

Factors associated with dropout rates in public secondary education in Minas Gerais¹

Tufi Machado Soares^{II}

Neimar da Silva Fernandes^{II}

Mariana Calife Nóbrega^{II}

Alexandre C. Nicolella^{III}

Abstract

This paper's aim was to identify the main factors influencing secondary education dropout. The data are based on a historical series developed by the National Household Sample Survey (PNAD) and a large survey conducted in Minas Gerais, Brazil, which collected diverse information from 3,418 interviewees (including students and dropouts). Cox proportional hazards regression models were estimated to identify potential correlations between intra and extra-school factors and early dropout. Another model used PNAD database, for which a Logit model was estimated, to verify non-dropout rates in high schools in Minas Gerais. Some significant factors are highlighted in the results explaining dropout, such as: difficulties faced with subjects, desire for a different school, perception of better job opportunities if studies are completed, and importance assigned to school choice.

Keywords

Dropout rates – Secondary education – PSAE – PNAD.

I - The authors thank Instituto Unibanco for project funding, and Capes and CNPq, which partially funded this study through project Rede Multinível de Pesquisa em Eficácia Escolar - Casadinho/Procad, and the support of Minas Gerais State Department and Centro de Avaliação de Políticas Públicas em Educação (CAEd).

II - Federal University of Juiz de Fora, Juiz de Fora, MG, Brazil.

Contact: tufi@caed.ufff.br

III - University of São Paulo, Ribeirão Preto, SP, Brazil.

Contact: anicolella@fearp.usp.br

Fatores associados ao abandono escolar no ensino médio público de Minas Gerais¹

Tufi Machado Soares^{II}
Neimar da Silva Fernandes^{II}
Mariana Calife Nóbrega^{II}
Alexandre C. Nicolella^{III}

Resumo

Este artigo tem como objetivo apontar quais são os principais fatores que influenciam o abandono escolar no ensino médio. Os dados utilizados baseiam-se em uma série histórica construída a partir do levantamento das Pesquisas Nacionais por Amostra de Domicílios (PNAD) e numa ampla pesquisa realizada no estado de Minas Gerais (MG), a qual culminou na coleta de diversas informações de um total de 3.418 entrevistados (entre alunos cursantes e não-cursantes). Com base nesses dados, estimaram-se modelos de regressão de risco proporcional de Cox, que permitiram identificar a correlação entre fatores intra e extraescolares com o abandono precoce. Outro modelo desenvolvido utilizou as bases de dados da PNAD, em que um modelo logito foi estimado, permitindo verificar a taxa do não abandono na trajetória do ensino médio também para o estado de Minas Gerais. Entre os resultados encontrados, destacam-se alguns fatores expressivos na explicação do abandono, tais como: a dificuldade nas disciplinas, ânsia por uma escola diferente, percepção de melhores oportunidades de trabalho com a continuidade dos estudos e a importância atribuída na escolha à escola.

Palavras-chave

Abandono escolar – Ensino médio – PSAE – PNAD.

I- Os autores agradecem ao Instituto Unibanco pelo financiamento do projeto, à Capes e ao CNPq, que financiaram parcialmente esse estudo por meio do Projeto *Rede Multinível de Pesquisa em Eficácia Escolar - Casadinho/Procad*, e ao apoio da Secretaria Estadual de Minas Gerais e do – Centro de Avaliação de Políticas Públicas da Educação (CAEd).

II- Universidade Federal de Juiz de Fora, Juiz de Fora, MG, Brasil.
Contato: tufi@caed.ufjf.br

III- Universidade de São Paulo, Ribeirão Preto, SP, Brasil.
Contato: anicolella@fearp.usp.br

Introduction

High school dropout is severe in Brazil and is a problem that affects even countries with a high level of economic development. In a review including 203 papers published in the United States in the preceding 25 years, Rumberger and Lim (2008) divided the factors predicting whether students would abandon school or successfully graduate from high school into two groups: those associated with the students' individual characteristics and those associated with the institutional characteristics of their families, schools and communities.

Among the individual characteristics the authors include educational achievement (academic performance and mobility in both elementary and middle schools, academic performance in high school and academic failure in elementary and middle schools; student's behavior and attitudes (e.g., involvement in academic and social activities, absenteeism, and educational expectations); demographic characteristics (race, gender); and prior experiences (e.g., having attended preschool).

Among institutional characteristics, the survey indicates three family aspects: 1) family structure (e.g. traditional or single-parent family) and changes in this structure (e.g., divorce) over the course of the school process; 2) family income and other resources; and 3) social capital (e.g., high educational expectations, parental monitoring of the children's school progress and participation in the children's school life). Additionally, four generic school characteristics are listed: 1) composition of its students; 2) resources; 3) policies and practices; and 4) other structural characteristics. The authors, however, stress that little evidence has been found concerning the importance of school resources, though there is strong evidence indicating that small classes (around 15 students per classroom) is a beneficial factor. On the other hand,

school and academic climate, reflected, for instance, in the percentage of students doing homework, is highly relevant. Finally, characteristics of the community in which the child lives, such as poverty or wealth levels, in addition to family characteristics, seem to play an important role in a student's decision regarding quitting or graduating from school.

Various studies in Brazil report that high school dropout¹ is influenced by the need for youngsters to enter the job market, be it because they need to contribute to the family budget or a desire to have their own money (ARROYO, 1993; MEKSENAS, 1998). This view may disregard diverse factors that young individuals take into account when they decide to quit school. The many factors accounted for driving students away from school already reported by the literature include a lack of interest in/for school.

