

Duration of the Transition from Regular Basic Education to Higher Education in Peru

Ángel Renato Meneses Crispín¹, Pablo Alfonso López Chau Nava², Santiago Aquiles Gallarday Morales³

¹Universidad Privada Del Norte, <https://orcid.org/0000-0002-4496-9003>, E-mail: angel.meneses@upn.edu.pe

²Universidad Nacional De Ingeniería, <https://orcid.org/0000-0002-6250-0380>, E-mail: alopezch@uni.edu.pe

³Universidad Privada César Vallejo, <https://orcid.org/0000-0002-0452-5862>, E-mail: sgallarday@ucv.edu.pe

Abstract

The dynamics of the transitions from regular basic education to higher education are analyzed in order to measure and characterize the degree of disconnection between the two training stages. Survival methods were used to predict higher education entry rates for recent high school graduates, addressing the issue of data censorship. It has been found that about 30% of young people transition to higher education in the second year after completing high school, then the rate of admission grows over the years until reaching 67.1% at the end of the process. The risk of dropping out of full secondary status and moving on to tertiary education is higher if moving on to university education and privately run institutions. In addition, there is evidence of a continuous process with respect to educational management, since graduating from private schools offers greater probabilities of transition to HE of the same management. On the other hand, the greater human capital of the parents translates into greater human capital for the children, this is invariant to the type of higher education that the young people go through. Instead, the role of socioeconomic conditions turns out to be significant only for the transition to universities and private institutions, while cognitive skills fail to explain entry into private management. Additionally, gender gaps are favorable for women, a consistent result across the different models evaluated, while labor participation turns out to be less restrictive for entry to non-university education and private institutions.

Keywords: Transitions, Education, Human Capital, Gaps.

DOI: 10.47750/pnr.2022.13.S03.047

INTRODUCTION

The Regular Basic Education (EBR) and Higher Education (ES) systems in Peru have been identified by their great disconnection (decoupling) from one another; that is to say, there is no sequence that characterizes the first stage as a fundamental condition for the access and success of the second, even more so if one takes into account the criticisms about the academic rigor with which adolescents reach the last years of their studies. high school, which has generated the growth of programs and institutions that offer corrections and ways to facilitate the transition between EBR and SE, minimizing the natural barriers to entry to higher education institutions.

In relation to access, the evidence indicates that the expansion of higher education is a global phenomenon. The entrance of adolescents and young people to higher education has continuously expanded in the Latin American region for more than two decades. The gross enrollment rate at the tertiary level, a comparative indicator of countries defined as the proportion of people between 18 and 24 years of age enrolled in higher education, doubles from 21% in 2000 to 43% in

2013, the highest record. A recent data from the World Bank indicates that this figure reaches 52%, placing it at a distance of 23 points below that of countries with high incomes (75%). In this situation, Peru is among the countries with the highest growth in enrollment with gross rates at the tertiary level close to 71%, being one of the countries in the region that has experienced the highest growth in enrollment.

When inquiring about the local dynamics, it was revealed that the gross enrollment rate is not capturing the population made up of high school graduates who do not report enrollment data in the year of execution of the survey; that is, those who may have opted for some condition of work activity that excludes post-secondary educational decisions. This adjustment is important since, unlike high-income countries, in Peru, the labor option falls in an environment with very high rates of informality. With this adjustment, the percentage of young people between 17 and 24 years old who register a higher level is reduced to 30% according to the 2018 ENAHO.

Based on the above, it was intended to study the access gaps, in the most detailed and up-to-date way possible, taking into account the nature of the phenomenon under study; that is, the duration of the transition between the EBR and ES. To

quantitatively address this issue, survival methods were used as they are sophisticated techniques with great acceptance and virtues, among which the fact that they offer an efficient treatment for data censorship stands out, a very common problem that is usually present in databases. about educational transitions and that other techniques would not be able to model. The information from the longitudinal subsample of the ENAHO between the years 2014 and 2018 was analyzed.

BIBLIOGRAPHIC REVIEW

For many years, traditional methods such as linear regressions, logistic regressions, discriminant analysis, among others, have been used to analyze transitions between different educational levels; however, these models cannot deal with censored information, leading to suboptimal estimates and/or predictions [1].

A slightly more complex relevant modeling technique is survival analysis, a subfield of statistics that aims to model longitudinal data. One of the benefits of using survival analysis is the ability to add the time or duration component of an event to the model, as well as effectively handle censoring of the data; however, the use of these techniques for studies of educational transitions is limited.

