Inequality in Access to Tertiary Education – Evidence from El Salvador

Bachelor Thesis

Livia Lea Jakob

13-131-578

livia.jakob@giub.unibe.ch

Student BSc in Geography

2017

Supervisor: Dr. Andreas Heinimann

Unit Geography of Sustainable Development, Institute of Geography,
Faculty of Science, University of Bern

Abstract

Research examining how social background and individual characteristics influence access to education has mostly focused on developed countries and disregarded spatial context. This is striking as sociospatial research in low- and middle-income countries can provide the scientific basis for an egalitarian development process.

The aim of this paper is to examine how spatial and socioeconomic factors shape inequalities in access to postsecondary education in the Latin American country El Salvador. The study conducts a logistic regression analysis with multiple predictors using data from the 2015 Salvadoran household survey Encuesta de Hogares de Propósitos Múltiples (EHPM). To include spatial context the number of accessible postsecondary institutions and study programs in different time radii are calculated.

The results indicate that individuals from low-income families and children of parents with lower educational attainment levels have lower propensities to participate in tertiary education. Furthermore, distance to tertiary educational institutions and urbanization (rural/urban) both play a role in determining study participation patterns. Individuals seem to be positively affected by each additional study opportunity up to a travel time radius of approximately two hours and the influence of spatial context tends to decline with increasing travel time radii. The findings suggest that research examining educational participation patterns should refrain from treating educational decisions as though they were unrelated to geographical context.

Content

1 Introduction 5

2 Theoretical Framework 8

2.1 Inequality in Educational Decisions 8
2.2 Formulation of the Theoretical Framework 10

3 Contextual Framework 12

3.1 Educational System in El Salvador 12
3.2 Supply of Postsecondary Institutions in El Salvador 13

4 State of Research and Hypotheses 15

4.1 Spatial Context 16
4.1.1 Regional Supply of Educational Institutions 17
4.1.2 Level of Urbanization 17
4.2 Socioeconomic Characteristics 19
4.2.1 Family Income 19
4.2.2 Social Class Origin and Intergenerational Mobility 20
4.2.3 Gender 21
4.3 Group-specific Differences in Sensitivity to Spatial Context 22
4.3.1 Family Income and Distance 22
4.3.2 Social Class Origin and Distance 22
4.3.3 Gender and Distance 23

5 Data and Methods 24

5.1 Data 24
5.2 Travel Distance Measurements 25
5.2.1 Previous Attempts to Measure Accessibility 25
5.2.2 Method to Measure Accessibility 27
5.3 Statistical Evaluation 29
5.3.1 Base Population and Sample 29
5.3.2 Logistic Regression Models 31
5.3.4 Variable Specifications 32

6 Results 33

7 Discussion 38

7.1 Spatial Context 38
7.2 Socioeconomic Context 39
7.3 Group-specific Differences in Sensitivity to Spatial Context 39
7.4 General Limitations 40
7.5 Policy Implications 41

8 Conclusion 43

Literature 45

Appendices 49

1 Introduction

Education is an indispensable building block in a sustainable development process. Being a key ingredient to economic productivity, it can be a powerful tool in the eradication of poverty. Moreover, education is of fundamental social importance. Enabling people to actively participate in society and its transformation process, good education fosters social inclusion and strong institutions (World Bank 2016). Nonetheless, as important as the overall level of education in a region is its distribution. Inequality in access to education not only results in an unequal distribution of education, it also tends to reproduce social classes and income gaps.

There is a large body of literature assessing how socioeconomic factors influence educational decisions. Typical findings are that the education and status of parents, gender and household income shape educational participation patterns. While the effects of many different socioeconomic background variables on educational participation have been widely researched in developed countries, there have been surprisingly few attempts to study the manifestation of these effects in Latin America. This gap in research is striking as Latin America is known as the most unequal region in the world in terms of income inequalities (Deininger and Squire 1996: Table 5; Hertz et al. 2008; OECD 2008; Brunori, Ferreira and Peragine 2013; Torche 2014). This fact makes it even more pertinent to consider whether these inequalities in Latin America persist across generations. Indeed, several recent studies have reported that the lowest intergenerational mobility regarding educational attainment is found in Latin America, i.e., children there are most likely to maintain the educational status of their parents (e.g., Hertz et al. 2008; Brunori, Ferreira and Peragine 2013).

Another issue, which has hardly been examined in relation to developing countries as well as in general, is the relation between spatial context and study participation. Research concerning educational participation patterns has typically been conducted by using decision models regardless of the geographical context. However, different authors have recently argued that the process by which students decide whether and where to attend tertiary education should not be examined without including spatial context (e.g., Turley 2009; Wessling 2016). Empirical findings show that the regional supply of education (Frenette 2006, 2004; Spiess & Wrohlich 2010), regional labor market conditions (Flannery & O'Donoghue 2009; Rizzica 2013) and the level of urbanization (Sá, Florax and Rietveld 2004) influence study decisions. In addition, some authors argue that spatial context is not equally relevant to different social groups. Empirical results, for instance, reveal that students from lower social backgrounds seem

to be more disadvantaged by the distance to school (Eliasson 2006; Frenette 2006; Cullinan *et al.* 2013; Wessling 2016).

Again, almost every study assessing educational participation patterns linked with spatial context has been conducted in developed countries, for instance, due to data limitations for developing countries. To the best of my knowledge, so far, no study examining how spatial characteristics influence study decisions has been made in the region of Latin America. This research gap should be filled in order to enhance the impact and efficiency of interventions aimed at reducing inequalities. Along these lines, Frenette (2004) has emphasized the importance of a deeper knowledge of how geographical barriers influence postsecondary participation patterns for policy decisions. He points out that understanding the geographical access gap and knowing whom is the most disadvantaged by geographical barriers provides an important base for evaluating student financial assistance programs. Above all, the poor infrastructure and the lack of public transport, which are often found in developing countries, result in longer travel times and thus may result in a larger role being played by distance.

In light of the above-mentioned research gaps, this thesis considers how spatial and socioeconomic characteristics shape inequalities in access to education in El Salvador, as an example of the Latin American region. The aim of this study can be divided into three research questions.

(1) Spatial context:

How do spatial and contextual characteristics influence access to tertiary education? This question not only addresses the issue about whether geographical context plays a role in educational decisions, but also the scope within which spatial context matters to an individual.

(2) Socioeconomic context:

How do socioeconomic characteristics influence access to tertiary education? This aims to examine how individual characteristics and family background shape postsecondary participation patterns.

(3) Sociospatial interactions and linkages:

For whom does spatial context matter in terms of access to tertiary education? This research question addresses group-specific variations in sensitivity towards spatial and contextual characteristics.

The questions will be answered for El Salvador, a small country in Central America. The region represents an interesting context for examining inequalities in education, since public expenditure in education and efforts to increase educational attainment are low compared to surrounding countries (Bashir & Luque 2012). The study focuses on the participation to tertiary education for two main reasons. First, since the number of educational institutions significantly reduces from secondary to tertiary education, spatial context can be better investigated. Second, as in the rest of Central America, primary and secondary education attains significant higher levels of coverage (Bashir & Luque 2012), while access to tertiary education becomes more selective. As a theoretical foundation, Boudon's (1974) decision model, along with the extension of the Breen-Goldthorpe model (Breen & Goldthorpe 1997), will be used, complemented by the author's own adjustments. The study conducts a logistic regression analysis with multiple predictors using data from the 2015 Salvadoran household survey *Encuesta de Hogares de Propósitos Múltiples* (EHPM).

Besides addressing the three research questions above, this study strives to achieve another goal. As discussed above, research concerning educational participation patterns has typically been conducted regardless of the geographical context. However, different researchers have recently included spatial characteristics, such as the distance to educational supply, in their analyses (e.g., Frenette 2004, 2006; Eliasson 2006; Spiess and Wrohlich 2010). From a geographical perspective, the attempts to include spatial context seem to have been highly simplistic, meaning that there is still room for improvement. In this contribution, the inclusion of spatial context for studies dealing with educational patterns will be discussed and previous attempts will be criticized. Further, the technique and geographical tools for modeling accessibility in this study will be described precisely. The method can be viewed as suggestion for modeling spatial context in future research in this field.

The thesis is structured as follows. The next section provides a theoretical background about educational decision models and formulates the underlying theoretical framework for this study. Section 3 covers the contextual framework, outlining the educational system and supply in El Salvador. Section 4 presents an overview of the current state of empirical research and derives hypotheses from these previous findings. Section 5 explains the data and methodology. Section 6 presents the main findings, while the results, limitations and policy implications are discussed in Section 7. The last section concludes.

2 Theoretical Framework

This section presents a brief overview of the theories about inequalities in terms of educational decisions and formulates the underlying theoretical framework for this study.

2.1 Inequality in Educational Decisions

Individual decisions are often explained by rational choice theories. The human capital theory (HCT) (Mincer 1958; Becker 1994) claims that investments in human capital, such as education or training, are based on the calculation of costs and benefits. On the one hand, education and training increase workers' productivity and in turn their future income and occupational prestige (*benefits*); on the other hand, to reach a higher educational level, monetary investments and effort in learning must be made (*costs*). The rational choice is the possibility that maximizes the utility of a person when evaluating costs and benefits. Differences in rational decisions related to educational participation are taken to result from variations in individual ability and opportunities (Becker 1994).[1] The HCT has been criticized widely, yet many sociologists have adopted its general presuppositions and extended the approach in different ways.

Probably the most prominent sociological approach for modeling educational decision-making can be traced back to Boudon (1974). Extending the HCT approach in his book *Education, Opportunity, and Social Inequality*, Boudon suggests that preconditions and cost-benefit evaluations vary between social groups, while distinguishing between *primary* and *secondary effects* that shape class-specific differences in educational attainment. The primary effects of social background involve differences in conditions of socialization that affect children's cognitive abilities and motivation towards learning. By way of illustration, when growing up, children from a higher social class origin tend to be equipped with more educationally relevant objects, such as books, and thus enter the educational system with different cognitive starting conditions. Boudon refers to *secondary effects* as the class based differentials in educational decisions, even when students do not differ in their cognitive abilities or performance (*primary effects*). While the HCT argues that equal opportunities and abilities must lead to equal educational

[1] For a productive overview on the history and limitations of the HCT within the field of education, Gillies (2015).

decisions, Boudon argues that children's ambitions and decisions depend on social background.