Eckstein and Wolpin (1999) show that less motivated students with low expectations concerning rewards from graduation are more likely to quit school. Additionally, poor quality schools, as perceived by students, tend to present higher dropout rates (HANUSHEK; LAVY; HITOMI, 2006). Other important aspects include curriculum, size, student-teacher ratio, infrastructure, and an excess of content (LEE; BURKAM, 2003). Nonetheless, little understanding regarding these questions has been achieved in Brazil so far.

Lack of interest, even with high rates of return from educational investment, also appears to be a strong factor influencing a student's decision to dropout from school and understanding this aspect is key both to acquiring a better understanding of the fragile state of these individuals and to producing an important indicator powerful enough to direct educational policies aimed

1- The concept provided by INEP (Anísio Teixeira National Institute of Education Studies and Research) was adopted, that is, dropout is a condition in which the student stops attending school during the course of the school year.

to reverse such a scenario (OREOPOULOS, 2007). Hence, a dichotomy between internal and external factors that cause school dropout becomes apparent. Additionally, Stearn and Gleinne (2006) show that these factors may vary according to the student's age, grade, and socioeconomic group. External factors include: job, social inequalities (BOURDIEU, 1998), pregnancy, and the need to look after family members. In turn, internal factors include differences in the language used by school actors, teachers' attitudes, characteristics of the school's management and pedagogical program, among others.

Various Brazilian studies addressed the influence of extrinsic factors on dropout rates (MESKENAS, 1998; ARROYO, 1993, 1999a; BRANDÃO, 1983; GATTI, 1981; LOPEZ DE LEON; MENEZES-FILHO, 2002; JANOSZ et al., 1997). Brandão (1983) states the family is the strongest determining factor in children's academic failure, whether because parents do not monitor the children's school activities or because of the living conditions they provide, showing a positive relationship between the mother's educational level and the child's academic performance and retention in school. Janosz et al. (1997) add that more permissive parents with little educational ambition also strongly influence dropout. In this same line of reasoning, Lopez de Leon and Menezes-Filho (2002) indicated other family characteristics that play a role in school dropout, such as size and type of family, the existence of a history of school dropout and socioeconomic level.

The job market plays a considerable role in one's decision to quit school, especially for those with families facing financial hardships. Meksenas (1998) corroborates this perception and adds that students attending school during evening hours, usually those from lower economic classes, arrive to school exhausted from daily work routines and many of these adolescents, discouraged

by the school's low quality, give up studies without graduating from high school. It is worth noting that Neri (2009), based on data provided by PNADs (2004, 2006), reports potential reasons for school dropout: lack of schools (10.9%), the need for students to have a job and income (27.1%), and a lack of interest in school (40.3%), among others (21.7%).

When the international production of studies addressing the topic is compared to the Brazilian production, especially in regard to high school education, a lack of studies is verified. Such studies can corroborate, clarify or even improve upon different explanations for the phenomenon of school dropout. The importance of each factor reported by the literature as influencing school dropout is particularly unknown. Furthermore, there is a lack of empirical studies. Neri (2009) advanced in this direction using data provided by PNAD; however, much more information concerning the different actors, school institutions and intra- and extra-school relationships is required to understand why students are not interested in school and do not graduate from high school.

Therefore, this study's aim was to indicate characteristics associated with school dropout that hamper the completion of high school by individuals with at least eight years of schooling in public schools in the state of Minas Gerais, Brazil. More specifically, the objective is to produce empirically based knowledge using secondary data provided by PNADs (1987 to 1992) and primary data produced by PSAB - *Pesquisa sobre Abandono Escolar* [Survey on School Dropout].

Based on these data, Cox proportional hazards regression models were estimated and enabled the identification of correlations between intra and extra-school factors and early dropout. The use of this type of model represents an important contribution to this methodological approach to the problem, unpublished so far, at least at a national level.

Another model used PNAD's database, in which a logit model was estimated and enabled verifying non-dropout rates in high schools in Minas Gerais. Some expressive factors found among the results stand out in the explanation for dropout, such as: having difficulty with subjects; a desire for a different school; a perception there are better job prospects with continuity of studies; and the importance attributed to school choice. All these forms of analysis enabled verifying the robustness of results found in each of the approaches.

Databases: PNAD and PSAE

Using the results of the surveys previously mentioned, the proposal of this study is to deepen analysis regarding factors that influence high school dropout. For that,

two databases were used. The first database is composed of data generated by the National Survey for Households Sample (PNAD) for the years 2001 to 2008. The second database was generated by PSAE (Survey on School Dropouts/MG) in 2009. The PSAE study resulted in a final report *Determinantes do abandono escolar no ensino médio de Minas Gerais [Determinants of High School Dropout in Minas Gerais]*, published in 2011. Data from this report were employed in the analysis that generated the following results.

The PNAD database was composed of individuals born in 1987, 1988, 1989, 1990, 1991 and 1992 in Brazil and in the state of Minas Gerais. Hence, theoretically, we would have the structure presented in Table 1, that is, age and grade represent an ideal situation, as if there were no gaps.

