to disadvantaged youth were created throughout the region, the so-called *Jóvenes Programs*. Unlike early VTIs, these programs are managed in a more decentralized way and "place a heavy emphasis on the private sector, both as a provider of training and as a demander of trainees" (Card *et al.*, 2011).

The Brazilian experience has been less studied in the literature. Only in 2005 Brazil created a youth-targeted program in the lines of the *Jóvenes Programs*, called

ProJovem.[14] In 2008, *ProJovem* was expanded and integrated with other similar programs.[15] On the other hand, in contrast with other Latin American countries, *Senai* has been able to somehow adapt to the new demand-driven challenges of the industrial sector and survived as the Brazilian main training institution. It is currently the largest educational network in Brazil.[16]

The implementation of the Apprentice Act in 2000 constituted the main youth-targeted ALMP in Brazil. The program shares some similarities with the *Jóvenes Programs* introduced in other Latin American countries. For instance, it involves many non-governmental organizations, philanthropy foundations, and private sector firms in several small-scale programs of training and placement of apprentices that are hired under the more general Apprenticeship program. On the other hand, *Senai* plays an active role in the program as the main provider of training.

In the next sub-section, we describe the program in more detail.

3.2. The Apprenticeship Program

Youth-targeted programs usually combine characteristics of several categories of ALMP's. The Brazilian Apprenticeship program is no exception to this rule. It is predominantly a professional training program. But it also has elements of other types of ALMP's. As described below, the program concedes employment subsidies through a reduction in payroll and firing costs. It also facilitates job search of participants, since it involves a network of formal sector firms that access data on apprentices.[17]

The Apprenticeship program has been part of the Brazilian labor legislation code *CLT* (*Consolidação das Leis Trabalhistas*) since 1943. However, it had a very small scope from 1943 to December 2000, when Law 11,180 - the Apprentice Act - was enacted. The program was initially designed for individuals aged 14 to 17 years old. It was

---

[14] See Gonzalez (2009) for an overview of youth-targeted labor market programs in Brazil.

[15] See Silva and Andrade (2009) for a detailed description of *ProJovem* and recent changes introduced in this program.

[16] *Senai* is financed by payroll taxes and run by the *National Confederation of Industry*. Since its creation, about 42 million students have enrolled in *Senai* training schools.

[17] Note that the design of the program focuses on training, which seems to matter most for a developing country with low levels of schooling. By contrast, an apparently successful UK young workers program, the *New Deal for Young People*, focuses on job search assistance and subsidized job placement (Blundell *et al*., 2003).

regulated in December 2005 by a more detailed legislation (*Decreto-Lei 5598*), when the maximum age for participation was increased from 17 to 23 years old. In 2010, around 200,000 workers had jobs under the program. The current Dilma Rousseff government works with a target of expanding it to one million young workers.

Young workers hired under the Apprentice Act program are required to take formal training courses outside the firm. Training courses are provided by official professional qualification agencies - the so-called *"S-System" (Senai, Senac, etc.)* - or by training institutions certified by the Labor Ministry. If an apprentice has not yet completed primary school (an eight-year schooling stage), she is required to enroll at school.[18]

The maximum number of working hours allowed for apprentices hired under the program is six hours per day for those still at primary school and eight hours per day for those with complete primary school. Payments must be at least the hourly minimum wage. There is a payroll subsidy in the form of a lower requirement of deposit on the worker's FGTS account (*Fundo de Garantia por Tempo de Serviço*, a job-separation fund). Firms should deposit only 2% of the basic monthly wage on this fund, instead of the rate of either 8% or 8.5% that prevailed for other workers during that period.[19]

Apprentices are hired under non-renewable fixed-term contracts with a maximum length of two years. As in other fixed-term contracts, there are no firing costs for job separations by the end of the contract.[20] This is one of the main benefits for firms to use temporary contracts, since the standard procedure in cases of unjustified separations induced by firms is to pay a fine equivalent to 40% or 50% of the accumulated amount deposited in the FGTS account during the employment relationship.[21]

---

[18] Despite the concern on school enrollment, the program has been criticized for including individuals with less than 16 years of age, which is prohibited by law in all other forms of labor contracts.

[19] Firms had to deposit 8% of the monthly wage on the worker's FGTS account from 1966 to October 2001, when the government introduced a temporary increase of 0.5 percentage points which lasted for five years (Gonzaga, 2003).

[20] Contracts must be terminated when the apprentice reaches the age limit (18 years old between 2000 and 2005, and 24 after 2005). The end of contract can be also anticipated in some exceptional cases, that include a non-adaptation or an insufficient performance of the apprentice in the training courses as attested by the certified training institutions (*Ministerio do Trabalho e Emprego*, 2006).

[21] The fine (to be paid to the worker) was 40% of the FGTS balance until October 2001, when it was permanently increased to 50%, with the additional 10 percentage points to be paid to the government. Since the FGTS fund approximately accumulates at the rate of one monthly wage per year, firing costs are around 50% of one monthly wage for each year of tenure. Almost all (more than 99%) firm-induced separations in Brazil are for unjustified reasons.

Firms' choices regarding the use of apprenticeship contracts are restricted by the following rule. A minimum of 5% (and a maximum of 15%) of the labor force employed in occupations requiring formal training should be composed of apprentices.[22] The inspection division of the Ministry of Labor is in charge of enforcement. Enforcement, however, is very low, especially in the early 2000s when firms could claim a lack of training agencies in the region/occupation they operate so as not to be penalized for employing less than the minimum amount required. Therefore, in practice the minimum threshold requirement was not binding in the period we analyze in this study. In particular, small firms tend not to hire workers under the program.[23]

3.3. Data Description

In this paper we use a very large restricted-access administrative file maintained by the Brazilian Ministry of Employment and Labor (*Ministério do Trabalho e Emprego*), the *Relação Anual de Informações Sociais* (RAIS). RAIS is a longitudinal matched employee-employer dataset covering by law the universe of formally employed workers in Brazil. All tax-registered firms have to report every worker formally employed at some point during the previous calendar year.[24] Apart from tax/social security compliance the data has no coverage limitation, as opposed to other similar databases that are limited by geographical region, size, or industry. We use data from 2001 to 2008. Over this period RAIS contains an average of 40 million worker-establishment records per year.

Firm and worker identification numbers provide a natural way to construct a matched employer-employee longitudinal dataset. With the establishment identification number (CNPJ) it is possible to follow all establishments that file RAIS over time. With the worker's national insurance number (PIS), it is possible to follow all workers that remain in the formal sector over time and to construct a panel of all matched establishment-worker pairs.

---

[22] The list of occupations requiring formal training can be found in the Ministry of Labor website at http://www.mtecbo.gov.br/cbosite/pages/home.jsf.

[23] Note that small firms tend to be overrepresented in remote places with lower supply of training agencies and lower enforcement of labor legislation.

[24] There are incentives for truthful reporting since the main purpose of RAIS is to administer a federal wage supplement (*Abono Salarial*) to formal workers.

Data on worker characteristics (age, education, gender) and establishment characteristics (industry, location at the municipality level) are available as well as detailed information for each employee-employer contract, such as wage, hours, tenure, month of admission, month of separation, reason of separation, occupation, type of contract (permanent or temporary, including whether it was an apprenticeship contract). There are two measures of wages: the average value over the year (or over the period of the year that the worker was with the firm) and the December wage.