Following Boudon's general principles, several extensions and reformulations of his model have been developed (e.g., Gambetta 1987; Erikson and Jonsson 1996; Breen and Goldthorpe 1997; Esser 1999). In terms of class-specific educational decisions, the Breen-Goldthorpe model (Breen & Goldthorpe 1997) is probably the most popular extension of Boudon's ideas. While Breen and Goldthorpe adopt Boudon's principles of *primary* and *secondary effects*, their focus is on the latter. They suggest that educational decisions are based on a simple rational choice model evaluating (1) costs, (2) benefits and (3) probability of success. In all three dimensions, social classes are assumed to differ in their typical evaluations. In relation to costs (1), this means that similar training costs create a higher burden for lower-income families (Breen and Goldthorpe 1997: 286). For the evaluation of benefits (2), the concept of status attainment plays an important role. Breen and Goldthorpe argue that parents primarily try to avoid downward social mobility, meaning that they do not want their child to reach a lower level of education or a lower social status than their own. The underlying assumption is that loss of status harms someone more than they benefit from status enhancement. Therefore, children with higher social backgrounds tend to see greater benefits in continuing their education (Breen and Goldthorpe 1997: 283ff). For the last factor (3), Breen and Goldthorpe suggest that the evaluation of chances of success also varies between social classes. Due to negative experiences or poorer school performance (e.g., because of *primary effects*), children with a lower level of parental education tend to evaluate their chances of succeeding in an educational program as being fewer than those from more highly educated parents (Breen and Goldthorpe 1997: 285f). According to the Breen-Goldthorpe model, then, the stated differences in evaluating these three factors indicate differing educational decisions between social classes. The Breen-Goldthorpe model has been empirically tested multiple times and, to some extent, seems to be suitable for examining educational participation patterns (e.g., Davies, Heinesen and Holm 2002; Breen and Yaish 2006; Stocké 2007).

Class-based differentials in educational outcomes are deeply related with patterns of intergenerational mobility. The term *intergenerational mobility* refers to changes in key characteristics and outcomes for individuals between one or more generations within the same family. Low intergenerational mobility means that children have a high tendency to acquire the same social attributes, e.g. levels of educational attainment, as their parents. The Breen-Goldthorpe Model provides an explanation for the tendency of inequalities to persist across generations.

2.2 Formulation of the Theoretical Framework

The rational choice approach of Boudon, along with the extension of Breen and Goldthorpe, will be used as a theoretical framework for this study, alongside the consideration of two limitations.

First, the focus of the Breen-Goldthorpe model and the empirical studies confirming the model lies on prosperous societies. The model describes educational *decisions*, based on the presupposition that students can decide whether they prefer to continue studying, which are affected by their social background. Breen and Goldthorpe do not take into account whether or not there is the opportunity to continue studying. In a context like Latin America, for low-income families, it may simply be unaffordable to send their children to a university, especially when there are large distances to the next educational institution.

This brings me to my next point. Besides social contexts, geographical contexts should also be explicitly included when examining educational participation patterns. The process by which students decide whether and where to attend tertiary education depends significantly on geographical characteristics. It can be assumed that effects similar to the *primary* and *secondary effects* of social background also exist in a geographical context. *Primary effects* regarding spatial context may involve differences in availability and the quality of establishments promoting education, such as libraries and schools, between spatial areas. These differences result in different cognitive abilities and motivation for learning. *Secondary effects*, in turn, may be interpreted in terms of the spatially dependent evaluation of costs, benefits and probability of success. In addition to the influence of costs (monetary and emotional costs), long travel distances or the inevitability of moving away to study could discourage students. Furthermore, spatial characteristics, such as regional labor market conditions, can also shape a student's estimation of benefits. By way of illustration, an individual living in a rural environment, where agriculture is carried out primarily for subsistence and the necessary knowledge can be acquired without a long formal education, will evaluate benefits differently than someone living in an area with higher non-farming employment.

Summing up, the theoretical framework of this study takes educational decision-making to be representable with a simple rational choice model, including costs, benefits and probability of success. Preconditions and the evaluation of these three factors both depend on social background and geographical context. As a consequence, inequalities in educational participation occur. Finally, in the Latin American region, it cannot always be assumed that students decide about whether to continue school, given that, for low-

income families, there may be no affordable opportunity. Note that, for the sake of linguistic simplicity, the term *decision* will nonetheless be used in this study.

3 Contextual Framework

This section provides an overview of the region of study, outlining its educational system and supply of tertiary educational institutions.

3.1 Educational System in El Salvador

The educational system in El Salvador can be referred to as a one-track system, meaning that students are not divided into different classes according to individual performance. *Figure 1* summarizes the school system in El Salvador. The first level (primary school) covers the first nine years and is split into three cycles, each of which lasts three years. Contrary to many other countries, any student who graduates from primary school can then transfer into secondary education. Students can choose from two types of secondary education: general high-school education (*educación media general*) takes two years to complete, whereas technical high-school education (*educación media técnico*) lasts three years. The additional year stems from the fact that technical high-school education offers a more practice-oriented program, with focus on a specific subject (e.g., accounting, mechanics, agriculture or tourism). Despite this focus, technical high-school education cannot be put on the same level as vocational training. After completing high school, students basically have two choices: (1) entering the labor market or (2) entering tertiary education. There is no alternative education, such as vocational training and education, which implies that anyone who does not complete tertiary education is left without vocational qualifications. Simply put, postsecondary institutions (universities, technical institutions and specialized institutions) offer three different study programs: (1) technical training (*técnico*, two to three years) (2) applied sciences (*tecnológico*, four years) or (3) an academic degree (*licenciatura/ingeniería*, five years). In this study, the three study programs are all treated as either tertiary or postsecondary education (United Nations Educational, Scientific and Cultural Organization (UNESCO) - IBE 2010).

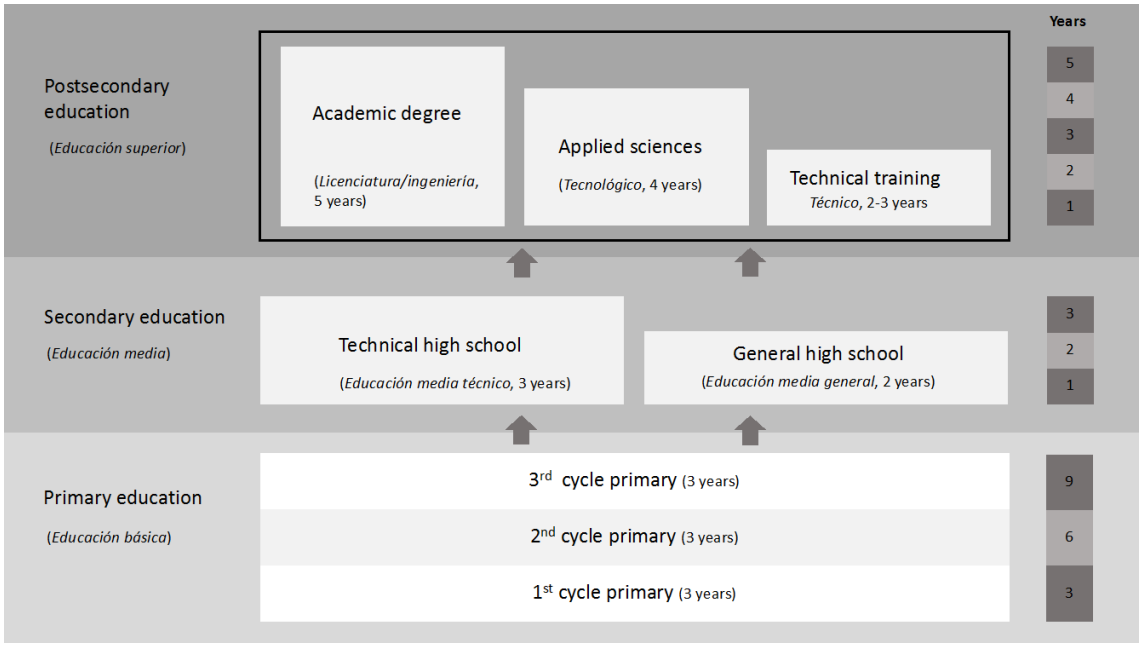


Figure 1: Educational system in El Salvador; representation according to data provided by the United Nations Educational, Scientific and Cultural Organization (UNESCO) - IBE (2010).

3.2 Supply of Postsecondary Institutions in El Salvador

Figure 2 shows the distribution of postsecondary institutions across El Salvador. In this country, postsecondary institutions can be divided into universities and technical institutions. In this study, as they both offer postsecondary education, they are both treated the same. As shown in the illustration, there is a wide gap between the number of public and the number of private postsecondary institutions. The University of El Salvador is the only public university in the country and includes the main campus located in San Salvador and three branches, which are spread throughout the country. In addition, there are about 42 private establishments for tertiary education. The fees and matriculation costs for public university add up to less than $5.30 per month, regardless of the chosen study program. Meanwhile, for private establishments, average fees for the least expensive study programs are about $100 per month (*Appendix 1*).[2] This gap represents an interesting structure for analysis, since it leads to larger distances or travel times for those who cannot afford to attend private educational institutions.

[2] Note that, in some postsecondary institutions, including the University of El Salvador, study fees depend on the socioeconomic status of each student's family. As such, calculations were made by computing the most inexpensive variant.