Tabela 1 – Estrutura teórica idade/série da base de dados utilizada, PNAD 2001 a 2008

| Birth cohort | Ideal age/grade | Survey year (PNAD) | | | | | | | |
|--------------|-----------------|--------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| | | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 |
| 1987 | Age | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
| | Ideal grade | 8 ^a | 1 ^o | 2 ^o | 3 ^o | Sup. | Sup. | Sup. | Sup. |
| 1988 | Age | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
| | Ideal grade | 7 ^a | 8 ^a | 1 ^o | 2 ^o | 3 ^o | Sup. | Sup. | Sup. |
| 1989 | Age | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
| | Ideal grade | 6 ^a | 7 | 8 ^a | 1 | 2 ^o | 3 | Sup. | Sup. |
| 1990 | Age | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| | Ideal grade | 5 ^a | 6 ^a | 7 ^a | 8 ^a | 1 ^o | 2 ^o | 3 ^o | Sup. |
| 1991 | Age | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| | Ideal grade | 4 ^a | 5 ^a | 6 ^a | 7 ^a | 8 ^a | 1 ^o | 2 ^o | 3 ^o |
| 1992 | Age | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| | Ideal grade | 3 ^a | 4 ^a | 5 ^a | 6 ^a | 7 ^a | 8 ^a | 1 ^o | 2 ^o |

Source: PSAE Final Report, 2011

It is known that the use of PNAD² data does not constitute a panel study, however, it does allow some analyses over the years with individual and family information, such as race, gender, parents' schooling, and *per capita* income, among others. The preceding structure enables analyzing changes that took place in the same group of individuals, for instance, those born in 1988, rather than utilizing school years; that is, we can observe the behavior of this specific group in regard to school attendance, as they get older.

The PSAE database is composed of data collected from 3,418 individuals, including high school students and dropouts, from 46 public schools in Minas Gerais, Brazil. Dropouts were individuals who left school in 2006, 2007, 2008 and 2009. The population³ of students was stratified according to the educational centers defined by the State Department of Education of Minas Gerais⁴; by state and municipal school systems; by modalities (regular, Youth and Adult Education [EJA], and career-technical education); by the last grade the student was enrolled; and finally, by the individual's status (i.e., student or dropout).

Once the schools were randomly drawn, those individuals who quit high school in 2009 and those who quit between 2006 and 2008 were recorded so that they could be located in their homes and interviewed according to a systematic procedure, until a total of 20 students per school were interviewed. These specific years were selected because of the difficulty in finding older dropouts and because more recent data are more reliable. A total of 60

students, distributed per grade, were selected in these schools.

The questionnaires⁵ (instruments) were composed with the objective to investigate individual characteristics, such as the individual's school history and socioeconomic situation; the school's characteristics, such as the individuals' perceptions concerning the teachers; and the family characteristics. These instruments were intended to capture intra- and extra-school conditions and their relationship with the school histories of both high school students and dropouts.

A questionnaire was structured according to general themes and varied concepts concerning the form of questions or statements. Usually, most of these questions are associated with the same factor, which is what one desire to assess. The responses assigned by the respondents may or may not meet the expectations of the experts who developed the questionnaire. In the first case, empirical data confirm the association of responses assigned to the set of questions originally proposed to assess the same factor. However, the responses assigned to certain sets of questions may be associated to each other, even if this is not the expectation of the instrument's authors. For both cases, it is more reasonable to use the individual responses to each question but identify subsets, the responses of which are associated, and use a general measure for the factor.

A more comprehensive definition used in this study refers not only to the individuals' exogenous social relationships but also to individual characteristics and the interactions of these actors in the school setting. The variables can be classified into two groups: directly observed simple variables (i.e.,

5- The sample was non-proportionally subdivided according to all these strata. For this reason, the analysis takes into account the weight of each sample element appropriately. The cities and schools were selected according to Probability Proportional to Size (PPT) – number of schools for cities and number of students for schools.

2- Even though PNADs from 2004 and 2006 provide a supplement regarding the dropout reasons, we believe this question is not appropriate for high school students because it is only asked to people under 17 years of age.

3- Data provided by the 2007 School Census (published in 2008) were used to build a reference system complemented by school records obtained after schools were randomly drawn.

4- Center, North, South, Triângulo Mineiro, Vale do Aço and Zona da Mata.

measured by one or two direct questions), or latent variables/factors (i.e., indirectly measured through a set of questions).

Taking into account this perspective and considering the instrument's limitations, in order to obtain a measure of the most important factors associated with the individuals' school history, we used a statistical technique labeled factor analysis, which enables reducing the number of variables used in the analysis. Factor analysis identifies common factors

in different questions summarizing the information of an entire set of statements/questions through a few measures or even a single one.

Afterwards, these factors are interpreted and assigned a semantic representation. The objective is to combine and synthesize the results of several indicators to provide a quality measure for factors that cannot be directly measured or observed. Table 2 presents a list of factors used in the statistical models proposed in this paper:

Table 2 – Description of PSAE factors, 2009

| Factor | Description |
|--|--|
| Socioeconomic Index (ISE) | Refers to the respondent's socioeconomic status and is established using questions that measure the possession of goods (cultural or comfort items), taking into account whether the family receives <i>Bolsa Família</i> (conditional cash transfer). |
| School's perceived quality | Refers to the student's perception concerning the school's quality of education and the teachers' work, attendance in vocational practical classes, encouragement provided by the school, and the importance of what is taught at school. |
| Interest and encouragement provided by the family to studies | Refers to the interest showed by the family in the individual's school life, the extent to which parents demand good behavior, good grades, attendance and the encouragement provided. |
| General difficulty faced with the subjects' content | Refers to the level of difficulty the student faces in various subjects (math, biology, history, geography, Portuguese language, physics and chemistry). |
| Desire for a dynamic and innovative school | Refers to the student's desire for after-school activities, practical classes, a more frequent use of handouts, contextualization of school content, and a preparatory program for the job market. |
| Need to work to help the family | Addresses the reasons individuals started working and one of the options is "need to help the family". |
| Intention to go to college | The students are asked what they intend to do when graduating from high school and one of the options is "to attend college". |
| School is chosen because of its Quality and/or Affinity. | The students are asked the main reasons they opted for that specific high school. |

Source: PSAE final report, 2011

The dropout variable (dependent variable) was constructed to accurately establish the students' grade⁶ : 10th, 11th, 12th or whether they were still attending middle school. A complementary question asked whether the individuals were enrolled at school and if yes, how they had completed their studies in that year: passed, failed, or

quit before the end of the year. In this way, data provided by PSAE enabled identifying how many times an individual quit school, what was the interval between dropouts, how many times the student failed, and how many years the student took to pass a given grade.