In order to exploit the age discontinuity of eligibility rules, we restrict the sample to workers who had their first jobs in the formal labor market at the ages of 17 or 18 years old in each of the first three years after the implementation of the Apprentice Act (from 2001 to 2003). We only keep information for those youths that were hired for a fixed-term (temporary) job. Apprentices hired under the Apprenticeship program constitute the treatment group, while other temporary workers are in the control group. In total, we have information on 11,366 apprentices (treatment) and 32,806 non-apprentices (control) that had their first jobs at the ages of 17 and 18 between 2001 and 2003.

We follow all workers in our sample for five years (in addition to the entrance year). This allows us to compute average program impacts for the short run (arbitrarily defined as 1 to 3 years after the first formal job) and the medium run (4-5 years after first formal job). We find each worker in the sample in all formal (temporary and non-temporary) jobs in subsequent years and keep all information for each matched employee-employer pair.

As in any study relying on longitudinal data, attrition is a crucial issue for our analysis. On average, RAIS' attrition rate in any two consecutive years from our sample period is approximately 5%.[25] One of the main sources of attrition in RAIS is due to occasional non-reporting by complier establishments. We identified several cases in which all employees from some establishments "disappear" from RAIS in a particular year and eventually return in subsequent years. We exclude these episodes of spurious establishments "births" and "deaths" from our sample.

[25] Attrition rate is defined as the share of workers who are not found in a given year despite having been registered as employed on the last day of the previous year.

3.4. Some evidence on scope, enforcement and compliance with the Apprentice program

Table 1 presents the actual number of workers registered in RAIS as apprentices for all years from 1998 to 2010. The table reveals that: i) the number of apprentices substantially increased throughout the 2000s; ii) the majority of apprentices are 16 and 17 year olds; iii) the age threshold of 18 years was respected between 2000 and 2005; and iv) the discontinuity at 24 years old after 2005 is not relevant with only a small number of apprentices aged 23 until 2008.[26]

Table 1: Number of Apprentices by Age

Number of Apprentices by Age

Age	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
14	215	82	99	143	582	803	937	1291	1497	2125	2242	2369	2918
15	1063	984	724	1061	2781	5279	7387	8962	11656	13228	16115	16252	20907
16	1616	3145	2917	2684	5747	12365	20341	26060	29776	35100	39100	41787	50723
17	1262	3182	3617	2156	4547	9117	16266	21122	30160	34376	38871	42393	50654
18	70	34	54	76	48	79	121	1392	7052	12972	16317	21968	27846
19	27	1	0	0	0	0	0	246	2521	6687	9565	13145	17758
20	22	0	0	0	0	0	0	125	1396	3560	5804	8250	10464
21	16	0	0	0	0	0	0	72	832	1970	3307	4902	6470
22	4	0	0	0	0	0	0	63	403	1125	1793	2621	3366
23	0	0	0	0	0	0	0	32	193	439	674	1057	1320
24	3	0	0	0	0	0	0	2	6	25	31	51	77
25	2	0	0	0	0	0	0	0	0	0	10	28	35
26	4	0	0	0	0	0	0	0	0	0	9	24	27
27	2	0	0	0	0	0	0	0	0	0	6	30	43
28	1	0	0	0	0	0	0	0	0	1	6	17	25
29	1	0	0	0	0	0	0	0	0	0	9	23	28
30	0	0	0	0	0	0	0	0	0	0	3	18	22
Total	4295	7428	7411	6120	13705	27643	45052	59365	85486	111582	133788	154744	192426

Source: Constructed by the authors based on microdata from RAIS.

Although the number of workers employed as apprentices has grown steadily over the 2000s, compliance was still very limited by the end of the decade. In 2010 only 1.5% of the establishments employed a number of apprentices in accordance with the minimum required by the Apprentice Act. Table 2 reports the shares of establishments, by establishment size, employing proportions of apprentices in the following ranges: no apprentices; more than zero but less than 5% of employees; between 5% and 15% (legal amount); more than 15%.[27] The first line of the table shows that the share of small establishments (7 to 20 employees) employing no workers as apprentices is very high,

[26] An informal conversation with an inspector from the Labor Ministry confirmed our prior that an apprenticeship job at age 23 is no longer attractive, which explains the declining number of apprentices as they approach that age.

[27] All quantities in Table 2 refer to the set of occupations requiring formal training, including the establishment sizes categories.

reaching 98.7%.[28] This share decreases monotonically as establishment size increases, reaching 74.1% for the group of largest establishments (more than 500 employees). This confirms the claim of limited enforcement, in particular for smaller establishments. The third column of the table shows that the share of establishments employing the legal amount of apprentices varies from 0.9% to 3%, averaging 1.5% as previously mentioned.

Table 2: Shares of establishments by proportion of apprentices and establishment size

	no apprentices	0% - 5%	5% - 15%	> 15%
7 - 20 employees	98.7%	0.1%	1.2%	0.1%
21 - 50 employees	93.9%	4.1%	2.0%	0.1%
51 - 100 employees	84.9%	11.8%	3.0%	0.2%
101 - 500 employees	75.5%	21.4%	2.6%	0.4%
501 + employees	74.1%	24.5%	0.9%	0.5%
Average	**95.6%**	**2.8%**	**1.5%**	**0.1%**

Source: Constructed by the authors based on microdata from RAIS.

Table 3 provides some information on enforcement from 1998 to 2011 provided by the Brazilian Ministry of Labor (MTE). The first column of the table shows the number of workers with labor contracts regularized after labor inspections conducted by MTE. One can see an upward trend until 2007 with a mild decrease afterwards. The second column displays the subset of labor contracts which were regularized as apprenticeship contracts. In contrast to the previous series, the number of regularized contracts of apprentices was very close to zero in the early 2000's and has monotonically increased throughout the period. As a result the share of apprentices in labor contracts regularized due to inspections rose steadily from less than 0.5% in 2001 to almost 25% in 2011 (see last column).[29]

[28] Establishments employing less than six employees in the set of occupations requiring formal training were excluded from our calculations. The reason is that for this group of establishments one apprentice would correspond to more than 15% of their employees. So in order to comply with the law the establishments in this group can't employ any apprentice.

[29] Table 3 also shows that before 2000 the requirement to hire apprentices was not binding. This confirms our claim that despite the availability of the apprenticeship contract since 1943, the use of such contract only became a reality after the enactment of the Apprentice Act in 2000.

Table 3: Labor contracts regularized due to labor inspections

Year	Total	Apprentices	%
1998	261,274	-	-
1999	249,795	-	-
2000	525,253	850	0.2
2001	516,548	1,919	0.4
2002	555,454	11,111	2.0
2003	534,125	18,146	3.4
2004	708,957	25,215	3.6
2005	746,272	29,605	4.0
2006	670,035	44,049	6.6
2007	746,245	52,676	7.1
2008	668,857	55,637	8.3
2009	588,680	68,926	11.7
2010	515376	87,823	17.0
2011	480,423	118,164	24.6

Source: Department of Inspection, Labor Ministry (MTE)

Summing up, this sub-section shows that the scope and enforcement of the Apprenticeship program are limited but are both rising throughout the years. The limited scope in the early 2000's is beneficial for our evaluation purposes as general equilibrium issues were probably not relevant then. The rising in scope and enforcement of the Apprentice Act underscores the importance of a rigorous impact evaluation as an increasing amount of resources is being devoted to this program.