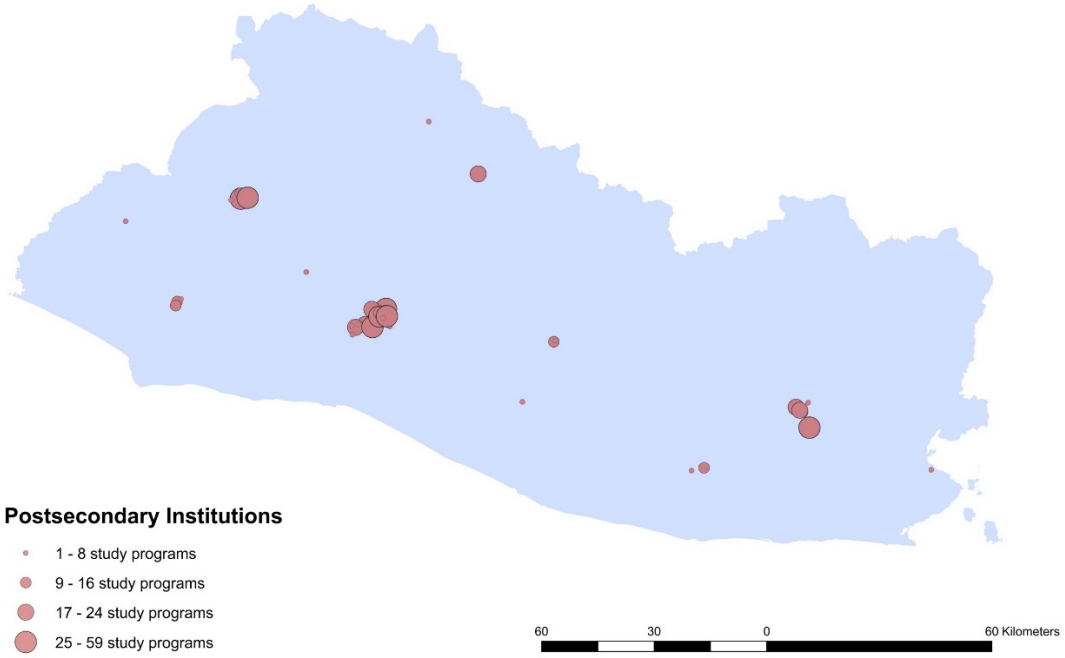


Figure 2: Distribution of postsecondary institutions across El Salvador, with symbol sizes being proportional to the number of study programs offered.

4 State of Research and Hypotheses

In the following, the results of previous studies will be discussed and, in turn, the hypotheses for this study will be derived. *Table 1* summarizes the hypotheses. The next section is divided into the three subsections, analogous to the three research questions: (1) spatial context, (2) socioeconomic context and (3) group-specific differences.

	Independent variable	**Mediation**	**Dependent variable**	**Mechanism**	**Variable type**
1a	↑ Regional supply of tertiary educational institutions		↑ Participation in tertiary education	Opportunities, reduction of costs	Spatial
1b	↑ Regional supply of tertiary educational institutions	↑ Travel time radii	↓ Effect on participation in tertiary education	Scope of spatial context	Spatial
1c	Rural		↓ Participation in tertiary education	Differences in sociospatial characteristics between rural and urban areas	Spatial
2a	↑ Household equivalence income (log)		↑ Participation in tertiary education	Costs	Family
2b	↑ Parents' highest educational attainment		↑ Participation in tertiary education	Intergenerational mobility	Family
2c	Gender (female)		↑ Participation in tertiary education	Gang culture, criminality, gender roles	Individual
3a	↓ Household equivalence income (log)	↑ Distance to nearest tertiary institution	↓ Participation in tertiary education	Costs	Family * spatial
3b	↓ Parents' highest educational attainment	↑ Distance to nearest tertiary institution	↓ Participation in tertiary education	Intergenerational mobility	Family * spatial
3c	Gender (female)	↑ Distance to nearest tertiary institution	↓ Participation in tertiary education	Females respond more strongly to contextual factors	Individual * spatial

Table 1: Summary of research hypotheses.

4.1 Spatial Context

In the existing research on the role of geographical variables in shaping educational participation patterns, two types of research questions can be distinguished. The first type of questions deals with how spatial variables influence the decision about whether or not to continue studying. The second type of research questions is concerned with how spatial context influences mobility choices, i.e., the decision about whether to continue studying in the home region or abroad.[3] Since the decision about where to study is not of particular relevance in the context of inequality, this study concentrates on the first type of questions.

In the literature, the influence of spatial contexts on educational participation patterns is examined in relation to three groups of explanatory variables: (1) regional supply of educational institutions, (2) level of urbanization (mostly measured by rural/urban or population density) and (3) regional labor market conditions.[4] Due to the unavailability of appropriate data, this study does not include regional labor market conditions in the analysis. Regional educational infrastructure (1) and the level of urbanization (rural/urban) (2), in turn, will be included in the analysis, while the empirical state of research is discussed below.

---

[3] For instance, Gibbons & Vignoles' (2012) findings suggest that, in England, the distance from a student's home to the nearest institution is very strongly linked to the decision about where to study.

[4] For the regional labor market conditions the empirical results from studies using data at a micro level are fairly mixed; The majority of studies show that higher regional youth employment levels increase students' chances of entering tertiary education (Giannelli & Monfardini 2003; Flannery & O'Donoghue 2009; Rizzica 2013). Others have found no influence of regional unemployment on educational enrolment (Micklewright, Pearson and Smith 1990), with some studies even reporting negative effects (Farley Ordovensky 1995). Wessling (2016) attributes these ambiguous findings to the assumption that poor labor market conditions influence enrolment decisions in two different ways. On the one hand, a high unemployment rate discourages young people from entering the labor market and induces them to continue studying (*discouraged worker effect*). On the other hand, poor labor market conditions influence parental wealth and could accelerate the entry of young people into the labor market (*wealth effect and risk aversion*). For a more detailed description on theoretical predictions about the effect of unemployment on school attendance, see Micklewright, Pearson and Smith (1990).

4.1.1 Regional Supply of Educational Institutions

Although research concerning educational participation patterns has typically been conducted without taking geographical context into consideration, several studies on mostly developed countries conclude that distance to the closest educational institution has a negative effect on educational attainment (e.g., Frenette 2004, 2006; Spiess and Wrohlich 2010). The underlying theoretical assumption is that distance influences both monetary (commuting or added direct and indirect costs of moving away) and emotional costs (distance from parents and peers etc.). Instead of calculating travel time to the closest university, Wessling (2016) uses different travel time radii to measure the regional supply of educational opportunities. She finds that, up to a scope of 60 min, every additional university has a positive effect on enrolment. She hypothesizes that the effect disappears for larger travel time radii because they exceed the commuting distance.

However, Gibbons and Vignoles (2012) point out that certain factors (e.g., family background or income) potentially influence parents' choice of residence and young people's educational decisions simultaneously. They argue that the linkage between distance and educational participation does not have to be causal, as it may also result from prior residential sorting. To get around this, other research has focused on the impact of newly founded educational institutions. Nonetheless, a positive effect of regional supply of institutions on educational attainment has been reported (Frenette 2009; Rizzica 2013). The aforementioned empirical evidence leads to the first two hypotheses for this study. *A large regional supply of postgraduate educational institutions (measured in terms of travel distance to institutions with different radii) increases the probability for a student to continue studying (Hypothesis 1a). The positive effects of regional supply of tertiary institutions on educational participation become weaker with larger travel distances and disappear for large radii (Hypothesis 1b).*

4.1.2 Level of Urbanization

Another strand of research is concerned with the level of urbanization and focuses on the differences in access to education between urban and rural areas. Typically, lower levels of educational attainment are reported for rural than for urban areas.[5] Stromquist

[5] For a cross-country study comprising 56 countries, including El Salvador, see Ulubaşoğlu and Cardak (2007).

(2004) even claims that, in Latin America, the highest inequality in terms of access to education is that between rural and urban areas. Many empirical studies concerning educational outcomes include rural/urban as a geographical variable in their models to roughly estimate the accessibility of educational institutions. Apart from the fact that students from rural areas have to deal with a scarcity of educational opportunities and thus longer travel times, there are other contextual characteristics that can explain the gap between rural and urban educational attainment. Sá, Florax and Rietveld (2004) find that, even when controlling for distance, the level of urbanization plays an important role in educational participation decision. Behrman and Birdsall (1983) argue that rural-urban inequality emerges due to the lower quality of schools in rural areas. Indeed, several studies conducted in Latin America conclude that the performance of students in rural areas is significantly lower than in urban or metropolitan areas. For instance, Schmelkes (2000) examines school performance in reading and mathematics in Mexican schools, finding that it is systematically better in urban than rural areas. A cross-country study about 13 Latin American countries conducted by Cassasus et al. (2002) shows the same regional differences in performance between rural and urban areas for third- and fourth-year children in primary school. Another classical explanation for the differences in educational attainment between different urbanization levels involves formal labor markets, which in rural areas depend on rural livelihoods and agriculture. If agriculture is carried out primarily for subsistence, the necessary human capital can be accomplished without receiving a long, formal education. Conversely, in areas with higher non-farm employment rates, the expected benefits from education are greater, meaning that a longer period of educational attainment is more attractive (Ulubaşoğlu & Cardak 2007).

Overall, it can be said that rural and urban variables influence educational decision patterns in ways other than simply in terms of the accessibility of educational institutions. Moreover, they involve sociospatial categories representing mechanisms such as rural livelihood and the poor quality of educational institutions in rural areas. From this argumentation, the following hypothesis can be derived: *Students from rural areas are less likely to attend tertiary education, even when controlling for distance to institutions (Hypothesis 1c).*

4.2 Socioeconomic Characteristics

Apart from contextual characteristics, nonspatial conditions are known to shape educational participation patterns. Factors influencing individual outcomes in educational participation patterns at a micro level can be divided into two groups: (1) individual conditions, such as gender and abilities, and (2) family conditions, typically household income, parental education or the number of siblings. This study will include variables relevant to both individuals and family. However, only socioeconomic variables will be taken into account, meaning that psychological and skill-related factors, such as students' school performance, will not be considered.[6]

4.2.1 Family Income

Theory suggests that similar training costs create a higher burden for lower-income families than for higher-income families, which results in different cost-benefit evaluations (Breen and Goldthorpe 1997: 286). Indeed, family income (typically operationalized as equivalent household income) has been found to be positively correlated with tertiary education attendance (e.g., Fuller, Manski and Wise 1982; Christofides, Canada and Hoy 2001; Carneiro and Heckman 2002; Kane 2006). These findings have also been confirmed for the Central American area. In a cross-country study of six Central American countries (Costa Rica, Nicaragua, Guatemala, Panama, Honduras and El Salvador) Bashir and Luque (2012) roughly compare factors determining the attainment levels in tertiary education. They conclude that, in each of these sampled Central American regions, attendance and completion rates for students from low-income family are significantly lower. Moreover, they suggest that, although the overall participation of tertiary education increased between 2001 and 2009 (except for Panama), the inequality gaps between different income groups remained static or even increased. In most examined countries, including El Salvador, the participation rate of the richest quintile has increased the most, while the enrolment rate of the poorest sections has hardly changed. Bashir and Luque relate these findings to the growth in the private sector offering tertiary education.