It is more limited, even in the international literature, when it is intended to study the dynamic process from the completion of basic education to admission to higher education, since most of the efforts have focused on studying issues such as dropout and/or or desertion of young people in higher education [2,3,4]. Data and the nature of the economy and educational structure analyzed that do not correspond to developing countries such as Peru [5,6,7] were reviewed.

Consequently, the local documentation regarding the study of access to higher education is also characterized by the use of classical regression methods. Thus, admission to university and non-university higher education has been studied using multinomial discrete choice models, focusing on the transitions of young people towards the educational-labour world [8]. While, based on the analysis of the ENHAB 2010 (National Survey of Labor Skills, which represents the urban population of working age), using multinomial regression and correcting endogeneity problems, it focuses on finding the relative importance of cognitive skills on economic factors, educational histories, and cultural elements at home to explain enrollment in higher education. For their part, others have focused on revealing inequities in access to and completion of higher education [9,10].

The closest thing that could be found in the local literature is a study that characterizes the access gap at the different levels of training due to ethnic differences in the population, controlling for other main covariates. These authors propose a basic survival model for age cohorts, which does not allow information censorship to be modeled, losing the virtues of the applied method [11].

The lack of studies on the dynamics of access in the limits between EBR and SE in Peru motivated the need to develop predictive models that can optimally identify students who will potentially enter higher education, and will also allow knowing when said event will happen and What are the factors associated with it? It is in this area of the economics of education that survival models have established themselves as an appropriate tool, gaining acceptance as a quantitative method.

MEASUREMENT METHODOLOGY

Method

Regarding the present study, and what will be considered for the formalization of the methods, a process with simple duration and two states (initial and final) was analyzed. All individuals reported the same state of origin and the same state of destination with information censored from the right. The application of the model was on the study of the transitions between secondary education and higher education, where the event of interest is the entry of individuals to the tertiary stage of education. The initial state corresponded to young people who recently graduated from high school, and the final state is the entrance to higher education institutions. The duration corresponded to the recorded time (available in the longitudinal information) that elapses from the completion of the last year of secondary school onwards. Because of this, the censored information was associated with those observations of duration where only the initial state of completed secondary education is recorded; that is, those cases where it is not known at what point the individual will potentially register higher education status.

Suppose a data is represented by the finite sequence $B_n(t) = \{d_i(t), x_i, z_i(t), T_i; i = 1, \dots, n\}$ where $d_i(\cdot)$ is a binary variable that identifies whether the observations are censored, x_i a vector of time-invariant covariates, $z_i(\cdot)$ a set of time-varying covariates, and T_i denotes the time of the observed event. Since $d_i(\cdot)$ it defines two subsets of data, one in which individuals manage to change state (uncensored), and another where they do not (censored), suppose that U_i is the survival time of the event that could be unobservable, and let be C_i the time elapsed in the censored data or the last follow-up time of the individual, then the observable duration of the event is in general $T_i = \min \{U_i, C_i\}$. Explicitly, the censoring variable is defined as,

$$d_i = \{1, U_i \leq C_i \ 0, U_i > C_i$$

Note that $T \geq 0$ is a continuous random variable and t a particular value of T . The survival function $S(t)$ is defined as,

$$S(t) = P(T \geq t) = 1 - F(t) \dots (1)$$

The survival function indicates the probability that the duration of the event T is equal to or greater than the value of t so that it turns out to be equal to the complement of the cumulative probability distribution $F(t)$, which comes from

its density function $f(t) = dF(t)/dt$. The distribution of the duration of an event can also be characterized in terms of the hazard function (hazard ratio or risk function), the one that expresses the probability of change of state after t units of time, conditional on it surviving in the initial state up to that period. In other words, the Hazard ratio is the instantaneous rate of occurrence of an event,

$$h(t) = \frac{P(t \leq T < t + \varepsilon)}{\varepsilon} = \frac{f(t)}{S(t)} \dots (2)$$

Survival models focus attention on this function, since it will represent the dependent variable reported for its interpretation.

Database and Variables

The main source was the National Household Survey (ENAHO) published on the website of the National Institute of Statistics and Informatics (INEI), this base is the most representative in terms of the study of the socioeconomic conditions of households in Peru. The ENAHO has a longitudinal sample subset, since it monitors a group of households and their members.