4. Methodology and estimation procedures

The main objective of this paper is to estimate the impact of the Apprenticeship program on labor market outcomes. As in many other similar studies, the main challenge in the absence of a randomized experiment is to separate causal effects from selection based on unobservables. In other words, the impact of the Apprenticeship program on youth employability is not trivially identified, since selection into the program is defined by firms and workers, and hence may be driven by unobservable characteristics. If these unobservable characteristics are not balanced among treated and non-treated workers, then methods relying solely on the comparison of the outcomes between the two groups produce misleading estimates.

In order to get consistent estimates of the effect of the Apprenticeship program we make use of a set of three somewhat related estimators. In all three cases we exploit the fact that the eligibility to the program switches as age crosses a threshold value.

The first is an estimator recently proposed by Dias *et al.* (2013), which combines the idea of matching on observables with exogenous variation provided by an instrument. The second is a semi-parametric version of the IV estimator applied to the context of a partially-fuzzy design as discussed by Battistin and Rettore (2008). The third is the traditional IV estimator, or 2SLS, also applied in a fuzzy design as discussed in Hahn *et al.* (2001).[30]

We are able to identify and estimate a version of the ATT parameter regardless of the procedure we choose. This is the case even when using the IV estimators, which is usually associated to the LATE parameter in program evaluation. The reason is that by design those above the age-threshold cannot and do not participate in the program. In this situation the group of always-takers does not exist, implying that the treated group coincides with the complier group, the one for whom the effect is identified in the LATE parameter.[31] In what follows we describe these estimators.

4.1. Adjusted Matching

In its ideal setting, the Dias *et al*. (2013) estimator uses an instrument which exploits boundary restrictions on eligibility rules based on individual characteristics (e.g., age, education, income). In this context, the instrument should drive participation into the program to zero for certain values of its domain and at the same time allows partial compliance for other values.[32] The idea is that by moving individuals in and out of the

[30] Hahn *et al.* (2001) relates the set of identification conditions in this context with those prevailing for the estimation of the LATE parameter, which in turn was proposed by Angrist and Imbens (1994). A summary on these topics can be found in Angrist and Pischke (2009).

[31] It is worth noting that in the setting of regression discontinuity design, as in the fuzzy designs we are dealing with in two of our estimators, identification is still local not because of the restriction to compliers but because of the validity around an age threshold. So we end up estimating a parameter that may be called a local average treatment on the treated (LATT).

[32] The Dias *et al.* (2013) approach is related to the partially-fuzzy regression discontinuity design proposed by Battistin and Rettore (2008). One of the main differences between the two approaches is that the former explicitly requires an exclusion restriction in the form of an instrument, while the latter is

program the variation in the instrument can correct for possible unbalances in unobservables due to self-selection into the program. Note that the standard matching (on observables) method does not take care of such unbalance.

To be more formal, we are interested in estimating the Average Treatment on the Treated (ATT) parameter:

$$\alpha = E[Y1|D = 1] - E[Y0|D = 1] = E_{X|D=1}E[Y1|X, D = 1] - E_{X|D=1}E[Y0|X, D = 1],$$

where Y1 and Y0 represent individual potential outcomes associated with assignment to treatment and non-treatment, respectively, D measures the actual treatment status, with D = 1 (D = 0) corresponding to actual participation (non-participation) in the program, and X is a vector of conditioning covariates. The notation $E_{X|D=1}$ means expectation over the X distribution for the D = 1 population.

The object $E_{X|D=1}E[Y1|X, D = 1]$ can be directly computed from the data through the mean of the outcome of interest among the treated group. However, as usual, the counterfactual object $E_{X|D=1}E[Y0|X, D = 1]$ is not directly available in the data, so it needs to be identified through the use of some assumptions. Dias *et al.* (2010) propose an estimator of the counterfactual object based on the existence of a variable Z for which two features are assumed to apply:

A1: $Y0 \perp Z|X$;

A2: There exists a set of points $\{z^*, z^{**}\}$ in the domain of Z where for all X:

$P\left[D = 1\middle|X,\, Z = z^*\right] = 0$ and $0 < P\left[D = 0\middle|X,\, Z = z^{**}\right] < 1$.

The first assumption is an exclusion restriction that imposes that the variable Z is not correlated with the counterfactual outcome Y0 after conditioning on the covariates in X.[33] Assumption 2 requires the existence of at least one value of Z that is capable of driving participation into the program to zero and at least another value for which participation is non-deterministic. It is interesting to note that these assumptions do not

based on an assumption of continuity near the cutoff point. Also, while the latter approach is cast in terms of the LATE (Local Average Treatment Effect), the former is cast in terms of the ATT (Average Treatment on the Treated).

[33] In fact, that condition could be stated in terms of mean (conditional) independence.

impose that there is no selective participation into the program. Indeed, they allow D to be correlated with Y0 when Z takes on the value z** (after conditioning on X).

Using A1 and A2, Dias *et al.* (2013) propose a constructive proof for the identification of the mean counterfactual outcome $E[Y0|X, D = 1]$.[34] They show that this object can be written as

$$E[Y0|X, D = 1] = E[Y0|X, D = 0] + \frac{E[Y0 \,|\, X, Z = z^{*}, D = 0] - E[Y0|X, D = 0]}{1 - P[D = 0|X]}$$

This expression shows that $E[Y0|X, D = 1]$ is equal to the mean outcome $E[Y0|X, D = 0]$, typically computed in matching estimation, plus what the authors call a correction term, which is given by the second term in the right hand side (RHS) of the equation. Note that all elements that compose this second term can be identified from the data, where $E[Y0 \,|\, X, Z = z^{*}, D = 0]$ is the mean observed outcome for ineligibles controls at given X and $\{1 - P[D = 0|X]\}$ is the propensity score. The object of interest $E[Y0|D = 1]$ is finally identified from $E[Y0|X, D = 1]$ by averaging the latter over the distribution of X for the treated group (D = 1).

We implement this estimator using age as the Z variable. This choice fits well in the ideal setting for the application of the Dias *et al.* (2013) estimator, since the eligibility rules of the Apprenticeship program impose a restriction on the maximum age for participation. As described in Section 4, the maximum age to participate in the program was 17 years old until September 2005, when the age restriction rose to 23. Recalling that the program is not compulsory, we have thus an appropriate setting in which the age of workers can be used as the instrument: while those aged above the cutoff value cannot participate, there is imperfect compliance for those below the cutoff.

Estimation results will be presented for both the standard propensity score matching estimator and the so-called adjusted matching estimator (Dias *et al.*, 2013). The covariates in X we use in the propensity score are dummies for gender, schooling, industry, occupation, geographical region, and the year in which the worker first entered the formal sector. The standard matching estimates were computed using Epanechnikov kernel weights. Only observations in the region of the common support of the

[34] The proof can be found in Appendix 1.

propensity score were used for computing standard and adjusted matching estimates.[35] Inference is based on standard errors estimated from bootstrap with 100 replications.