A very interesting study by Murakami and Blom (2008) has investigated the affordability of tertiary education in Latin America, with specific focus on Brazil, Colombia, Mexico

[6] There is theoretical (*primary effects*) and empirical evidence (Brunori, Ferreira and Peragine 2013) to confirm that performance is also socially selective.

and Peru. They estimate student living costs and study costs in relation to GDP per capita. Their findings suggest that an average family must pay 60% of per capita income for tertiary education per student, whereof less than half is used for living costs. Meanwhile, families in high-income countries pay 19% of GDP per capita for tertiary education. In other words, the affordability of tertiary education in Latin American countries is significantly lower than in high-income countries. *Regarding these findings, it can be hypothesized that a higher level of family income (operationalized as log-transformed equivalent household income) increases the probability for a student to enter or complete tertiary education (Hypothesis 2a).*

4.2.2 Social Class Origin and Intergenerational Mobility

Different strands of research have focused on different indicators to measure intergenerational mobility. The economic literature has mainly measured intergenerational mobility by comparing parents' and children's income. The sociological literature, however, has focused on parents' occupational status and educational attainment. In developing countries, researchers have mainly used parents' educational attainment as a measure of social background for practical reasons. First, accurate data on parents' and children's education are often available and easy to observe and quantify. Conversely, income or wealth is more likely to be measured with large errors (Daude & Robano 2015). A second advantage is that educational attainment is often fixed for adults, while income and parents' occupational status can change over time. In particular, since income is age-dependent, the determination of long-term income is much more complicated than the assessment of when educational attainment (Hertz *et al.* 2008). Several studies conclude that intergenerational mobility is lower in Latin America compared to other regions.[7] For instance, in a cross-country study, Hertz et al. (2008) measure educational mobility for 42 countries, finding the lowest intergenerational mobility to be in the seven Latin American countries included in their analysis. Hence, parental education seems to be a key factor determining children's educational patterns. *It is expected that higher levels of parental education increase a child's probability of attending tertiary education (Hypothesis 2b).*

[7] For a useful overview on intergenerational mobility in Latin America, see Torche (2014).

4.2.3 Gender

Regarding access to education, it is noteworthy that the large inequalities found in Latin America do not reflect large gender differences. Several studies conclude that the educational gender gap in Latin America closed at the end of the last millennium. Since then, in most regions, women have even a higher participation rate than men. With a cross-country analysis, Duryea et al. (2007) provide a useful overview of the evolution of the gender gap in Latin America.

A number of researchers has tried to explain why boys are disadvantaged in terms of school attendance. As the reasons are highly dependent on contextual characteristics, such as normative frameworks, they are only partly transferrable to other regions. In Latin America, boys' lower educational participation rates and school performance have often been linked to the prevalent gang culture. Gangs distract boys from school, which results in higher dropout rates for boys (United Nations Children's Fund (UNICEF) 2003). The related explanation (crime and violence) is especially relevant in the case of El Salvador, since the country is known to have a particularly high homicide rate (Igarapé Institute 2016). Typically, males are more affected by violence and crime, which could lead to higher dropout rates from school. However, it is difficult to determine whether crime and violence are causes or effects in terms of boys' lower participations rates. Furthermore, masculine identity and gender roles (Latin American machismo) are mentioned as reasons for the reversed gender gap. Education might be viewed as feminine and therefore less attractive for boys (Jha, Bakshi and Faria 2012). Another explanation is that, in poor families, pressure to enter the labor market and earn money is particularly high for men. In addition, male privileges in the labor market, such as higher salaries and lower unemployment rates, which persist in Latin America, can encourage boys to quit school (Jha, Bakshi and Faria 2012).

One exception to girls' advantages in terms of access to education concerns Latin American communities with a significant percentage of indigenous populations, where women still receive, on average, fewer years of schooling (Duryea *et al.* 2007, Rama Vitale 2007). Furthermore, findings suggest that, in Latin America, women are still underrepresented on postgraduate programs (Rama Vitale 2007), while labor market discrimination for women persists (Camou 2015). Even though women still seem to be discriminated in the labor market and on (post)graduate programs, the following can be hypothesized: *Women are advantaged in terms of entering and completing tertiary education (Hypothesis 2c).*

4.3 Group-specific Differences in Sensitivity to Spatial Context

As discussed above, there is vast empirical evidence that contextual characteristics, such as regional supply of tertiary education institutions and levels of urbanization, influence educational participation patterns. But, due to mechanisms such as affordability and differing cost-benefit evaluations, these conditions may not be equally relevant to every socioeconomic group. To further explore inequalities in access to education, this section will discuss group-specific differences in terms of how spatial context shapes educational participation decisions.

4.3.1 Family Income and Distance

For emerging regions, affordability plays a major role in educational decisions. Distance to educational institutions may translate into higher financial burdens, due to added financial costs for commuting or moving away. Low-income families may not be able to afford these additional expenses or decide that the high level of investment is not worth it. Furthermore, the small offer from public compared to private postsecondary institutions in El Salvador may involve greater distances or travel times, and thus higher commuting costs, for those who are unable to pay high tuition fees. To the best of my knowledge, no studies on the relation between family income classes and geographical barriers have been conducted in Latin America. Studies in other regions, however, indicate that the effect of the distance to school on school attendance is linked to family income. The negative effect of distance on educational participation seems to be much stronger for students from a lower-income family, while wealthy students tend to be less affected (Frenette 2004; Frenette 2006; Eliasson 2006). *Concerning these findings, it can be hypothesized that children from a lower-income family are particularly disadvantaged by distance (Hypothesis 3a).*

4.3.2 Social Class Origin and Distance

The relation between distance to school and social background seems to be crucial for policy decisions about establishing new educational institutions or improving infrastructure, which result in shorter travel times. Knowing how distance to an educational institution affects students from different family backgrounds could help to undermine the persistence of inequalities across generations. Findings from developed countries suggest that, analogous to family income, students from a lower social class origin or lower family background seem to be more disadvantaged by distance to school

(Eliasson 2006; Frenette 2006; Cullinan *et al.* 2013; Wessling 2016). As intergenerational mobility is measured in this study with regard to parents' educational attainment, it will be used also as an indicator for measuring the relation between social background and regional supply of educational institutions. *Regarding tertiary education, it is expected that children with parents with a lower educational attainment level are particularly disadvantaged by distance (Hypothesis 3b).*

4.3.3 Gender and Distance

As discussed in Section 4.2.3, in Latin America, boys seem to be slightly disadvantaged in terms of attending tertiary education. In return, studies on developed countries suggest that women respond more strongly to contextual characteristics and changes than men. For instance, Rizzica (2013) examines how the expansion of tertiary education supply, as part of a reform process in Italy, influenced enrolment decisions of young Italians, while highlighting significant gender differences. The increase in educational supply in proximity resulted in an increase in women's participation rates in tertiary education, while rates for men were unaffected. The author attributes these findings to the prevalence of traditional gender roles, where women are more strongly bound to their homes. However, she warns against generalizing these findings to countries without similar cultural models to those of Italy. Likewise, research in other countries has found that women are more advantaged by proximity to educational institutions (e.g., Kazeem, Jensen and Stokes 2010 for Nigeria; Burde and Linden 2013 for Afghanistan). *Even though these findings should be interpreted with caution, it is hypothesized that women are particularly disadvantaged by distance (Hypothesis 3c).*

5 Data and Methods

5.1 Data

The empirical analysis in this paper is based on three different types of data: (1) the 2015 Salvadoran household survey, EHPM, (2) spatial data on population density, road networks and administrative borders, and (3) the location and other attributes of each postsecondary institution in the country. *Table 2* provides an overview of the data used in this study.

The EHPM (1) is a cross-sectional household survey conducted once a year by the Salvadoran Office for Statistics and Censuses, DIGESTYC and the Salvadoran Ministry of Economy. The EHPM is conducted in each of the 14 departments and covers urban and rural areas. It contains data on economic and social variables for 229 of the 262 Salvadoran municipalities. The 2015 edition covers 23,670 households comprising a total of 88,184 people. For the EHPM, the Salvadoran Office for Statistics and Censuses uses a two-stage sampling design. Based on the population census from 2007, the country is divided into collection districts (for urban areas with between 120 and 150 inhabitants; for rural areas with between 50 and 70 inhabitants) and some of these districts are selected (first stage). In a second step, households are listed within each selected district and some households are chosen (second stage). Each member of the randomly selected households is then interviewed (Dirección General de Estadística y Censos (DIGESTYC) & Ministerio de Economía (MINEC) 2016). Spatial data (2) is gathered from different sources, as listed in *Table 2*. The spatial data involves a population density grid, two road networks, and information on administrative units and borders. Information on postsecondary institutions (3) is also collected by combining different sources. Institutions that only offer postgraduate programs, military schools, and institutions that opened after 2014 were excluded. The list can be seen in *Appendix 1*.

Data source	Data type	Information on
Salvadoran Office for Statistics and Censuses (DIGESTYC) and the Salvadoran Ministry of Economy (MINEC) (2016)	Household data: EHPM (Encuesta de Hogares de Propósitos Múltiples) Survey 2015	Economic and social variables on household and individual level
OpenStreetMap (2014)	Spatial data: roads (feature)	Highway road network
Centro Nacional de Registros (2015b)	Spatial data: transports (feature)	Road network Different road types (paved/not paved)
European Commission, Joint Research Centre (JRC) & Colombia University, Center for International Earth Science Information Network - Ciesin (2015)	Spatial data: population grid (raster)	Population density with a resolution of 1 km^2
Centro Nacional de Registros (2015a)	Spatial data: administrative borders (feature)	Department borders Municipality borders
(1) Ministerio de Educación El Salvador (MINED) (2014) (2) Altillo (2017), a portal for students in Latin American countries (3) Websites of postsecondary institutions (see *Appendix 1*)	Spatial and nonspatial information about postsecondary institutions	Location of postsecondary institutions Number of study programs

Table 2: Overview of the data used in this study.