Another set of variables was used for the models using PNAD data as presented in Table 3.

6- The age recommended to initiate high school is 15 years old.

Table 3 – Description of variables extracted from PNAD, 2001-2008

| Factor | Description |
|-----------------------------|--|
| Number of siblings | Numbers of siblings |
| Older than 18 years old | Whether the individual is older than 18 years old |
| Age/grade gap | Grade the student is enrolled minus the grade in which s/he should be enrolled according to year of birth. |
| Occupation | Work situation: 1 if the individual has a job and 0 (zero) if not. |
| Mother's age | Mother's age |
| Mother's level of education | Mother's years of schooling |
| Fathers age | Father's age |
| Father's level of education | Father's years of schooling |
| Energy | Whether the household has electricity |
| Water filter | Whether the individual has a water filter at home |
| Working age | Whether the individual is 18 years old or older |

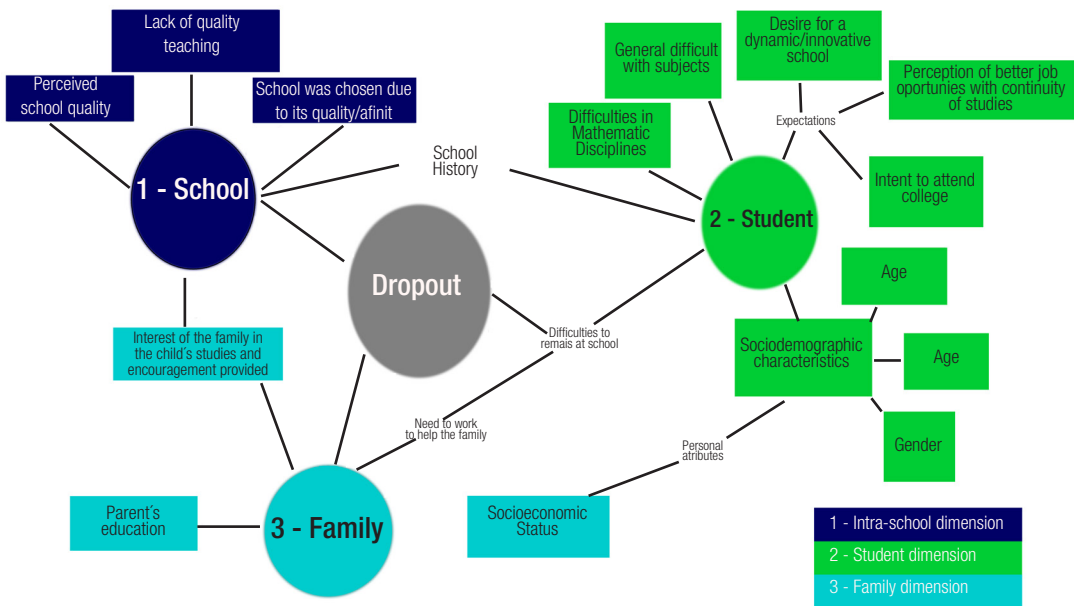
Source: PSAE report, 2010

The conceptual model based on all the variables and factors is presented in the next section.

Conceptual model

Figure 1 presents the conceptual model developed in this study. It summarizes the factors and their interconections that appeared associated with high school dropout in this study. These factors are conceptually divided into three groups labelled here as dimensions. The first dimension includes factors related to the school such as: students' perceived quality of school; student's perceived lack of quality of the teacher's work; and choice of school due to quality/affinity. The student dimension includes the individuals' characteristics such as difficulties faced in regard to the subjects' content, future prospects, and sociodemographic characteristics. Finally, there are factors related to the family list: socioeconomic status, parents' education, and interest and encouragement provided by the parents.

Figure 1 – Conceptual model – instrument's dimensions



Source: Developed by the authors

The interconnections described in the model – school history, difficulties remaining in school, need to work to help the family, and personal attributes – also aggregate factors associated with the decision to dropout of school.

The figure is a simplified representation of a more developed theoretical model constructed by the authors in two stages. The first stage, an *a priori* theoretical model, was constructed based on the literature review and previous qualitative studies performed by the authors, such as focal groups. Further details concerning the development and theoretical model can be found in the PSAE report (PSAE, 2011).

Based on the *a priori* theoretical model, we aggregated new conjectures originating from questions based on students' reports collected in focal groups. This information supported the development of PSAE's instruments. Therefore, the PSAE's data collection allowed the *a priori* theoretical model to be adjusted to take into account the survey's empirical evidence, which resulted in the final model previously presented.

Even though this model is much more an expression of relationships among associated factors than causal relationships, it is interesting to note how important elements of two models with a causal nexus for school dropout, extensively revised, discussed and synthesized by Finn (1989), unintentionally and naturally emerged from the prior conceptual proposal.