4.2 Semi-parametric IV

The age cutoff condition for eligibility in the Apprenticeship program fits directly into a framework of regression discontinuity design (RDD). In particular, it fits well the framework put forward by Battistin and Rettore (2008), where on one side of the eligibility threshold individuals are precluded from participating, while on the other side eligible individuals may self-select into the program. In fact, the main idea behind their estimator exploits the imposition of the non-participation condition near the threshold for eligibility to solve the selection problem. In our context, this implies that those aged 18 years old will help identifying the average impact of the program on the 17 years old youths that choose (or are chosen) to become apprentices in the formal labor market. The fact that their framework is based on a design where on one side of the cutoff point there are ineligibles and on the other side there are eligible participants and eligible non-participants configures what the authors call a partially fuzzy design.[36]

To see how Battistin and Rettore's (2008) estimator operates, let program eligibility be defined by an observable, continuous variable Z[37] (age in our case). Let $\bar{z}$ be the value of Z that defines the eligibility threshold, that is, the discontinuity point in the domain of Z below which individuals can participate in the program. Let $\bar{z}^{-}$ and $\bar{z}^{+}$ refer to the groups of individuals that are marginally below and above the cutoff point of eligibility, respectively. In our estimation, they are represented by workers with 17 and 18 years old, respectively.

---

[35] Since the denominator of the correction term of the adjusted matching estimator is the estimated propensity score, estimates of the correction term can become quite imprecise for low values of the propensity score. Hence, following a suggestion in Dias *et al.* (2013), we asymmetrically trimmed the common support interval to be between the maximum of the 5th percentiles and the minimum of the 99th percentiles of the propensity score distributions of the treated and control groups.

[36] Typically in the RDD literature there are two types of designs: i) the sharp, where the probability of participation in the program changes from zero to one as the value of the eligibility variable crosses the threshold; and ii) the fuzzy design, where the change in the participation probability is less than one at the discontinuity threshold. The partially-fuzzy design combines features of these two designs. Classic recent references in the RDD literature are Hahn *et al.* (2001) and van der Klaauw (2002).

[37] In their paper this variable is denoted by "s". We use "z" to be consistent with the notation in the previous section.

Using the notation from the previous section, if $\alpha = Y1 - Y0$ denotes the impact of the intervention, our interest centers in identifying the average treatment on the treated effect (ATT): $E[\alpha|D = 1]$, where $D = 1$ denotes program participation. Using the usual counterfactual notation, the observed outcome of any individual in the population can be written as

$$Y = Y0 + D(z)\alpha,$$

where $D(z)$ is an indicator function of treatment status which explicitly recognizes that it depends on the variable Z.

Consider the difference in mean outcomes $E[Y|\bar{z}^-] - E[Y|\bar{z}^+]$. Using the previous expression, this difference can be rewritten as

$$E[Y|\bar{z}^-] - E[Y|\bar{z}^+] = E[Y0|\bar{z}^-] - E[Y0|\bar{z}^+] + E[D(s)\alpha|\bar{z}^-] - E[D(s)\alpha|\bar{z}^+].$$

By construction of the program design, those that are marginally ineligibles cannot participate (those who are 18 years old in our context). Hence, $D(\bar{z}^+) = 0$ and the previous expression becomes:

$$E[Y|\bar{z}^-] - E[Y|\bar{z}^+] = E[Y0|\bar{z}^-] - E[Y0|\bar{z}^+] + E[D(z)\alpha|\bar{z}^-].$$

The only condition needed to identify a local version of the parameter of interest is:

C1: $E[Y0|Z]$ is a continuous function of Z at $\bar{z}$.

This assumption, which is the main condition for identification of the mean impact of treatment for those at $\bar{z}^-$ in the usual sharp RDD, simply requires that there is no discontinuity in counterfactual outcomes at the threshold for eligibility. It is typically considered a weak condition.

Noting that $E[D(z)\alpha|D = 0, \bar{z}^-] = 0$, we can write the last term of the previous expression simply as $E[D(z)\alpha|\bar{z}^-] = E[\alpha|D = 1, \bar{z}^-].P[D = 1|\bar{z}^-]$. Now, by condition C1, $E[Y0|\bar{z}^-] = E[Y0|\bar{z}^+]$, so the parameter of interest can be locally identified for individuals with $Z = \bar{z}^-$ (the group of 17 years old) by

$$E[\alpha|D = 1, \bar{z}^-] = \frac{E[Y|\bar{z}^-] - E[Y|\bar{z}^+]}{P[D=1|\bar{z}^-]}$$

Notice that all objects in the RHS of this expression can be computed from the data. In particular, the denominator can be seen as the propensity score for participation for those marginally eligible. In practice, it is calculated for this group using the same propensity score that was estimated for the adjusted-matching estimator of the previous section.[38] For comparison purposes, inference is based on the same 100 bootstrap replications that were used in the computation of the adjusted-matching estimator. We also compute the partially-fuzzy estimator using the same common support of each replication of the adjusted matching estimator.

4.3 Parametric IV

As P[D=1|$\bar{z}^{+}$] = 0 in our context, expression (1) above can be re-written as:

$$E[\alpha|D=1,\bar{z}^{-}] = \frac{E[Y|\bar{z}^{-}] - E[Y|\bar{z}^{+}]}{P[D=1|\bar{z}^{-}]} = \frac{E[Y|\bar{z}^{-}] - E[Y|\bar{z}^{+}]}{P[D=1|\bar{z}^{-}] - P[D=1|\bar{z}^{+}]}$$

The last term is the traditional formula for the fuzzy regression discontinuity identification strategy, which in turn motivates the use of 2SLS estimation procedures by applied economists. Therefore we also use this estimator for the sake of comparability with a standard framework to deal with self-selection issues. We apply it in a fully parametric 2SLS framework, where a dummy for being 17 years old is used as the instrument for the apprentice's treatment dummy. Note that this identification strategy could be applied even for a complete fuzzy design, whereas the two previous strategies rely on the partial fuzzy design formally expressed by assumption A.2 in section 5.1.

[38] *Mutatis mutandis,* all objects presented in this section could be conditioned on the vector of observable characteristics X without changing the essence of the identification of the object of interest.

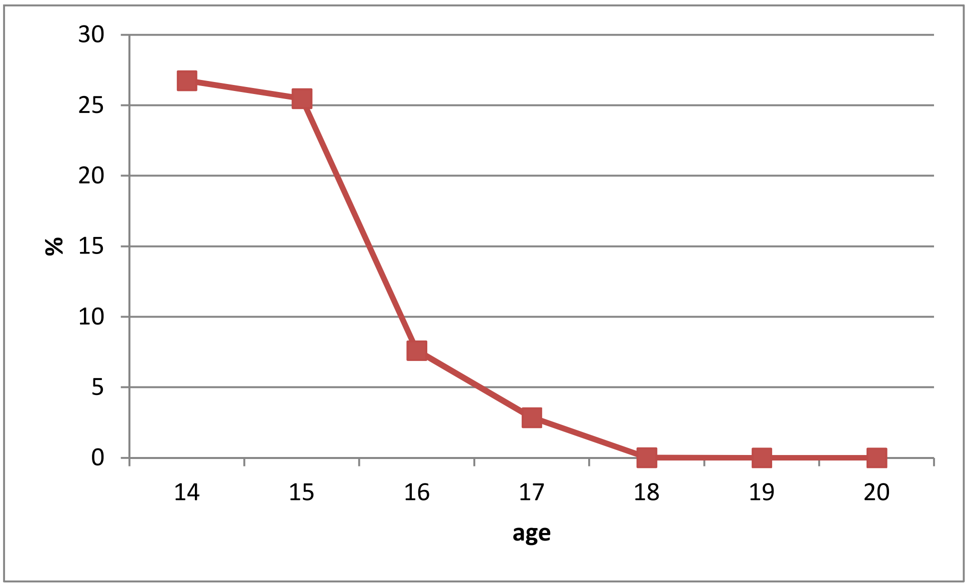
5. Descriptive Statistics and Econometric Results

In this section we show the results of the estimation of the impact of the Apprenticeship program on several labor market outcomes derived from applying the three identification strategies described in the previous section. Before turning to the results we discuss the plausibility of two important assumptions that permeate the identification strategies using some useful descriptive statistics.