5.2 Travel Distance Measurements

5.2.1 Previous Attempts to Measure Accessibility

In the literature on educational decisions, different approaches for measuring accessibility to education institutions can be distinguished. Given the fact that this area of research is dominated by sociologists, due to a lack of geographical tools and knowledge in many studies, distance measurements are rather crude.

Probably the most commonly used approach is to calculate straight-line distances using geographical coordinates from students' homes to the nearest postsecondary institution (e.g., Frenette 2004, 2006; Spiess and Wrohlich 2010). This approach can be criticized at least from three perspectives:

(i) In educational decision-making, students may not take linear distances into account, but instead estimate the travel time from their home to possible study places. Especially in regions with poor infrastructure and major differences in the density of the road network, travel distances can differ significantly from straight-line distances. Further, Euclidian measures do not

consider differences in rural and urban travel distances due to differing infrastructure. Taking this into consideration, authors such as Eliasson (2006) and Cullinan et al. (2013) calculated the travel distance to the nearest educational institution using road networks.

(ii) The distance approach can further be criticized for not considering the number of institutions within commuting distance. The basis of this stance is formed by assuming that numerous educational opportunities increase the possibility that one of them matches an individual student's interests. There is also empirical evidence to indicate that each additional educational institution in proximity has a positive effect on a student's enrolment decisions (Turley 2009).

(iii) Assuming that a greater selection of study opportunities in proximity has an impact on educational decisions leads to the conclusion that the size of universities should be considered. An educational institution in proximity with only a few study programs should not result in the same size of effects as an institution with a wide range of different study programs.

In order to model travel distances in this study, the above-mentioned points of criticism will be considered. To live up to the first critique (1), the calculation is based on the time by car using a road network.[8] Following the approach of Wessling (2016), besides estimating the shortest path, (2) the number of postsecondary institutions in different radii (30, 45, 60, 90, 120, 150 and 180 min) are identified. The radii primarily serve to estimate how much different distances matter and to determine the scope within which the supply of educational institutions matters to an individual. Taking the size of the educational institutions into account (3), I further included the number of reachable study programs in the travel time radii (for a simplified illustration, see *Figure 3*).

[8] Note that it could be argued that students are more likely to take public transport into account when estimating the travel time to prospective study places. This critique becomes even more important as El Salvador is a region where having a private car is not common among students. Unfortunately, due to the lack of electronic data on bus timetables in El Salvador, it was not possible to include public transport in this study.

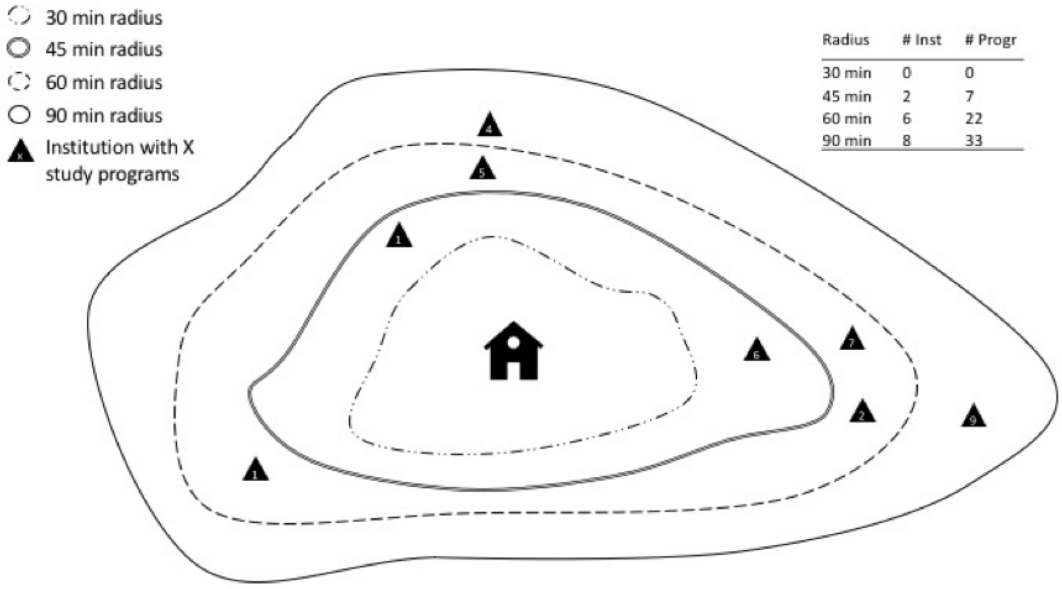


Figure 3: Illustration of travel time radii; the number of postsecondary institutions and the number of study programs were calculated for the different travel time radii.

5.2.2 Method to Measure Accessibility

In this section, I will outline the technique as to how travel distances were estimated using a geographic information system (GIS). All of the GIS analyses and calculations for this study were undertaken using ArcGIS 10.3.

(i) Combining road network data sets:

In the distance calculation for this study, the road network plays an important role in travel time outcomes. Missing streets in the data set ought to distort the results enormously, given that, in the model, off-road traveling is expected to be very slow. Especially in studies on developing countries, dealing with fragmentary and erroneous data is a challenge that many researchers have to face. To ensure a more accurate and complete road network in this study, two data sets from different sources were merged. For an overview of the data sources, see *Table 2*. One source was more precise for urban regions, while the other data set contained more accurate information on streets in peripheral and rural regions. Combining the two data sets resulted in a fairly detailed and exact road network for further analysis (see *Appendix 2*).

(ii) Speed raster:

In a next step, I assigned an average speed to the different road types (see *Appendix 3*). Subsequently, a speed raster of the whole country was generated (100x100 m resolution).

(iv) Cost distance raster:

In order to measure travel times, the postal address of every postsecondary institution was collected and then geocoded to provide precise spatial (longitudinal, latitudinal) coordinates for spatial analysis (see *Appendix 1*). By means of the *cost distance function* (Spatial Analyst Toolbox) with regard to each institution, the least accumulative distance for each cell was calculated.

(iii) Municipality average:

Since the spatial resolution of the EHPM data is based on municipalities, average speeds for each municipality, divided between rural and urban, had to be estimated. Given the fact that, in the rural area, many parts of the country are uninhabited, calculating an average for each cell in a municipality could have distorted the results and overestimated travel times. Therefore, by means of a population density data set, cells with less than 10 inhabitants per km^2 were removed. Subsequently, I evaluated the average distances from the rural and urban areas of each municipality to each institution using the *Zonal Statistics* tool.

(v) Travel time radii:

From the calculated travel time matrix (consisting of 47 institutions x 262 municipalities divided into rural/urban), the accessible postsecondary institutions and study programs in the time radii (30, 45, 60, 90, 120 and 180 min) for each municipality could be estimated (*Figure 4*).

(iv) Matching travel times with individual data:

Finally, the distance to the nearest postsecondary institution (*Appendix 4*) and the radii were added to the individuals according to their municipality and its subdivision into either rural or urban.

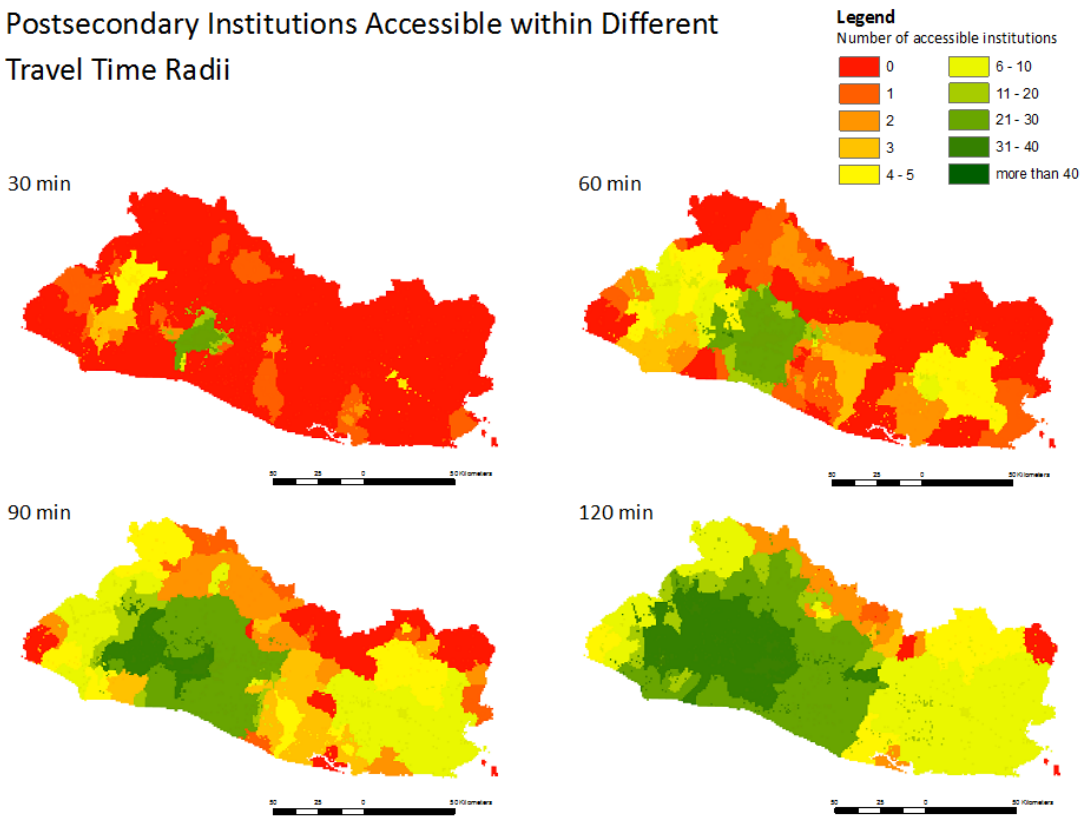


Figure 4: Number of postsecondary educational opportunities accessible in 30, 60, 90 and 120 minutes in El Salvador. Spatial resolution: municipality, divided into rural/urban.