The first model discussed by Finn is the frustration/self-esteem model. In short, it proposes that students' poor academic practices lead to poor academic results, which decrease self-esteem, producing associated and/or reinforced negative influences from peers and behavior problems that feedback with poor academic results and low self-esteem. Persistence in this cycle progresses to a gradual disengagement of the student from school activities and eventually to

dropout. In the conceptual model proposed, these elements appear to be expressed by the variables associated with academic history and difficulties faced in academic subjects.

More elements are found in the second model, called the participation-identification model, which proposes that the quality of school education, the student's specific individual abilities, and participation in school activities (social and academic) contribute to academic success and performance, which in turn, influence one's identification with the school in the sense that there is a greater perception of belonging and the construction of individuals' values, which, when associated with appropriate and quality school education, reinforce academic success.

Naturally, when this virtuous cycle weakens, it favors disengagement from school activities and consequently results in school dropout. The elements of the conceptual model previously presented (Figure 1) that are common to the participation-identification model include school history and difficulty in subjects and also a desire for a dynamic and innovative school and perceived quality of school. These variables are probably proxies of desire for encouragement or motivation, for more intense and pleasurable academic participation and inappropriate pedagogical/instructional processes for students in unfavorable situations.

Lamborn et al. (1992), on the other hand, analyze the influence of the family and family friends on engagement with school activities, highlighting that this influence affects the individual's aspirations, represented in the conceptual model that guides the analysis of this study by one's intention to attend college and the perception of better job opportunities. Naturally, these aspirations enhance the school engagement of students, who then establish greater educational objectives than those who do not have such aspirations.

Results of the statistical models for school history

In this paper we opted to discuss the results of only three of the statistical models developed within the PSAE sphere constructed to explain high school dropout in Minas Gerais, Brazil. Two of the different types of statistical models are used: Cox proportional hazards model for PSAE data and the binary dependent variable model (logit) using pooling data with cohort dummies and year for the PNAD data.

Results of the logistic model for retention of students in high school (PNAD data)

School dropout is a particularly severe problem in the Brazilian educational system. Even though, according to secondary data provided by PNADs, the percentage of individuals who do not graduate up to 21 years of age is relatively low in comparison to Brazilian historical standards (about 14% for the cohort of 1987, in Minas Gerais), dropout is still high if compared to international standards, particularly with countries belonging to the Organisation for Economic Cooperation and Development (OECD).

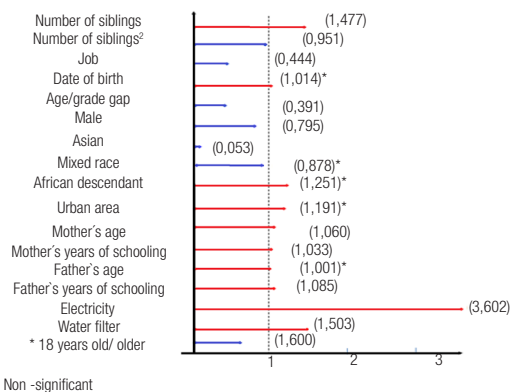
Furthermore, since few of those individuals who have not graduated from middle school are still attending school, we infer that almost all these individuals have quit middle school, at least at the time PNAD survey was conducted. When analyzing data from high school, an even more severe situation is presented: for the cohort of 1987, about 39% of the individuals in Minas Gerais had not graduated from high school at the age of 21 years old. Of these, only 20% continued in school, while the remaining were out of the school system. Hence, the situation is that 32% of the individuals had dropped out of school before graduation up to the age of 21 years old⁷

⁷- See PSAE final report, 2011..

The logit econometric model for permanence at school was built using PNAD historical series (2001 to 2008 for birth cohorts from 1987 to 1992).⁸ This model (Figure 2) explains the effect of co-variables concerning retention in high school of those born from 1987 to 1992. In the group of variables included in the model, we highlight: number of siblings, age/grade gap, work and gender.

The horizontal lines represent the coefficient⁹ of relative risk of each co-variable. This coefficient represents the effect of each co-variable on the ratio between probability of remaining in school and the probability of dropping out of high school. This ratio will be labeled risk of retention or risk of remaining in school. The inverse is the risk of dropping out of or not remaining in school. Hence, when the relative risk is greater than 1, that variable increases the likelihood of an individual remaining in school. In turn, when the coefficient (relative risk) is smaller than 1, that co-variable is associated with a lower likelihood that the individual will remain in school.

Figure 2 – Analysis of high school history: explanation of the effect on permanence at school



Source: Developed by the authors

⁸- For the interviewees who reported between 8 and 10 years of schooling. It is assumed these individuals graduated from middle school but not from high school.

⁹- The coefficient presented is known as the odds-ratio from the logit regression..

The relationship between number of siblings and retention in school is significant but not linear. Having up to three siblings increases the likelihood of remaining in school, but more than that reduces it. Students with gaps are more likely to abandon school. In fact, each gap in years reduces the risk of remaining in school to about 61%. Those who work are less likely to remain in school (about 44% of the risk compared to those who do not work). Male individuals are 20% less likely to remain in school than females. Children of older and more educated women are more likely to remain in school. Evidently, those with lower economic statuses are less likely to remain in school. (This is captured by the model through the variables “father’s years of schooling”, “electricity” and “water filter”.) The risk of remaining in school is 40% lower for those above the age of 18 years old.

COX models for dropout rate (PSAE)

Focus on sociodemographic characteristics

Session 2 presents the procedures used in data collection concerning the school history of all those interviewed for the PSAE. Among the information for the period between 2000-2009, concerning the academic performance of each interviewee was: whether the individual was attending school, was approved or failed, or quit before the end of the year. In this section, we attempt to explain school dropout using intra- and extra-school variables: socioeconomic index, gender, pregnancy, age/grade gap, job, number of siblings, order of birth among siblings, difficulty in subjects, desire for a dynamic/innovative school, interest and encouragement provided by the family, perceptions of students concerning better work opportunities if they remain in school, whether there is an intention to attend

college, and whether the school was chosen due to affinity.