In addition to the ENAHO, complementary bases have been processed such as 1) National Survey of University and University Graduates (ENEUU), 2) School Census (CE), 3) Student Census Evaluation (ECE). From these, variables on the quality of higher education institutions (1 and 2) and the cognitive abilities of students (3) can be obtained at different levels of aggregation. However, when these data have been extrapolated, the coincidences with respect to bases 1 and 2 have not been as expected, so the variables that characterize the final state (higher educational level) of the young people have been omitted from the analysis; Meanwhile, it has been possible to recover the performance indicators of the students in the secondary stage, through the academic scores in Letters and Mathematics.¹

On the extrapolated basis, the questions from the ENAHO education module were used. There are two main groups that made up the valid observations for the final sample: a) the first is validated from the question about enrollment in the year of the survey, that is, those young people who are enrolled in the last year of high school ; b) the second is validated by means of the question on maximum education achieved for those who did not answer in the first case, of these there are at least two sub-cases: those who have completed high school and are at an early age at the end of said stage, and those who have higher education at an early age at the conclusion of the secondary stage; for the latter, the minimum age for completing secondary school is assumed as the start date of the duration. This means that the important thing for the size of the sample is that one of the two states (secondary and/or higher) is registered at the opportunely selected age that prioritizes the group of young people in the

time horizon with high probabilities of entering the higher education, a maximum of 24 years of age, aligning with the calculation standards for admission rates to HE in official reports.

With this consideration (addition of group b), it was possible, at least, to double the sample, which allowed, on the one hand, to considerably improve the inference of the reports and, on the other, to obtain a greater variety of results using different subsamples. In total, 4,688 recent high school graduates between 15 and 24 years of age are followed up, in the 2014-2018 time horizon, of which 41.7% register admission to higher education considering the data pool.

The decoupling between EBR and SE is captured through the study of the entry rate to higher education, and is represented by the time that elapses (duration) since the completion of secondary school, where the conditions that identify to the two sample groups.

In addition to the dependent variable, in the survival data, data censorship indicators must be constructed, in this case binary variables that indicate whether or not the individual has experienced the transition to higher education, that is, information is considered censored to observations where only the initial state is recorded. The size of the sample allows different possibilities to be evaluated through different censorship indicators, so that entry to university higher education, non-university higher education, public higher education, and private higher education will be modeled.

The presence of available independent variables that guarantee a good structure in the modeling of the duration of the transition between EBR and ES has been prioritized: geographic and gender controls (initial or final state), school background such as educational management and the skills of students, human capital of parents (time average), socioeconomic strata (initial state), and working conditions (final state).

CHARACTERIZATION OF THE PERUVIAN CONTEXT

From the longitudinal structure, the final sample at the level of individuals amounted to 4688 young people between 15 and 24 years of age who have graduated from secondary education, of which 41.7% register the two educational states (complete secondary and superior), the remaining 58.3% is considered potentially censored information. Shows the description and summary of survival data. The duration is analyzed by years, the maximum duration in the sample is 7 years and the average duration is 2.7 years, an average that is equal to the mean time at risk, since the input time for all observations is naturally zero. The total time at risk is the sum of all the durations (records) registered in the sample, in this case, it takes the value of 1274.

This last piece of information made it possible to calculate

¹The methodological exercise of joining the bases continues to be useful for future research since the sample can be increased in the

process of joining the bases, which suggests finding alternative identifiers in the respective bases.

the incidence ratio indicator, which is the quotient between the number of exits (changes in state) and the total time at risk. For the analyzed period, it rose to 15.3% and, if we consider some covariates, it was found that the incidence ratio is higher among urban residents, women, graduates of private schools, people from households with a high socioeconomic level, educational level head of household, and for those young people who are out of work. The distribution of the sample in the different covariates was appreciated, and on the other hand the main percentiles of the duration variable that in median value give the first indications of how the main variable (duration) is conditioned through the variables. independent variables indicated.

Kaplan-Meier Descriptive Method

In the first year after finishing high school (initial status), the probability of remaining in the same status is 83.5%. In other words, the probability of changing from secondary to higher status is 16.5% one year after graduating from school. By property of the survival curve, said probability of remaining in the initial state is decreasing and mathematically it converges to zero as time goes by. According to the data, the probability of surviving (remaining in the secondary state) is reduced to a greater extent in the first two years ², reaching 69.1% on average, after that time, the survival rate reductions are between 6 and 8 percentage points. It is expected that the vast majority of young people who are going to transition to higher education do so up to 24 years of age or 7 years of duration, counting from the year they graduate from high school. Likewise, a percentage of young people have preferences for only working or studying careers not cataloged as higher education (for example, one-year programs in technological institutes, cetpros, etc); For these two reasons, the statistical convergence of the probability of surviving in the secondary state is a value below but close to 32.8%; that is, a rate of admission to higher education between 64.6% and 69.7%, compared with the tabulated panel (41.7%) it follows that the Kaplan-Meier calculations are correcting a censorship of at least 20 percentage points on the rate of income ³.