5.3 Statistical Evaluation

This section describes the sample and the models and explains how certain variables were calculated. All of the statistical analyses were made using Stata; the script can be viewed in *Appendix 9.*

5.3.1 Base Population and Sample

The base population was set to young adults between 18 and 30 years who graduated from secondary school. Note, that the way to completing secondary education (e.g. transition from primary to secondary education) is already selective, but as this study focuses on access and transition to tertiary education only young adults who are eligible for tertiary education are of interest. The range of ages was set to between 18 and 30 years for several reasons. First, during this stage in life, it is usual to be studying or to have recently graduated from tertiary education (only a few adults start their studies after

reaching the age of 30). Second, certain characteristics, such as family income, change over time. This study aims to include changing characteristics as close as possible to the time of study decisions, which is why the study is limited to those younger than 30. Third, a huge study gap exists between generations. As this study is interested in the current state, rather than the historical development, of educational inequalities, it makes more sense to limit the analysis to the current study generation.

Incomplete data represented a challenge for this study. As the survey was conducted at the household level, it only contains data on family and parents of adolescents officially living at the parental home. Even though it is not that common in El Salvador to move out while studying, this has led to some gaps in the data. On the one hand, complete case analysis can lead to distortions and sample selection bias. In complete case analysis, cases with missing values are removed, resulting in the reduction of the sample size. In our case, this would have meant to reduce the sample to young adults living at the parental home, resulting in a non-representative sample for the base population of young adults. On the other hand, as the data was *missing not at random* (MNAR), it was not appropriate to replace missing data with substituted values using a multiple imputation procedure.[9] In order to deal with the missing values for the statistical evaluations, various models with different samples were calculated. The estimation of the effects of spatial and individual variables was based on the complete sample (n=8,526), while estimations of family background effects were based on a reduced sample (n=4,951). The latter was reduced to young adults living at the parental home, and coded as “son” or “daughter” in the respective household.

The comparison of characteristics in the two different samples in *Appendix 5* and *Appendix 6* shows that, as expected, there was a higher participation rate in tertiary education among those living in the parental home. Conversely, spatial context, such as minimal distance and area (rural/urban), was similar in both groups. Within the reduced sample, the average age was only slightly lower and contained a few more men than women.

[9] The data was MNAR because it can be assumed that students have a higher tendency to stay in the parental home compared with young adults who enter the labor market, i.e., the distribution of missing data was probably affected by the dependent variable of this study.

5.3.2 Logistic Regression Models

To estimate the effects of the variables discussed in Section 3 on study decisions, this research uses logistic regression models with multiple predictors. The independent variable I used in this case is binary coded. I attributed 1 to young adults who are attending tertiary education or have already graduated from tertiary school, and 0 to those who have received no tertiary education. Regression analysis now attempts to model the effect of different predictors on the *probability* to attend tertiary education. As linear regression can produce probabilities that are less than 0, or even bigger than 1, logistic regression is more suitable. For an understandable interpretation of the logistic regression results, the average marginal effects (AMEs) were calculated. The AMEs measured the average change in probability in terms of attending tertiary education when the independent variable increased by one unit. With binary independent variables, the AMEs measured discrete change effects, i.e., the change in the predicted probabilities as the variable change from 0 to 1.

To test the hypotheses on the effects of regional educational supply (*Hypothesis 1a*) and their spatial scale (*Hypothesis 1b*), I calculated a logistic regression for all travel time radii between 30 and 180 min (hereinafter referred to as *Model 1*). These regressions with different radii were conducted for both the number of educational institutions and the number of study programs (controlled by the number of institutions). In the other logistic regression models, instead of including the educational supply within different radii, the educational supply at a regional level was represented by a single variable: the travel time to the nearest postsecondary institution. *Hypotheses 1c* and *2c* were tested by using the complete sample containing young adults who were either living with their parents or on their own (hereinafter referred to as *Model 2*). Conversely, on account of containing family characteristics, *Hypotheses 2a* and *2b* were based on the reduced sample (hereinafter referred to as *Model 3*). Note that family income (*Hypotheses 2a* and *3a*) was only assigned to those living in the parental home, since this variable should reflect the family background and not one's own income after moving out. For *Hypotheses 3a*, *3b* and *3c*, three different models, including interactions between variables, were calculated (hereinafter referred to as *Interaction 1, Interaction 2* and *Interaction 3*).

5.3.4 Variable Specifications

In many (in particular, earlier) studies, social mobility is measured by the father's education. But, in the case of El Salvador, the father's absence from the household is not unusual, meaning that the mother's education should be taken into account. If both parents are present, the variable *parent's educational attainment* represents the higher educational attainment of both parents. Note that, for this analysis, those who were listed as sons or daughters in a household, with the presence of both, the head and the spouse, were assumed to be the children of both. Therefore, some of the children were not the biological children of both of their assigned parents. As this study does not aim to investigate genetic mechanisms, there was no need to exclude stepparents.

The *household income* was calculated for each individual separately as follows. First, an individual's own earnings were subtracted from the total family income. This was based on the presumption that young adults who had dropped out of school or graduated from university to start working made their family wealthier than they were at the moment of the study decision. Subsequently, due to economies of scale, instead of dividing the household income by the number of household members, weightings were estimated with the commonly used *OECD-modified equivalence* scale (first proposed by Hagenaars et al. 1994). According to this scale, the weightings in a household are calculated as follows: 1.0 to the first adult, 0.5 to each subsequent person aged 14 and over, 0.3 to each child aged under 14. The household income is then divided by the sum of these weightings, resulting in the equivalence income. Finally, as income is no normally distributed, a log-transformation was applied.

6 Results

This section summarizes the main results of the statistical evaluations. An overview of the hypotheses and the corresponding results is provided in *Table 3.*

	Independent variable	**Mediation**	**Dependent variable**	**Model**	**Results**
1a	↑ Regional supply of tertiary educational institutions		↑ Participation in tertiary education	*Model 1, Model 2*	Confirmed for institutions; study programs have no effect when controlling for the number of institutions
1b	↑ Regional supply of tertiary educational institutions	↑ Travel time radii	↓ Effect on participation in tertiary education	*Model 1*	Confirmed
1c	Rural		↓ Participation in tertiary education	*Model 2*	Confirmed
2a	↑ Household equivalence income (log)		↑ Participation in tertiary education	*Model 3*	Confirmed
2b	↑ Parents highest educational attainment		↑ Participation in tertiary education	*Model 3*	Confirmed
2c	Gender (female)		↑ Participation in tertiary education	*Model 2*	No significant results, but a tendency that women have higher chances to attend tertiary education
3a	↓ Household equivalence income (log)	↑ Distance to nearest tertiary institution	↓ Participation in tertiary education	*Interaction 1*	Opposite significant effects: individuals from a higher-income family seem to be more sensitive to distance
3b	↓ Parents highest educational attainment	↑ Distance to nearest tertiary institution	↓ Participation in tertiary education	*Interaction 2*	No significant results
3c	Gender (female)	↑ Distance to nearest tertiary institution	↓ Participation in tertiary education	*Interaction 3*	No significant results, but an opposite tendency: men may be more sensitive to distance

Table 3: Hypotheses with results.

Model 1 calculates a logistic regression for every travel time radius, including individual and contextual control variables (age, area rural/urban, gender). *Figure 5* summarizes the effects of the number of tertiary institutions within different travel time radii, displayed as AMEs. A higher regional supply of educational institutions has significant positive

effects on educational participation, which supports *Hypothesis 1a*. Within a 30-min radius, every additional tertiary institution increases the probability of attending tertiary education by almost 0.7 percentage points. The positive effect of educational supply on study participation remains significant up to a 120-min radius. With exception of a 90-min radius, the effects decline with every larger radius. This means that the positive effects of educational supply weaken with longer travel distances (*Hypothesis 1b*). Using the number of study programs instead of the amount of institutions within the different radii, similar effects are observed (*Appendix 8)*. But, when controlling for the number of institutions, these effects disappear. These results indicate that the number of study programs does not have an additional effect on educational decisions.

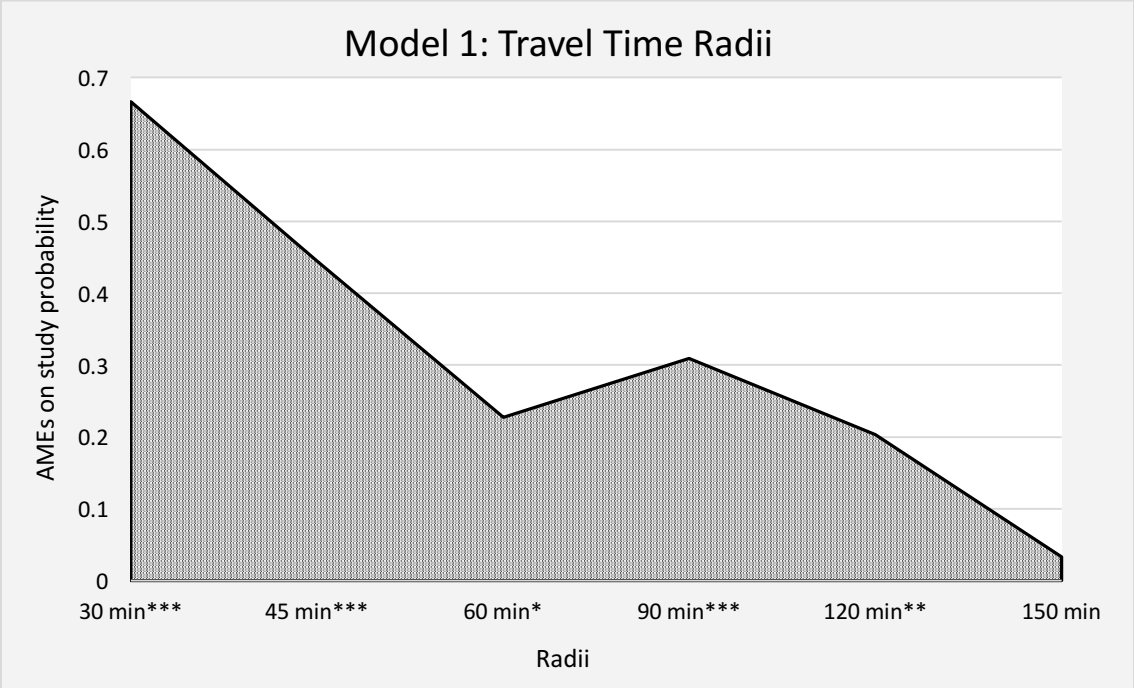


*Figure 5: Effects of additional institutions on the probability of participating in tertiary education within different travel time radii, presented as AMEs; significance level: *p<0.05 **p<0.005, ***p<0.001, based on the complete sample (n=8,526), additional control variables: age, sex, area (rural/urban).*