5.1. Descriptive Statistics

Two of the three methods we use in the paper are based on the partial participation of youths under 17 years old and the non-participation of youths over 18 years old. To confirm this, Figure 3 shows the participation rate in the Apprenticeship program by age for the period 2001-2003. The figure reveals that, although the probability of participation declines for eligibles, it is always positive below the 17 years old cutoff and becomes virtually zero for youths older than this threshold. Since the estimators we use are local, we only used information on youths aged 17 and 18 in all estimations.

Figure 3: Participation rate in the Apprentice's program by age – 2001/2003



Source: Constructed by the authors based on microdata from RAIS.

Another important identifying assumption relies on the comparison of unobservable characteristics between the 18-years-old and the 17-years-old groups. The precise statement of the assumption varies according to the method but one way or another they require some sort of similarity in this comparison, which can also be stated in terms of the outcome variable in the absence of the program. In what follows we refer to this assumption as the exclusion restriction. This sort of comparison cannot be implemented with either variable. Some indirect evidence based on comparisons implemented with observable variables is usually provided by applied economists using such type of methods.

Table 4 displays how observable characteristics are balanced among alternative groups for all temporary workers that had their first jobs at the ages of 17 or 18 years old between 2001 and 2003. Our sample has information on 11,366 apprentices and 32,806 non-apprentices.

We split the 17 years group in two sub-groups: the 11,366 workers hired under an apprenticeship contract and those hired under another type of temporary contract (10,138 workers). The results for these two sub-groups are shown in the first two columns of Table 4. The last two columns of the table compare the 17 years old group with the 18 years old (22,668 workers). A good balance for observed characteristics across groups would support our identifying assumption.

The first row of the table shows that gender is balanced across groups, with a proportion of around two thirds of males in the sample. This equivalence between groups is not observed for the other characteristics reported in the remaining rows. We note, however, that in some cases differences may be induced by the program.

Schooling distributions are very different when one compares the 18- and 17-years old groups. Differences in the bottom part of the schooling distribution, though, seem to be induced by the program, as the shares of non-apprentices in the first two schooling categories are very close to those registered for workers aged 18. Nonetheless, an important difference remains in the top part of the schooling distribution, as can be attested by the shares in the last two schooling categories: incomplete secondary school and complete secondary school. This is probably explained by the fact that the completion of secondary schooling in Brazil tends to occur in the students' eighteenth year of life.

Table 4 – Observable Characteristics: Temporary Workers, 1st Job at age 17 or 18

Characteristics	Age 17 Apprentices	Age 17 Nonapprentices	Age 17 - All	Age 18 - All
Male	0.64	0.67	0.66	0.67
Schooling				
Less than 5	0.00	0.15	0.07	0.14
6 to 8	0.11	0.22	0.16	0.20
9 to 11	0.84	0.53	0.70	0.33
More than 12	0.04	0.10	0.07	0.33
Industry				
Agriculture	0.00	0.23	0.11	0.20
Construction	0.01	0.01	0.01	0.03
Manufacturing	0.06	0.05	0.05	0.03
Trade	0.17	0.10	0.14	0.06
Services	0.62	0.57	0.59	0.65
Occupation				
0.Armed forces	0.00	0.00	0.00	0.01
1.Managers (public & private)	0.01	0.02	0.02	0.01
2.Professionals	0.01	0.00	0.01	0.00
3.Technicians	0.28	0.16	0.22	0.14
4.Clerical	0.39	0.18	0.29	0.12
5.Service and sales	0.04	0.13	0.08	0.11
6.Agricultural	0.00	0.30	0.14	0.26
7.Craft	0.09	0.09	0.09	0.13
8.Machine operators	0.08	0.04	0.06	0.04
9.Elementary occupations	0.09	0.08	0.08	0.17
Region				
North	0.04	0.03	0.03	0.04
Northeast	0.10	0.10	0.10	0.16
Southeast	0.70	0.66	0.68	0.60
South	0.11	0.16	0.13	0.14
Mid-West	0.06	0.05	0.06	0.07
Sample size	11,366	10,138	21,504	22,668

Source: Constructed by the authors based on microdata from RAIS.

The table also shows that the apprenticeship program is concentrated in non-agriculture activities. This is probably related to the logistics required for the implementation of an apprenticeship contract. Long commuting between workplaces and training centers can make this type of contract prohibitive for young workers in rural areas. It should be noted that a higher share of workers in the service activity within the 18 years old group persists even after comparing with the non-apprentices group.

Occupational distributions are also not well balanced between groups. Apprentices are more concentrated in clerical and technical occupations compared to other forms of temporary contract. This is also the case when one compares the 17 and 18 year old groups. Finally, the table shows that the regional distribution of workers aged 18 seems to be more concentrated on the Northeast and less on the Southeast than for those aged

17. In this case the difference does not seem to be induced by the program, with similar numbers observed for apprentices and non-apprentices.

Overall it seems fair to say that we cannot reject the hypothesis that in the absence of the program workers aged 17 and 18 would be similar. Although the table shows some important differences on observable characteristics between these workers, most of them seem to be induced by the program. This consideration reinforces the need for a method that takes into account a non-random selection of individuals into the program.

The exclusion restriction deserves some further consideration. We want to stress that the sample used in the regression analysis is restricted to youth entering the labor market for the first time and hired under a temporary contract. We think the first restriction minimizes concerns of selection induced by employers since little (or nothing) is known about worker productivity except the characteristics that we are able to control for. Moreover even if you assume that employers are able to extract relevant information that may induce better opportunities for one or another group of workers, the last restriction tends to homogenize these opportunities, as everyone in the sample was hired under the same type of contract.

Finally, one may argue that the 18th anniversary introduces a discontinuity in employability since individuals take more responsibilities at this age.[39] However it should be stressed that we are comparing individual's employability two to five years after the entrance year. Therefore we should expect that everyone in our sample would already have incorporated any discontinuous jump in employability experienced when they turned 18.[40]

5.2. Econometric Results

In this sub-section we present our estimates of the effects of the Apprenticeship program on several labor market outcomes, such as wage growth and measures of the degree of attachment to the formal labor market in subsequent years following the treatment for all three estimation procedures described in Section 4.

The outcomes of interest can be classified in four groups of variables: i) formal employment probability (overall and for non-temporary jobs); ii) measures of

---

[39] In Brazil 18 years old is a threshold defining criminal responsibilities and permission to drive vehicles.