The different end states can be non-university higher education, university higher education, public higher education and private higher education. ⁴In general, transitions to university education and private institutions occur more quickly.

Going back to the initial censoring model, the survival function can be controlled by different covariates, in this case, non-parametric; It is interesting to appreciate the

differences by origin of educational management in the school stage. It was observed that those young people who come from private schools are less likely to survive in the secondary state, otherwise it means that individuals from private schools are more likely to transition to higher education as time goes by or that the transition done in less time. Thus, in the first year after high school, graduates of private secondary education recorded 10 percentage points more than their peers graduates of state education; in the second year, the distance increases to 15 percentage points and, in the following two years, the situation is not so different.

Cox Semiparametric Method

Once the descriptive model of the transition to higher education has been analyzed, the dependent variable is modeled, in this case, the hazard ratio or risk function or marginal rate of transition, based on a group of covariates, mostly categorical. The results of the Cox semiparametric model are reported, also considering the stratification by school management. Previously, the proportionality tests were validated ⁵, which allow the main assumption of the model to be met.

The results indicated that stratifying the model does not yield such different results. For this reason, it has been preferred to consider from now on the model without school management strata, therefore, placing this variable as one more regressor in the specification. The reported hazard ratios indicated not only the magnitude of the effects of the covariates (equal to 1 minus the hazard), but also indirectly indicate the sign, for those values less than unity it means a negative effect, and for values greater than or equal to the unit positive effect. Thus, the school background obtained a level of significance at 1%, and its ratio is 1.23, this means that in young people who have studied in private schools, compared to those who studied in public schools, the risk of leaving is increased. or transition to higher education in 23.2%. The risk of exiting from the secondary state to the higher education state decreases for men, otherwise, men present a 12% less probability than women in relation to the rate of transition. On the other hand, being a resident of urban areas improved the insertion towards higher schools in the order of 19.3%.

The socioeconomic conditions of the young people's homes, in short, had a positive role in improving the possibilities of access to higher education. This is manifested in a hazard greater than one, although the data is only significant in the comparison between the high and low strata. Regarding the abilities in the school stage, the statistical significance was

²The non-linearity of the survival curve cannot be appreciated by the annual frequency of the data.

³The calculations are also reasonable with other aggregate data based on cross-sectional information, such as the case of the annual ENAHO 2014-2018, where the aggregate admission rate, in the range between 17 and 24 years of age, is around 60%.

⁴For example, in the first censorship model, the transition from complete secondary education to non-university higher education is analyzed. Regardless, the other models are analyzed in the same way.

⁵Residual Scaled Schoenfeld tests were performed.