Instead of travel time radii, *Model 2* includes the distance to the closest postsecondary educational institution as a single variable for accessibility. In *Figure 6*, the AMEs of individual and spatial independent variables are presented. Again, travel distance has a significant negative effect. With every additional hour to the closest educational institution, the probability of attending tertiary education decreases by 20 percentage points. This effect again supports *Hypothesis 1a*. Furthermore, the likelihood that young adults living in urban areas participate in tertiary education is 15 percentage points higher than for those from rural areas. The effect is highly significant with p(two-sided)<0.001, providing strong support for *Hypothesis 1c*. Women seem to be slightly advantaged over

men, but this effect is not significant at a 5% level. Hence, *Hypothesis 2c* cannot be confirmed.

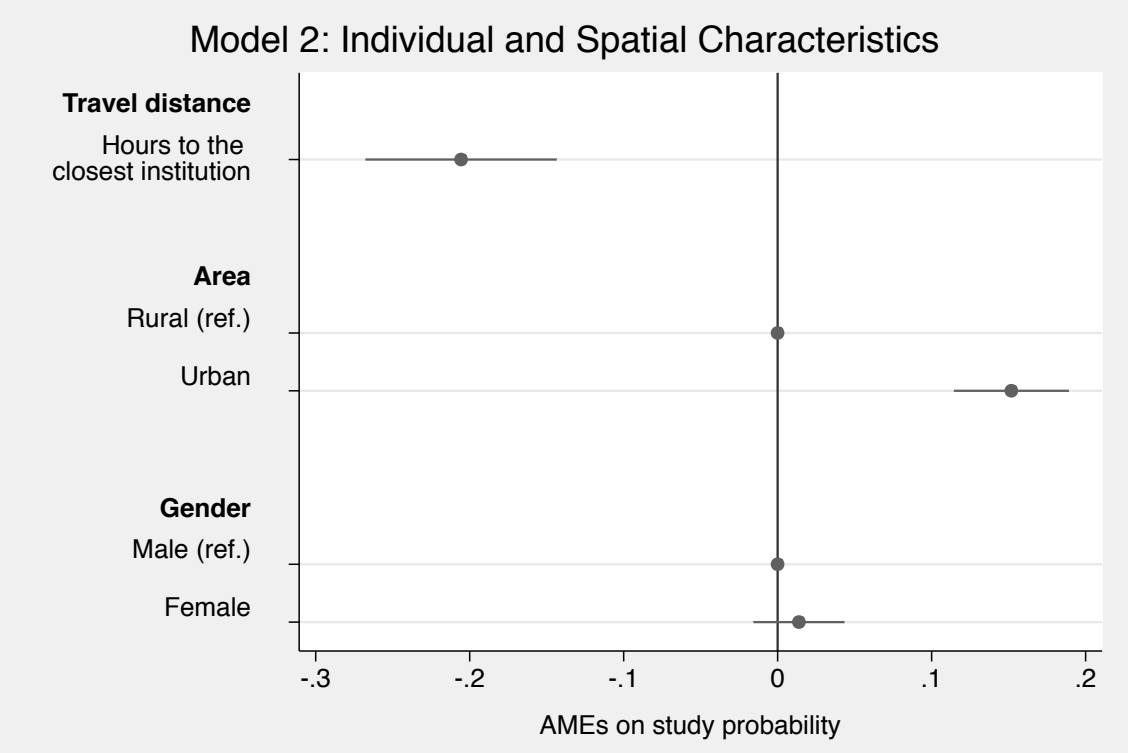


Figure 6: Effects of individual and spatial variables on the probability of participating in tertiary education, displayed as AMEs; significance level: p<0.05, based on the complete sample (n=8,526), additional control variable: age.

Model 3 is calculated with the reduced sample containing only those young adults who are still living in the parental home. The AMEs are presented in *Figure 7*. An increase in the family income by 1% results in the likelihood of attending tertiary education being 0.14 percentage points higher.[10] The effect is highly significant with p(two-sided)<0.001 (*Hypothesis 2a)*. Furthermore, parents' education seems to influence educational decisions. While there is no significant increase from "no education" to "primary education", children with more highly educated parents have a higher probability of attending tertiary education. These effects are very strong. Children with parents who have a tertiary degree have a likelihood of participating in tertiary education that is 52 percentage points higher than for those with parents who have received no education. The results are in line with *Hypothesis 2b*.

[10] Note that this interpretation is possible because family income is log-transformed, but it is only approximate.

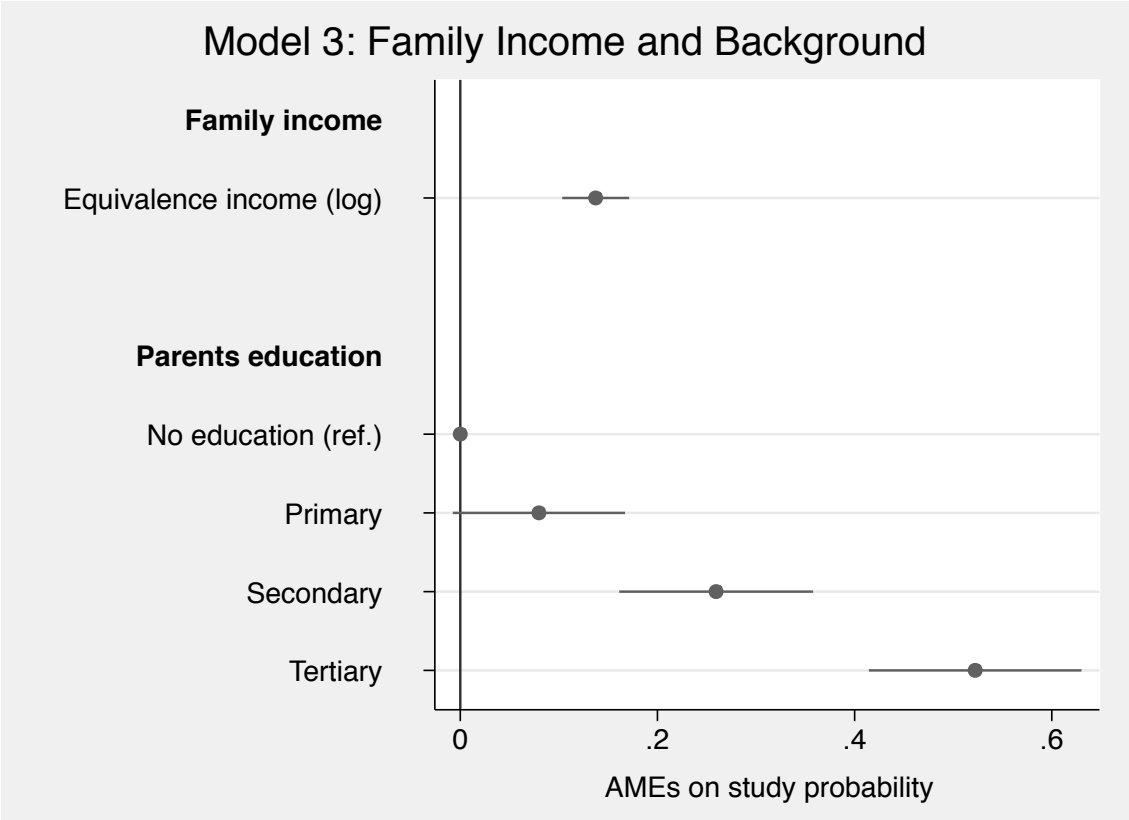


Figure 7: Effects of family background variables on the probability of participating in tertiary education, displayed as AMEs; significance level: p<0.05, based on the reduced sample (n=4,951), additional control variables: age, distance to the closest institution, sex, area (urban/rural).

The results of the interaction models (*Interaction 1*, *Interaction 2* and *Interaction 3*) are displayed in *Table 4.* The first two interaction models (for *Hypotheses 3a* and *3b*) are based on the reduced sample; the third interaction model is based on the unreduced sample, as *Hypothesis 3c* does not involve family variables. To calculate the effects from an average income, instead of zero income, *Interaction 1* uses the centered log-transferred equivalence income.[11] *Table 4* shows that distance to the closest educational institution has no significant negative effect for those with average family income. Individuals from a higher-income family, however, seem to be significantly more disadvantaged by distance (or more advantaged by proximity). *Hypothesis 3a* states that children from a lower-income family are particularly disadvantaged by distance. Since the results indicate a significant effect in the opposite direction, *Hypothesis 3a* must be rejected. Furthermore, the interactions between parental education and distance (*Hypothesis 3a*) are not significant (*Interaction 2*). Parental education does not seem to

[11] The income is mean-centered and, as such, the interaction effects can be better interpreted.

play a role in shaping one's sensitivity to distance. Last but not least, contrary to *Hypothesis 3c*, there is a tendency that men are more sensitive to than women. But the interaction effect is not significant (*Table 4*, *Interaction 3*).

Overall, the *Hypotheses 1a, 1b, 1c, 2a* and *2b* can be confirmed. By contrast, the conducted analysis does not provide support for *Hypotheses 2c, 3b* and *3c*. As for *Hypothesis 3a*, an effect in the opposite direction is significant.