Note that there may be incomplete information because individuals are enrolled in high school in different years, while others quit and may or may not go back to school later. This type of information cannot be treated the same way as cross-sectional information, as is the case of data provided by PNAD. Data in this case clearly require treatment that takes into account the point in time when each student is observed, from the time the individuals enter in high school.

Data also require what is known in the literature as censorship; that is, the time of observation naturally imposes a limitation on the study, as not all students will be observed over the course of the entire high school period. Therefore, those who have not entered high school yet will be censored, meaning this information is incorporated into the model differently than that of those who quit.

School dropout can be a recurrent event. Hence, we should allow that the information of a student who dropped out and then resumed school be incorporated into the analysis as a recurrent measure. For this reason, we opted to produce the dropout analysis based on the COX model, adjusted for recurrent events (KAPLAN; MEYER, 1958).

This model explains the probability, labeled dropout, of a student leaving school in every enrolled year in high school based on his/her co-variables. Hence, $\lambda(t)$ represents the student’s dropout rate, COX model is given by:

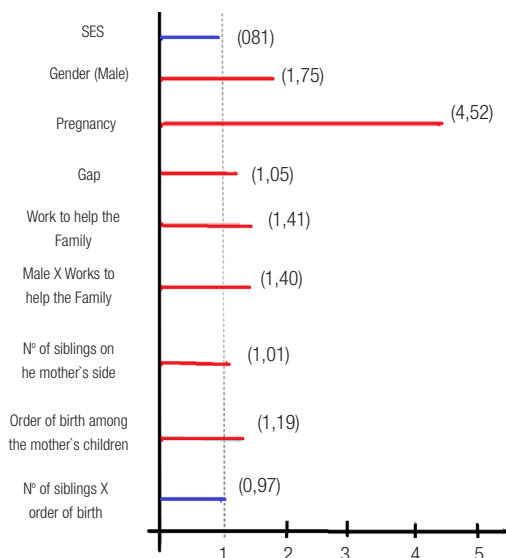
$$: \lambda_0(t) e^{\sum \beta_i x_i}$$

in which $\lambda_0(t)$ refers to dropout rate observed in the population and calculated by Kaplan Meyer estimator (COLOSIMO, 1982); x_i is an i -th covariate considered in the model.

The following is emphasized in the presentation and analysis of the second model (Figure 3): socioeconomic status, pregnancy, grade/age gap, gender, job, order of birth

among siblings, and number of siblings. For a more appropriate reading of the graphics, we suggest interpreting the dotted line as the dropout rate of the population computed in the model. In regard to the horizontal lines, they represent the effect of an increased or decreased dropout rate associated with the corresponding covariate (e^{β_i}). Variables with coefficient equal to one do not influence this rate, hence, such a characteristic is not represented by the statistical model. When the coefficient is greater than one, it tends to increase dropout rates. The inverse is applied when the coefficient is less than one.

Figure 3 – Effect on dropout rate



Source: Developed by the authors

The effect of socioeconomic status – already exhaustively explicated in the literature as influencing performance – is confirmed in the results presented in figure 3. The higher one's socioeconomic status, the lower the dropout risk rates. Another determinant factor confirmed here is pregnancy, which, in this group of interviewees, increases dropout rates by 352%.

It is worth noting, in this second model, the relationship of the age/grade gap because each year a student is delayed from entering high school increases the dropout risk by 5%. Therefore, if the student is three years/grades behind in his/her school history, dropout risk rates rise by 16%.

In model 1, male individuals were more likely not to remain in school; the same result is found in Model 2 because male individuals are 75% more likely to quit school compared to females. Additionally, working to help the family increases dropout rates by 41% in the case of female students and by 97.4% in the case of male students.

The interesting fact in this model is that it suggests that order of birth among siblings influences dropout rates. Older siblings tend to present higher dropout rates. When the family is bigger, however, this effect is minimized.

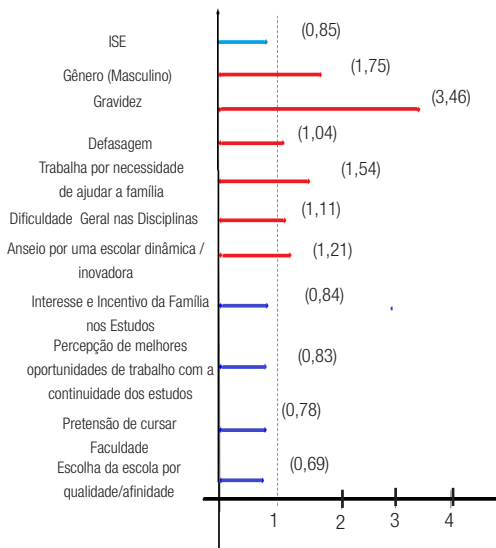
Focus on intra-school characteristics

The third statistical model (Figure 4) presents several more complex associations among the variables linked to the context closer to school that influence dropout rates. The model explains dropout rates through individual characteristics, intra-school and family variables and, basically, seeks to fill in a gap in the literature regarding the measurement of these characteristics. Some sociodemographic variables, the most influential ones, are used as controls, including ISE, gender, pregnancy, need to work to help the family, and the age/grade gap at the time when the student enrolls in the 10th grade.