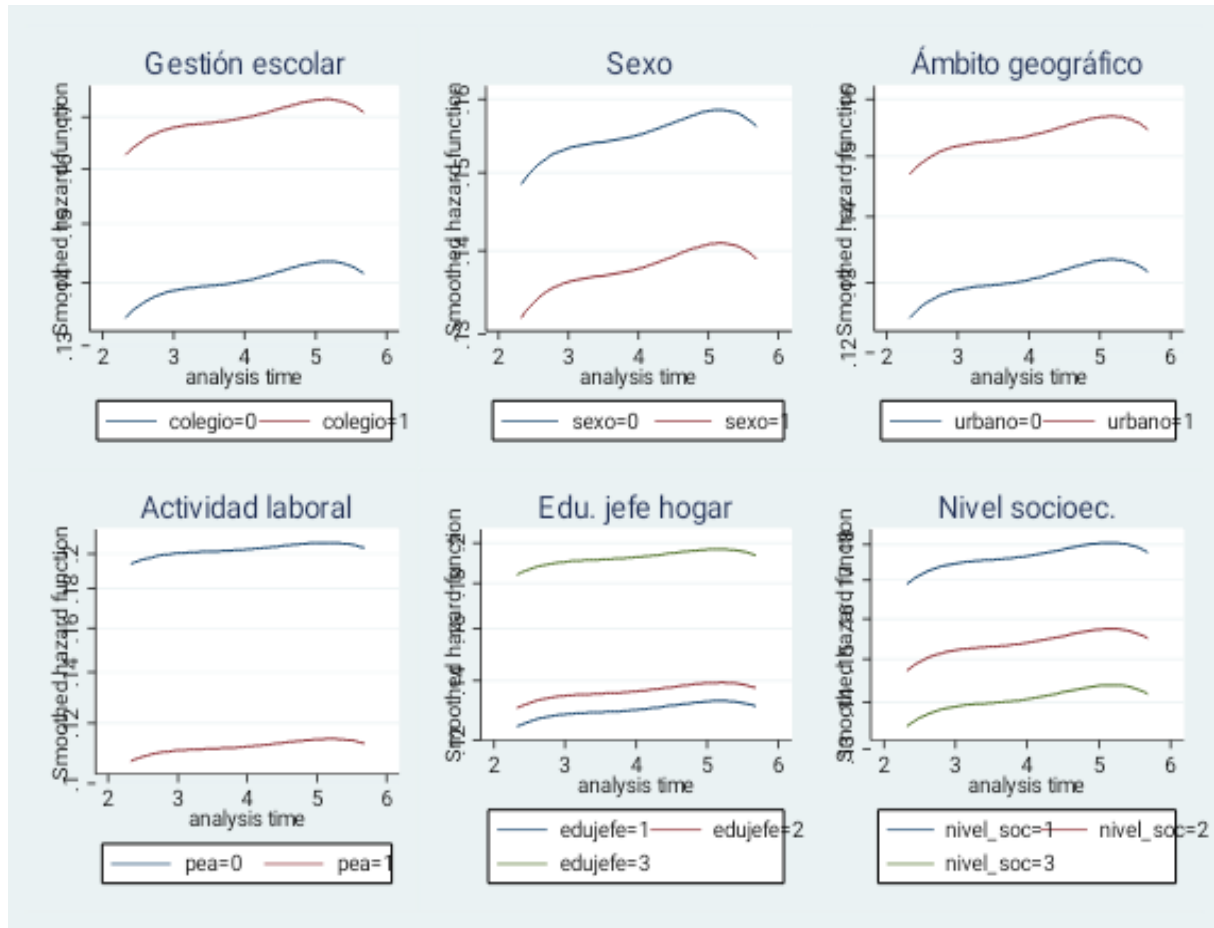
only revealed for the grades in mathematics, although the sign is as expected, very little can be said about the impact on the risk of leaving. On the other hand, the human capital of those responsible for the household (most likely the parents) was conclusive. Indeed, the higher educational levels of the heads of households were directly related to greater probabilities of transition. The estimation showed that what is relevant for young people is that their parents attain higher education and not a lower level, since this would allow secondary school graduates to increase their chances of transition by up to 54% compared to a primary education situation of parents, and 46.3% if the parents only had secondary education, which reflects the growth of human capital of young people based on the human capital of household heads.

Preferences for work, after graduating from school, decreased the hazard ratio by up to 47.8%. This sign is capturing the

substitution of efforts involved in entering a higher school, since it is required to disaccumulate work assets to achieve this purpose.

The results presented can be expanded qualitatively with the following graph, which shows the hazard ratios smoothed over time and considering the different controls. Here it is necessary to emphasize that the term "increasing" is used in the sense of the ordered units of duration of individuals. There are then different smooth profiles of the hazard ratio and with greater notoriety the effects of the regressors on the dependent variable. As the years go by, the transition possibilities of all young people increase, but unevenly, which is reflected not only in the distances between the categories of each variable, but also in the slopes of the risk function.

Graph N° 1: Hazards smoothed in the Cox model in the sample of young recent graduates of secondary education according to variables



Source: ENAHO 2014-2018

Own elaboration.

Finally, for different censorship models, some interesting comparisons were made. Thus, to transition to the higher

technical level, geographical residence, socioeconomic status and education of the head of household did not turn out to be

significant in explaining the changes in the probabilities of transition of young people; but it is for the university case where the variable that most reacts to the risk of leaving is the higher education of the parents, the socioeconomic level and private school management, in that order. Likewise, finishing high school in a private school was more familiar with transitions to university than to technical education, as shown by the negative (positive) effect of the coefficient of said variable in the case of the non-university level (university). In numbers, leaving a private school reduced the probability of going on to technical schools by a third, while it increased them by 66.4% for university education. Another detail is that the work activity, during the transition to HE, slowed down, to a greater extent, the possibilities of studying university higher education compared to studying at a technical level, the difference in the hazard ratio is around 26 percentage points, which would otherwise be pointing out that the individual's preferences for work and study are more acceptable in non-university education.