[40] The outcome variables are compared when the youth who entered the labor market with 17 years old is about 19 to 22 years old.

experience in the formal labor market (accumulated number of hours and months in formal sector jobs; probability of staying in the same firm or occupation); iii) measures of turnover (accumulated number of admissions and dismissals; probability of quits); and iv) real hourly wages. All impacts are estimated for the short run (2-3 years after the program) and the medium run (4-5 years after the program).

Before presenting the econometric results, Table 5 displays the averages of the outcomes for the sub-groups in our sample analyzed in Table 4: apprentices; non-apprentices age 17; 17 year olds; and 18 year olds. The last column presents the averages of all variables for the non-treated group of non-apprentices.

Table 5 – Outcomes: Temporary Workers, 1st Job at age 17 or 18

	Age 17				
Outcomes	**Apprentices**	**Non-apprentices**	**Age 17**	**Age 18**	**Non-treated**
Employment					
Employment probability in years t+2 or t+3	0.72	0.73	0.72	0.77	0.76
Employment probability in years t+4 or t+5	0.74	0.72	0.73	0.74	0.73
Employment probability under a non-temporary job in years t+2 or t+3	0.68	0.55	0.62	0.57	0.56
Employment probability under a non-temporary job in years t+4 or t+5	0.71	0.58	0.65	0.58	0.58
Experience					
Accumulated number of hours worked in years t+2 and t+3	1867	1749	1812	2026	1940
Accumulated number of hours worked in years t+4 and t+5	1740	1761	1750	1854	1825
Accumulated number of months worked in years t+2 and t+3	10.86	10.29	10.59	11.86	11.37
Accumulated number of months worked in years t+4 and t+5	10.03	10.24	10.13	10.80	10.63
Probability of staying in the same establishment in years t+2 or t+3	0.08	0.10	0.09	0.13	0.12
Probability of staying in the same occupation in years t+2 or t+3	0.07	0.08	0.07	0.07	0.07
Turnover					
Accumulated number of dismissals between years t+2 and t+3	0.59	0.81	0.69	0.88	0.86
Accumulated number of dismissals between years t+4 and t+5	0.53	0.75	0.63	0.75	0.75
Accumulated number of admissions between years t+2 and t+3	0.83	0.95	0.89	0.94	0.94
Accumulated number of admissions between years t+4 and t+5	0.58	0.79	0.68	0.79	0.79
Probability of dismissal by quit in years t+2 or t+3	0.22	0.26	0.24	0.23	0.24
Probability of dismissal by quit in years t+4 or t+5	0.21	0.27	0.24	0.24	0.25
Wages					
Real hourly wage in year t	1.95	2.25	2.09	3.05	2.80
Real hourly wage in year t+1	2.28	2.67	2.45	3.40	3.17
Real hourly wage in year t+2	3.34	3.06	3.21	3.63	3.45
Real hourly wage in year t+3	3.82	3.52	3.68	3.94	3.81
Real hourly wage in year t+4	4.35	3.86	4.13	4.39	4.23
Real hourly wage in year t+5	4.89	4.28	4.55	4.78	4.63
Sample size	11,366	10,138	21,504	22,668	

Raw comparisons of outcomes between apprentices (first column) and non-apprentices (last column) reveal that apprentices tend to have:

- similar probabilities of being employed in a formal job both in the short and in the medium run (slightly less in the short run);
- larger probabilities of being employed in a non-temporary formal job both in the short and in the medium run;
- less accumulated hours and months worked in formal jobs both in the short and in the medium run;

- a smaller probability of staying in the same firm in the short run, but rates are very small for both groups (8% for apprentices and 12% for non-apprentices);
- a similar and very small probability (7%) of staying in the same occupation in the short run;
- much lower turnover, with smaller accumulated numbers of dismissals and admissions both in the short and in the medium run;
- a slightly lower probability of quitting;
- a larger increase in real hourly wages over time: wages of apprentices were 30% lower than of non-apprentices in year *t* but 6% larger by year *t+5*.

Table 6 presents the estimation results for the average treatment effect on the treated parameter for the three estimation procedures discussed in Section 4. For comparison purposes, the first two columns display, respectively, the simple differences in outcome variables between treatment and control groups as reported in Table 5, and standard matching estimates based on propensity score.

Columns 3, 4 and 5 present, respectively, the adjusted matching estimate proposed by Dias *et al.* (2013), the partially-fuzzy estimate proposed by Battistin and Rettore (2008), and a standard IV coefficient. The covariates used in the propensity score are dummies for gender, schooling, industry, occupation, geographical region, and the year in which the worker first entered the formal sector.[41]

---

[41] We have also computed the correction term of the adjusted matching method, with its respective standard errors. Results are available upon request.

Table 6: Estimates of the Impact of the Apprenticeship Program on Selected Outcomes

Outcomes	Raw Diffs	Std_match	Adj_match	Part_fuzzy	IV
Employment					
Employment probability in years t+2 or t+3	**-0.04**	**-0.027**	**-0.085**	**-0.102**	**-0.093**
		0.006	0.012	0.013	0.010
Employment probability in years t+4 or t+5	**0.01**	**0.026**	**-0.023**	**0.013**	**0.007**
		0.006	0.011	0.012	0.010
Employment probability under a non-temporary job in years t+2 or t+3	**0.12**	**0.062**	**0.068**	**0.100**	**0.046**
		0.006	0.017	0.017	0.011
Employment probability under a non-temporary job in years t+4 or t+5	**0.13**	**0.088**	**0.104**	**0.177**	**0.110**
		0.007	0.017	0.017	0.011
Experience					
Accumulated number of hours worked in years t+2 and t+3	**-72.94**	**-95.61**	**-379.17**	**-491.63**	**-497.77**
		22.76	62.86	62.58	38.81
Accumulated number of hours worked in years t+4 and t+5	**-85.27**	**101.04**	**-351.95**	**-53.88**	**-27.58**
		20.04	63.12	65.43	35.26
Accumulated number of months worked in years t+2 and t+3	**-0.51**	**-0.541**	**-2.065**	**-2.693**	**-2.761**
		0.139	0.371	0.356	0.217
Accumulated number of months worked in years t+4 and t+5	**-0.60**	**0.601**	**-2.044**	**-0.287**	**-0.141**
		0.104	0.350	0.348	0.197
Probability of staying in the same establishment in years t+2 or t+3	**-0.04**	**-0.049**	**-0.058**	**-0.116**	**-0.102**
		0.004	0.009	0.010	0.007
Probability of staying in the same occupation in years t+2 or t+3	**0.00**	**-0.031**	**0.021**	**-0.015**	**-0.021**
		0.004	0.010	0.010	0.006
Turnover					
Accumulated number of dismissals between years t+2 and t+3	**-0.27**	**-0.194**	**-0.371**	**-0.498**	**-0.368**
		0.013	0.032	0.033	0.024
Accumulated number of dismissals between years t+4 and t+5	**-0.22**	**-0.083**	**-0.307**	**-0.253**	**-0.145**
		0.011	0.034	0.037	0.023
Accumulated number of admissions between years t+2 and t+3	**-0.11**	**-0.054**	**-0.126**	**-0.143**	**-0.094**
		0.014	0.027	0.030	0.025
Accumulated number of admissions between years t+4 and t+5	**-0.21**	**-0.066**	**-0.291**	**-0.227**	**-0.122**
		0.013	0.036	0.038	0.023
Probability of dismissal by quit in years t+2 or t+3	**-0.02**	**-0.006**	**-0.005**	**0.021**	**0.015**
		0.008	0.019	0.023	0.017
Probability of dismissal by quit in years t+4 or t+5	**-0.04**	**-0.027**	**-0.073**	**-0.049**	**-0.023**
		0.018	0.031	0.025	0.023
Wages					
Wage variation (in R$) between t+2 and t	**0.74**	**0.818**	**1.686**	**2.115**	**1.295**
		0.063	0.205	0.234	0.137
Wage variation (in R$) between t+3 and t	**0.86**	**0.848**	**1.887**	**2.472**	**1.616**
		0.056	0.199	0.218	0.129
Wage variation (in R$) between t+4 and t	**0.98**	**0.786**	**1.947**	**2.391**	**1.454**
		0.143	0.236	0.301	0.200
Wage variation (in R$) between t+5 and t	**1.12**	**0.499**	**2.963**	**2.673**	**1.992**
		0.321	0.492	0.433	0.358