Hence, the model seeks to analyze the net influence of these new variables, disregarding the effect of control variables, already reported in the literature and confirmed here through the models represented in figures 2 and 3. In this way, it is possible to deepen the analysis seeking to identify individual

intrinsic factors that influence the decision to dropout, reducing the chances this effect is motivated by a non-explicit association with some of the control variables. Once the peculiarity of the construction of the statistical model 3 is explained, we proceed to the relationships found.

Figura 4 – Efeito sobre a taxa do abandono com variáveis ligadas ao contexto da escola



Source: Developed by the authors

For those who experience general difficulties with academic subjects, the dropout risk is greater. It may be an indication that some of the students who quit school experience an overall difficulty, that is, difficulty in keeping up with all the subjects more so than just one or a subset of subjects. When constructing this factor, interviewees were asked about the difficulties accompanying each of the subjects and two preponderant factors emerged from this analysis. The first factor is associated with all the subjects and the second is associated with exact sciences. It seems that students with a higher propensity to quit school make this decision when facing difficulties in all the subjects more than simply those who experience difficulties specifically in the exact sciences.

It is worth noting that the construct “desire for a dynamic/innovative school”, which is interpreted as being a desire on the part of students for a school different from that they experience, and which is associated with a desire for after-school activities, vocational classes, use of handouts, contextualization of school content, and preparation for the job market, was also associated with school dropout. We do not intend to judge or assess the public school system, only interpret a desire expressed in the responses and reports of these students.

On the other hand, intent to go to college; having an encouraging family interested in the students’ school affairs; being able to perceive better job opportunities related to more years of schooling; and, having chosen the school due to its quality or affinity, whether the individual believes it has better quality of education or because s/he attended this school for his/her entire life, are associated with lower dropout rates.

Final Considerations

This study’s aim was to deepen knowledge regarding the reasons¹⁰ that lead students to dropout from high school. For that, factors traditionally indicated in the literature were investigated based on data provided by both PNADs and PSAE. We also sought to advance in the intra-school causes that influence dropout, seldom available in surveys such as PNAD or similar sources.

A general profile of those individuals more vulnerable to high school dropout was outlined: those with lower economic statuses, male individuals, those with prior history of school failure, dropout, or poor academic performance, those with a lack of interest or motivation, with poor participation in academic activities, and those experiencing special situations such as early pregnancy. Apparently, the school has little to do in regard to the students’ socioeconomic conditions, but should be attentive to the more vulnerable groups.

10- The complete study regarding this survey concerning school dropout is represent in PSAE Final Report, 2011.

Specific projects and programs devised to increase student interest and encourage individuals in at-risk situations to participate in academic and social activities beginning in elementary school years are needed before this risk situation becomes irreparable. Many authors tend to agree that dropout is a process that begins in first grade.

Immediate policies to reduce failure in elementary and middle schools are needed to improve school flow. These policies should not be based on less demanding requirements intended to promote education but on ongoing and efficacious recovery of the students' cognitive skills in the various subjects and increased motivation to complete academic stages. Students need to feel they are learning and that this effort will be eventually rewarded. Failing marks or grades should be seen as an exception and last resort in the school process, and when adopted, be based on very clear and standardized criteria. In our opinion, failure should not take place in elementary school and avoided as much as possible in the subsequent stages.

It is clear that family mediation is very important in the case of dropouts because of the importance family attributes to education. Parental interest and encouragement may be determinant in ensuring the continuity of education, for students to make the effort necessary to complete academic stages, persist to overcome obstacles, and enhance their ability to deal with frustrations. Young individuals in situations of risk belonging to lower economic classes lose twice: first, due to the family's background, which does not have prior experience to construct a cultural capital that is relevant and related to professional success and potential social mobility. Probably, the importance the family attributes to school is more associated with acquiring a diploma per se than to the quality of education. Additionally, given the need to contribute to the family income or, at least, reduce expenditures, it is difficult to continue encouraging an individual to remain at school when s/he has experienced successive academic failures. Hence, the need to work and contribute to the family income may be a factor that has in its origin

not only the need to provide for oneself or the family's livelihood, but also the fact that children of poorer families become uncomfortable and feel unproductive spending many hours at school.

Based on the reports and responses provided by the students to the questionnaires, it seems that the school that is traditionally structured makes learning a burdensome task for many students. This group of students includes those at the risk of dropping out from school. Even if students glimpse the possibility of accomplishing greater success after graduating from high school, it does not seem to be sufficiently clear for them that the effort to complete this stage of education will be duly rewarded. Hence, keeping these individuals in school is a task that cannot be left to the family alone.

Wehlage and Smith (1992) address the need for a more innovative, authentic and intellectually stimulating curriculum for students in situations of risk. An analysis of successful schools and programs directed to students in unfavorable conditions reveals that small schools adapted to this profile of students are more efficacious in encouraging participation and motivating students than attempts to restructure the model of existing schools. On the other hand, Lamborn et al. (1992) emphasize evidence that participation in after-school activities encourages students to remain engaged. These are, however, little explored questions in the research conducted in Brazil in the field of education and require further investigation.

Perhaps, a way to encourage the engagement of young individuals in high school is to provide them different alternatives and let them choose part of their educational process. High schools could be structured around a few basic common competencies but additional activities, including recreational activities, which are considered to be important for an individual's education, the students themselves would be able to choose. Additionally, it is possible to allow schools to innovate, and/or meet the community's special demands, providing non-mandatory but complementary alternatives within their natural vocation. Again, this is a topic that needs to be discussed further and different models can be compared through experimentation.