Regarding higher management, the continuity of private management from the school stage to the higher stage was corroborated as the hazard is significant in the case of private higher institutions, not changing position in educational management is in the order of 48 %, while the non-statistical significance in the public case reinforced the same idea. Another noteworthy point was that, at the aggregate level, mathematical skills turn out to be relevant in the transition to public institutions and not so for private ones; The opposite effect was revealed in the socioeconomic level, which is a significant variable for the risk of leaving for private management and, relatively, it is not for public education; but what they have in common in terms of effects is the role of parental education in explaining transitions, while work was correlated with study to a greater extent in the case of exits towards private training, which could reflect the need to supplement education expenses not covered by those responsible for the young people in their homes.

CONCLUSIONS AND RECOMMENDATIONS

New evidence of the decoupling between EBR and ES has been shown by analyzing the transitions between both educational stages through more sophisticated methods such as survival analysis. A variety of results are observed on the rates of entry into higher education by young recent graduates. The probability of entering higher education is around 30% in the first two years after graduating from high school, increasing over the years and converging at entry rates between 67% and 70%. The risk of abandoning the complete secondary condition and going on to tertiary education is greater if one goes to university education (10.0%-39.4%) and to private management institutions (10.9%-49.0%).

The factors most used by the literature turn out to be significant to control the duration of the transition between

EBR and ES. Thus, the access gap by gender has not only been closed, but now it is women who have greater probabilities of entering a higher career, also involving a shorter duration in the process to enter HE, this result being more notable when access is to non-state education. On the other hand, a continuous process is observed in terms of preferences for private education, as it is a study center that offers greater probabilities of transition to HE of the same management as the years go by. The role of parents' education is crucial since they turn out to be significant regardless of the type of higher education to which young people go; what does not happen with the role of socioeconomic conditions and cognitive skills that are key to differentiatingly explain the dynamics of admission to private HE and public HE, respectively.

Regarding the recommendations for future work, it is suggested to analyze the transitions to higher education seeking to increase the frequency of the durations to appreciate non-linear behaviors in the survival curves; To this can be added the idea of incorporating new explanatory variables in the search to find differentiated results for the different final states in higher education.

Regarding modeling, two important advances can now be followed: the first has to do with the study of transitions with final states that combine educational and work activities and the second consists of analyzing the performance of higher education with the approach of transitions to explain relevant issues such as abandonment and/or delay in higher education.

REFERENCES

- Lin, J., Imbrie, P., Reid, K. (2009). Student retention modelling: An evaluation of different methods and their impact on prediction results. *Research in Engineering Education Symposium*, pages 1–6.
- Tinto, V. (2006). Research and practice of student retention: what next? *Journal of College Student Retention: Research, Theory & Practice*, 8(1):1–19.
- Kellogg, P., Radcliffe, R., Huesman, D. (2006). Identifying students at risk: Utilizing survival analysis to study student athlete attrition. In *National Symposium on Student Retention*, Albuquerque, NM, pages 1–11.
- Kellogg, D., Jones-White, P., Radcliffe, R., Huesman, J. (2010). Redefining student success: Applying different multinomial regression techniques for the study of student graduation across institutions of higher education. *Research in Higher Education*, 51(2): 154–174.
- Kim S, Chang M, Park J (2017). Survival analysis for Hispanic ELL students' Access to postsecondary schools: discrete model or Cox regression? *International Journal of Research & Method in Education*.
- Berkowitz, R., Jacobs, J. (2002). "Age and College Completion: A Life-history Analysis of Women Aged 15–44." *Sociology of Education* 75 (3): 211–230.
- Tseng, Y., Polidano, C., Tabasso, D. (2015). "A Second Chance at Education for Early School Leavers." *Education Economics* 23 (3): 358–375.
- León, J., Sugimaru, C. (2013) Between study and work: The decisions of young Peruvians after completing regular basic education. *Advances in research, education and learning*. Development Analysis Group.
- Castro, F., Yamada, G. (2011). Higher Education Decisions in Peru: On the Role of Financial Constraints, Skills and Family Background. Discussion Paper. Lima: Pacific University Research Center.

- Benavides, M., Etesse, M. (2012). Intergenerational educational mobility, higher education and social mobility in Peru: recent evidence from household surveys. Book: Higher education, social mobility and identity. Institute of Peruvian Studies.
- Castro, F., Yamada, G. (2010). Ethnic and gender differences in access to basic and higher education in Peru. Lima: Research Center and Department of Economics of the Universidad Pacífico.