Notes: Column (1), raw differences, presents the simple differences in outcome variables between treatment and control groups as reported in Table 5. Standard matching (column 2) refers to propensity score matching based on an *Epanechnikov* kernel with a bandwidth of 0.02. Adjusted matching (column 3) adjusts column 2 with the correction term proposed in Dias *et al.* (2013). The covariates used for matching were dummies for gender, schooling, industry, occupation, geographical region, and the year in which the worker first entered the formal sector. Partially Fuzzy (column 4) is based on Battistin and Rettore (2008). The last column presents a standard IV estimation. The instrument for all estimates is a dummy that assumes value 1 (0) if the age of the worker is 17 (18). Standard errors (in parentheses) were computed from bootstrap with 100 replications.

Results in Table 6 are organized as in Table 5 for the four groups of outcomes of interest: employment probability, experience, turnover, and wages. The results in the last three columns of the table show that the impact of the Apprenticeship program was:

- negative and statistically significant on the probability of being employed in a formal job in the short run (two or three years after the program);
- not statistically significant on the probability of being employed in a formal job in the medium run (four or five years after the program) according to the last two estimates, while the adjusted matching estimate was negative and significant, although small;
- positive and statistically significant on the probability of being employed in a non-temporary formal job both in the short and in the medium run, with much larger estimates (from 10.4 to 17.7%) for the medium run;
- negative and statistically significant on accumulated hours and months worked in formal jobs in the short run;
- not significant on accumulated hours and months worked in formal jobs in the medium run according to the last two estimates, while the adjusted matching estimate was negative and significant;
- negative and statistically significant on the probability of staying in the same firm in the short run (estimates in the range of -5.8% to -11.6%);
- mixed and small on the probability of staying in the same occupation in the short run (estimates in the range of -2.1% to 2.1%);
- negative (large in absolute terms) and statistically significant on the accumulated numbers of dismissals and admissions both in the short and in the medium run;
- not significant on the probability of quitting in the short run;
- negative and significant on the probability of quitting in the medium run, with the exception of the standard IV which renders insignificant estimates;
- positive, large and statistically significant on real hourly wages levels both in the short and in the medium run. Estimates vary from R$1.30 to R$2.12 in year *t+2*, and increase with the time horizon, varying from R$2.0 to R$ 2.97 in year *t+5*. These numbers correspond to substantial increases with respect to the average real hourly wages of apprentices in year *t* which was R$1.95.

Note that these findings are very similar across estimation procedures.

Altogether the results suggest that the program is capable of increasing the employability of apprentices. In particular, the program has a large impact on real wages. It is also very effective in increasing the probability of treated youth of being employed in a non-temporary job in the formal sector, especially in the medium run. We also find a much lower turnover for participants in the program both in the short and in the medium run. On the other hand, we find a negative effect on accumulated formal labor market experience in the short run, which tends to vanish after 4-5 years.

These findings are compatible with the interpretation that the program increases either the reservation wage or the “reservation job quality” of participants. As a result apprentices tend to spend relatively more time in the short run searching for stable/high wage jobs, possibly as non-employed. After a while they tend to find these better-quality jobs.

This interpretation goes in line with: i) slightly lower employability in the short run; ii) higher chances of getting a non-temporary contract both in the short and medium run; iii) lower levels of experience in the short run, but not in the medium run; iv) lower turnover in the short and in the medium run; and vi) higher real wages that increase over time.

6. Concluding comments

Youth-targeted ALMPs have been implemented all around the world, reflecting evidence of scarring effects of early unemployment experiences. In developing countries the focus has been on training programs which make sense given a general scarcity of skills. However, there are just a few evaluations of the effectiveness of youth-training programs in developing countries in the literature.

We provide a first evaluation of the Apprentice Act, a subsidized youth-targeted training ALMP that has been implemented in a large scale in Brazil since 2000. We make use of a very large restricted-access longitudinal matched employee-employer dataset (*Rais, Relatório Anual de Informações Sociais*), based on administrative data collected by the Labor Ministry, that contains information on the employment histories of all formal workers in Brazil from 1998 to 2010. We measure the impact of the program on four groups of outcomes that represent formal labor market attachment and remuneration, using other temporary workers as a control group. The analysis is carried on for the short run (two and three years after the program) and the medium run (four and five years after the intervention).

We employ three distinct estimation procedures which deal with self-selection by exploiting a discontinuity by age in the eligibility to the Apprenticeship program. Our main estimator is the one proposed by Dias *et al.* (2013), which is an adjusted matching estimator that corrects the standard matching approach with an IV estimated correction term based on a sharp observed cutoff criterion. For robustness purposes we also use a partially-fuzzy regression discontinuity estimator due to Battistin and Rettore (2008) and a standard parametric IV.

We find that the program increases the employability of apprentices. In particular, we find a very large impact on real wages that increase over time. We also find that the program is effective in increasing the probability of employment in a non-temporary job in the formal sector, especially in the medium run. The impact of the program on turnover is negative both in the short and in the medium run. On the other hand, we find a negative effect on accumulated formal labor market experience in the short run, which tends to vanish after 4-5 years.

These results are robust to our choice of methods that deal with selection into the program, holding for the whole set of estimation procedures.

Appendix: Identification Result in Dias *et al.* (2010)

This appendix informs the reader how to use assumptions A1 and A2, described in section 5.1 above and reproduced below, to reach the identification of the counter-factual component of the ATT parameter. The identification conditions are:

A1: $Y0 \perp Z|X$;

A2: There exists a set of points $\{z^*, z^{**}\}$ in the domain of Z where for all X:
$P\big[D = 1\big|X,\, Z = z^*\big] = 0$ and $0 < P\big[D = 0|X,\, Z = z^{**}\big] < 1$.

Following Dias et al. (2010), we first have that

$$\begin{aligned} E[Y0|X] &= E[Y0|X,Z] \\ &= E[Y0|X,Z,D=0]P[D=0|X,Z] + E[Y0|X,Z,D=1]P[D=1|X,Z] \\ &= E[Y0\,\big|\,X,Z=z^*,D=0], \end{aligned}$$

where the first equality comes from A1. The second equality holds for any z, in particular for Z = z*. Hence the third inequality comes from A2 when Z = z*.