References

- ARROYO, Miguel G. Educação e exclusão da cidadania. In: BUFFA, Ester (Org.). **Educação e cidadania: quem educa o cidadão**. 4. ed. São Paulo: Cortez, 1993.
- ARROYO, Miguel G. Experiências de inovação educativa: o currículo na prática da escola. In: MOREIRA, Antônio Flávio. (Org.) **Currículo: políticas e práticas**. Campinas: Papyrus, 1999. p. 131-164.
- BOURDIEU, Pierre. A escola conservadora: as desigualdades frente à escola e à cultura. In: NOGUEIRA, Maria Alice; CATANI. Afrânio (Orgs.). **Escritos de educação**. Petrópolis, Vozes, 1998. p. 40-64.
- BRANDÃO, Carlos R. **Casa de escola: cultura camponesa e educação rural**. Campinas: Papyrus, 1983.
- COLOSIMO, Enrico Antonio. Análise não paramétrica. In: COLOSIMO, Enrico Antonio; GIOLO, Suely Ruiz (Orgs.). **Análise de sobrevivência aplicada**. São Paulo: Blucher, 2006. p. 29-68.
- ECKSTEIN, Zvi; WOLPIN, Kenneth. Why youths drop out of high school: the impact of preferences, opportunities, and abilities. **Econometrica**, v. 67, n. 6, p.1295-1339, 1999.
- FINN, Jeremy D. Withdrawing from school. **Review of Educational Research**, v. 59, n. 2, 1989.
- GATTI, Bernadete A. et al. A reprovação na 1ª série do 1º grau: um estudo de caso. **Cadernos de Pesquisa**, São Paulo, v. 38, p. 3-13, ago. 1981.
- HANUSHEK, Eric. A.; LAVY, Victor; HITOMI, Kohtarō. **Do students care about school quality?** Determinants of dropout Behavior in developing countries. Cambridge: [s. n.], 2006. NBER: working paper 12737.
- IBGE - Instituto Brasileiro de Geografia e Estatística. **Pesquisa Nacional por Amostragem de Domicílios (PNAD)**. Rio de Janeiro: IBGE, 2006. Disponível em: <<http://www.ibge.gov.br/downloads>>. Acesso em: abr. 2010.
- JANOSZ, Michel et al. Disentangling the weight of school dropout predictors: a test on two longitudinal samples. **Journal of youth and adolescence**, v. 26, n. 6, p. 733-762, 1997.
- KAPLAN, Edward L.; MEYER, Paul. Nonparametric estimation from incomplete observations. **Journal of the American Statistical Association**, v. 53, p. 457-81, 1958.
- LAMBORN, Susie et al. Putting school in perspective: the influence of family, peers, extracurricular participation, and part-time work on academic engagement. In: NEWMANN, Fred M. (Org.). **Student engagement and achievement in american secondary schools**. New Yourk, Teachers College Press, 1992. p. 153-181.
- LEE, Valerie E.; BURKAM, David T. Dropping out of high school: the role of school organization and structure. **American Educational Research Journal**, v. 40, n. 2, p. 353-393, 2003.
- LOPEZ DE LEON, Fernanda L.; MENEZES-FILHO, Naércio Aquino. Reprovação, avanço e evasão escolar no Brasil. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v. 32, n. 3, p. 417-451, 2002.
- MEKSENAS, Paulo. **Sociologia da educação**. 8. ed. São Paulo: Loyola, 1998.
- NERI, Marcelo. C. **O tempo de permanência na escola e as motivações dos sem-escola**. 1. ed. Rio de Janeiro: FGV/IBRE, CPS, 2009.
- OREOPOULOS, Philip. Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling. **Journal of Public Economics**, v. 91, n. 11-12, p. 2213-2229, 2007.
- PESQUISA SOBRE ABANDONO ESCOLAR. **Relatório**, Juiz de Fora, 2011. Relatório elaborado pela Coordenação de Pesquisa do Centro de Políticas Públicas e Avaliação da Educação/CAEd.

RUMBERGER, Russell W.; LIM, Sun Ah. **Why students drop out of school: a review of 25 years of research.** Santa Barbara: University of California, 2008. California dropout research project: Report 15.

STERN, Elizabeth; GLENNIE, Elizabeth, J. When and why dropouts leave high school. **Youth & Society**, v. 38, n. 1, p. 29-57, 2006.

WEHLAGE, Gary G.; SMITH, Gregory A. Building ner programs for studentes at risk. In: NEWMANN. Fred M. (Org.). **Student engagement and achievement in american secondary schools.** New Yourk: Teachers College Press, 1992. p. 92-118.

Received on August 31th, 2014

Approved on March 10th, 2015

Dr. Tufi Machado Soares has a PhD in mathematical theory of control and statistics from the Pontifícia Universidade Católica do Rio de Janeiro and a post-doctorate in statistics from UFRJ. He is an associate professor at the Department of Statistics and Graduate Program in Education. Coordinator of research of CAEd/UFJF – Center for Evaluation of Education Public Policy at the Federal University of Juiz de Fora.

Mariana Calife Nóbrega has an MSc in education from the Graduate Program in Education from the Federal University of Juiz de Fora (UFJF) and is currently a research assistant at CAEd/UFJF.

Neimar da Silva Fernandes is an undergraduate student in electrical engineering at the Federal University at Juiz de Fora (UFJF) and a research assistant at CAEd/UFJF.

Alexandre Chibebe Nicolella has a PhD in applied economics and is a professor of economic theory at the School of Economics, Business and Administration and Accounting at Ribeirão Preto – FEA-RP/USP.