Since it is always true that

$E[Y0|X] = \ E[Y0|X,D=0]P[D=0|X] + E[Y0|X,D=1]P[D=1|X]$, we can write

$$\begin{aligned} E[Y0|X,D=1] &= \frac{E[Y0|X] - E[Y0|X,D=0].\,P[D=0|X]}{P[D=1|X]} \\ &= \frac{E[Y0\,\big|\,X,Z=z^*,D=0] - E[Y0|X,D=0].\,P[D=0|X]}{P[D=1|X]}, \end{aligned}$$

where the last equality comes from the previous result. Now, with some algebraic manipulation of the last expression we obtain that

$$E[Y0|X,D=1] = E[Y0|X,D=0] + \frac{E[Y0\,\big|\,X,Z=z^*,D=0] - E[Y0|X,D=0]}{1 - P[D=0|X]}$$

This expression corresponds to equation (4) in Dias *et al.* (2010).

References

Adda, J., C. Dustmann, C. Meghir, and J.M. Robin (2013), "Career Progression, Economic Downturns and Skills", Cowles Foundation Discussion Paper n. 1889.

Almeida, R. and P. Carneiro (2012), "Enforcement of Labor Regulation and Informality", *American Economic Journal: Applied Economics*, 4(3): 64–89.

Attanasio, O., A. Kugler, and C. Meghir (2011), "Subsidizing Vocational Training for Disadvantaged Youth in Colombia: Evidence from a Randomized Trial", *American Economic Journal: Applied Economics,* 3: 188–220

Battistin, E. and E. Rettore (2008), "Ineligibles and eligible non-participants as a double comparison group in regression-discontinuity designs". *Journal of Econometrics*, 142: 715–730.

Bell, D. and D. Blanchflower (2010) "Youth Unemployment: Déjà Vu?", IZA Discussion Paper n. 4705, Institute for the Study of Labor.

Betcherman, G., M. Godfrey, S. Puerto, F. Rother, and A. Stavreska (2007), "A review of interventions to support young workers: Findings of the youth employment inventory", World Bank Social Protection Discussion Paper 0715.

Biavaschi, C., W. Eichhorst, C. Giulietti, M. Kendzia, A. Muravyev, J. Pieters, N. Rodríguez-Planas, R. Schmidl, and K. Zimmermann (2012), "Youth Unemployment and Vocational Training", IZA Discussion Paper n. 6890, Institute for the Study of Labor.

Blundell, R., M. Costa-Dias, C. Meghir and J. van Reenen (2003), "Evaluating the Employment Effects of a Mandatory Job Search Program", *Journal of the European Economic Association*.

Caliendo, M., S. Kunn, and R. Schmidl (2011), "Fighting youth unemployment: the effects of active labor market policies", IZA Discussion Paper n. 6222, Institute for the Study of Labor.

Card, D., P. Ibarrarán, F. Regalia, D. Rosas, and Y. Soares (2011), "The Labor Market Impacts of Youth Training in the Dominican Republic: Evidence from a Randomized Evaluation." *Journal of Labor Economics*, 29(2): 267-300.

Card, D., J. Kluve and A. Weber (2010). "Active Labor Market Policy Evaluations: A Meta-Analysis." *Economic Journal*, 120(548): F452-F477, November.

Centeno, L., M. Centeno, and A. Novo (2009) "Evaluating job-search programs for old and Young individuals: heterogeneous impacto n unemployment duration", *Labour Economics*, 16(1):12-25.

Cruces, G., A. Ham, and M. Viollaz, (2012), "Scarring effects of youth unemployment and informality: Evidence from Argentina and Brazil", CEDLAS Discussion Paper.

Cunningham, W. and J. B. Salvagno (2011), "Youth Employment Transitions in Latin America", Policy Research Working Paper No. 5521, The World Bank.

De Giorgi, G. (2005) "The New Deal for Young People Five Years On", *Fiscal Studies*, 26(3): 371-383.

Dias, M., H. Ichimura and G. van den Berg (2013), "Treatment Evaluation with Selective Participation and Ineligibles". *Journal of the American Statistical Association*, Journal Articles.

Dorsett, R. (2006) "The new deal for young people: effect on the labour market status of young men", *Labour Economics*, 13(3): 405-422.

Eliason, M. and D. Storrie (2006), "Lasting or Latent Scars? Swedish Evidence on the Long-Term Effects of Job Displacement", *Journal of Labor Economics*, 24(4): 831-856.

Farber, H. (1998), "Mobility and Stability: The Dynamics of Job Change in Labor Markets", Working Paper, Industrial Relations Section 400, Princeton.

Gonzaga, G. (2003), "Labor Turnover and Labor Legislation in Brazil", *Economía: Journal of the Latin American and Caribbean Economic Association*, 4(1): 165-207.

Gonzalez, R. (2009), "Políticas de Emprego para jovens: entrar no mercado de trabalho é a melhor saída?", in Abrahão, Aquino e Andrade (org.), *Juventude e Políticas Sociais no Brasil*, IPEA.

Hahn et al. (2001)

van der Klaauw (2002)

Kletzer, L. and R. Fairlie (2003). “The Long-Term Costs of Job Displacement for Young Adult Workers," *Industrial and Labor Relations Review*, July 2003.

Kluve, J. (2010), “The Effectiveness of European Active Labor Market Programs”. *Labour Economics*, 17(6): 904-918.

Ibarrarán, P., L. Ripani, B. Taboada, J. M. Villa, and B. García (2012), “Life Skills, Employability and Training for Disadvantaged Youth: Evidence from a Randomized Evaluation Design”, IZA Discussion Paper No. 6617.

Larsson, L. (2003) “Evaluation of Swedish Youth Labor Market Programs”, *The Journal of Human Resources*, XXXVIII: 891-927.

Ministerio do Trabalho e Emprego (2009), *Manual da Aprendizagem: O que é preciso saber para contratar o aprendiz,* 4. ed. – Brasília: MTE, SIT, SPPE, ASCOM.

de Moura Castro, C. and A. Verdisco (1998), “Training Unemployed Youth in Latin America: Same Old Sad Story?”, Inter-American Development Bank, Washington D.C.

Nilsen, O. and K. Reiso (2011), “Scarring effects of early career unemployment”, mimeo.

OECD (2010) “Off to a good start? Jobs for Youth” OECD, Paris.

Silva, E. e Andrade, C. (2009), “A Política Nacional de Juventude: avanços e dificuldades”, in Abrahão, Aquino e Andrade (org.), *Juventude e Políticas Sociais no Brasil*, IPEA.

Skans, O. (2011), “Scarring Effects of the First Labor Market Experience”, IZA Discussion Paper n. 5565, Institute for the Study of Labor.

Urzua, S. and E. Puentes (2010), "La evidencia del impacto de los programas de capacitación en el desempeño en el mercado laboral," IDB Publications 38038, Inter-American Development Bank.

von Wachter, Till and S. Bender (2006) ‘At the Right Place at the Wrong Time: The Role of Firms and Luck in Young Workers’ Careers.’ *American Economic Review* 96(5): 1679-1705.

World Bank (2002). *Brazil Jobs Report*, vol. 1. Washington.