	Interaction 1: income and distance	*Interaction 2:* parental education and distance	*Interaction 3:* gender and distance
Distance: hours to the closest	-0.0497	-0.147	-0.179***
educational opportunity	(0.0348)	(0.0921)	(0.0332)
Family income (centered)	0.286***		
	(0.0239)		
Interaction 1: income and distance			
Family income (centered) * distance	-0.135***		
	(0.0364)		
Parental education (ref. no education)			
Primary education		0.0364	
		(0.0833)	
Secondary education		0.258**	
		(0.0887)	
Tertiary education		0.604***	
		(0.0839)	
Interaction 2: parental education and distance (ref. no education)			
Primary education * distance		0.105	
		(0.100)	
Secondary education * distance		0.113	
		(0.106)	
Tertiary education * distance		0.0372	
		(0.123)	
Sex (female)	0.0348	0.0274	-0.00948
	(0.0185)	(0.0193)	(0.0262)
Interaction 3: gender and distance (ref. male)			
Sex (female) * distance			0.0633
			(0.0450)
N=individuals	4,914	4,951	8,526

*Table 4: Interaction effects on the probability of participating in tertiary education, displayed as AMEs; significance level: *p<0.05 **p<0.005, ***p<0.001, additional control variables: age, area (urban/rural).*

7 Discussion

7.1 Spatial Context

As shown in *Model 1* and *Model 2*, the supply of education at a regional level and the distance to the closest educational institution strongly shape educational decisions (*Hypothesis 1a*). Nevertheless, one ought to be careful with causal interpretations. This effect may be partly driven by the process in which students from remote areas move closer to the educational institution when they start tertiary education. The available household survey data does not allow us to distinguish between these two effects. But this is not a serious problem, since the positive effect of additional educational institutions is found within larger radii as well.[12] Thus, it can be assumed that the effects do not exclusively occur with relocations.

With regard to scope, I found that educational supply has an impact on study decision up to a radius of 120 min. Additional educational institutions in a radius of 150 and 180 min no longer seem to have a positive effect. When examining tertiary education decisions in Germany, Wessling (2016) found a similar pattern, but with a scope of 60 min. Spatial scopes may differ between these two regions due to differences in the quality of infrastructure and public transport, as well as differences in the accessibility of educational institutions. Individuals in El Salvador may be used to longer commuting times and thus be affected by characteristics in a larger spatial scope. More research is needed to confirm this.

Although controlled for travel distance, individuals who live in rural areas seem to be particularly disadvantaged when it comes to participating in tertiary education (*Hypothesis 1c*). The rural and urban spatial context variables seem to influence decision patterns in ways other than in terms of the accessibility of educational institutions. As reasons for this gap between rural and urban areas, the literature refers to the lower quality of schools in rural areas (Behrman & Birdsall 1983), urbanization and urban lifestyles (Sá, Florax and Rietveld 2004), lower parental education levels and family income in rural areas (McCracken & Barcinas 1991), and the role of agriculture and associated informal employment in rural areas (Ulubaşoğlu & Cardak 2007). In Section 2, *primary* and *secondary effects* of geographic context were discussed. With the available data for this study it was not possible to distinguish between these two

[12] Students who relocate due to study reasons can be expected to live near a tertiary educational institution (presumably within a radius of at least 60 min).

mechanisms that might underlie regional differentials in educational attainment. However, it would be highly relevant to know to what extent the mentioned educational disparities can be attributed to differences in ability and motivation (e.g. because of availability and quality of establishments that promote education) or divergent cost-benefit evaluations respectively. Particularly regarding policy implications, the specific reasons and sociospatial mechanisms behind this rural-urban gap should be explored by conducting further research with appropriate data.

7.2 Socioeconomic Context

Model 3 shows that family income plays an important role in access to education (*Hypothesis 2a*). The strong effect of income on study participation rates raises the question about whether tertiary education is affordable to lower-income families. As the comparative study conducted by Murakami and Blom (2008) suggest, affordability of tertiary education seems to be a major problem in Latin America.

In addition to family income, parental education seems to be a key factor determining an individual's educational patterns (*Hypothesis 2b*). Furthermore, the findings are in line with the concept of status attainment (Breen & Goldthorpe 1997), as discussed in Chapter 2, since the probability of entering tertiary education increases by more percentage points between the secondary and tertiary parental education variables than between the primary and secondary parental education variable (see *Figure 7*). Moreover, parental education is significant even when controlling for family income. Consequentially, not only economic but also social factors, such as status maintenance, seem to be responsible for inequalities in education.

Although there is a tendency that women are advantaged over men, gender does not seem to play a major role in determining access to tertiary education. In contrary to other developing regions, the large inequalities in Latin America are not manifest in large gender differences.

7.3 Group-specific Differences in Sensitivity to Spatial Context

For the interaction between family income and travel distance, a reverse effect to the one expected was found. Empirical studies conducted in developed regions, found that the negative effect of distance on education participation is much stronger for students from lower-income families (Frenette 2004, 2006; Eliasson 2006). However, the results in this study indicate that individuals from higher-income families seem to be more

disadvantaged by distance. There are two possible explanations for these findings. On the one hand, this interaction effect may in fact be caused by differences in cost-benefit evaluations relating to travel times. Other than in developed countries, young adults from lower-income families may care less about the additional travel time, such that travel distance may not matter too much in terms of affordability. On the other hand, this effect may also emerge from the selection of the reduced sample. It is plausible that wealthier students have a higher tendency to move out from the parental home in order to attend educational institutions, which would exclude them from the reduced sample. With the available EHPM data, it is not possible to determine where the interaction effect arises from. In contrast to family income, no such effects have been observed with regard to social background.

Finally, women do not seem to respond more strongly to spatial characteristics (*Hypothesis 3a*). There is even a tendency that men are more disadvantaged by distance to educational institutions, but this effect is not significant. This shows that, as already highlighted by Rizzica (2013), gender differences seem to depend largely on regional and cultural context; thus, generalizations should be made with caution.

7.4 General Limitations

In addition to the above-mentioned difficulties, I would like to point out some general issues and limitations.

A limitation concerning travel times is that municipalities, even though divided into rural or urban, are still large local units. Using districts seems to be a better approach for examining spatial patterns and characteristics. Especially in developing countries, however, there is a lack of accurate spatial data and the localization of existing survey data is often made at the municipality level. The inaccuracies of the measuring process possibly lead to the underestimation of the effect.

Another issue is that, in this study, only the effects of certain variables, and not the mechanisms behind these effects, could be determined. Understanding these mechanisms is especially relevant to policy interventions. Different survey data or even a qualitative design would be necessary to examine the specific mechanisms and processes that lead to unequal access to education. Nevertheless, knowing which variables affect study participation levels is particularly important for designing prospective investigations.

In this study, individuals are taken to participate in tertiary education when they either hold a tertiary degree or are attending a tertiary educational institution. In terms of inequality in education, it is not only of great interest to determine enrolment rates, but also to investigate failure and dropout rates from tertiary education. For this kind of study, a panel data set, which follows the education journey of young adults from a certain cohort (with a focus on transition points), would be fruitful. Unfortunately, for El Salvador, no such data is available.

This brings me to my last point. Although the EHPM data was not optimal for the purposes of this study, there were no better data available for El Salvador. The lack of good data is a common problem when conducting research in developing countries. Obtaining better data would be highly desirable. Nevertheless, it is very important to carry out research in these regions to inform policies with scientific evidence.

7.5 Policy Implications

Since the field of research examining inequalities in education has significant policy relevance, this section will outline some policy strategies that follow from the results in this study. The suggestions made in what follows should not be regarded as complete, but rather as a general overview of ideas.

Since individuals from rural and remote areas are particularly disadvantaged when it comes to receiving tertiary education, explicit policy interventions in these areas seem appropriate. To overcome spatial inequality in access to education within Germany, Spiess and Wrohlich (2010) propose the implementation and improvement of distance learning programs. But, contrary to developed countries, in the context of Latin America, this intervention is only reasonable if computers and Internet access are available in remote areas. Other policy options for reducing spatial gaps are the supply of affordable or free accommodation for students from remote areas near to educational institutions or improvements in public transport facilities on the way to schools.

Besides spatial inequality, family background and family income were also found to play a role in access to education. To prevent inequalities from persisting across generations, specific incentives and monetary support for students from low-status and low-income families should be offered. In line with the theory discussed in Section 2, policy interventions could address this issue by strengthening benefits, reducing costs or raising the chances of success for students from low-status families. One such intervention could be to offer scholarships, which are not only based on academic merit,

but also based on financial need (Winkler 1990). Conditional cash transfer programs might tackle affordability problems and give lower-income students the chance to go to university. On account of the large private educational sector in Central America, Bashir and Luque (2012) propose that specific funding should be available for students from low-income backgrounds who are studying at private institutions. A different approach would be to enhance the chances of completing tertiary education. Bashir and Luque (2012) calculated that approximately 50% of students who start tertiary education in El Salvador do not complete it (Figure 9, 2001 and 2009). This is an enormous waste of resources. Study assistance and free tuition for low-income students could be effective measures.

8 Conclusion

The aim of this paper was to examine how spatial and socioeconomic factors shape inequalities in access to postsecondary education in the Latin American country El Salvador. As expected, distance to tertiary educational institutions and urbanization (rural/urban) both play a role in determining study participation patterns. Individuals seem to be positively affected by each additional study opportunity up to a travel time radius of approximately two hours, which is a noticeably larger radius than other studies have found in developed regions. Furthermore, the influence of spatial context seems to decline with increasing travel time radii. Meanwhile, regarding family background, large inequalities were found. Individuals from low-income families seem to be disadvantaged in terms of attending tertiary education. Apart from that, children of parents with lower educational attainment levels were found to have lower propensities to participate in tertiary education. Conversely, gender is not a major concern with regard to inequality of access to tertiary education in El Salvador. The presumptions about group-specific differences in sensitivity towards spatial context could not be confirmed.

Finally, I would like to conclude by pointing out two issues:

First, research examining educational participation patterns should refrain from treating educational decisions as though they were unrelated to geographical context. The results from this study indicate that young adults are embedded in family contexts as well as spatial context, both of which influence their study opportunities and decisions. Future research should therefore link individual and family characteristics with spatial characteristics. A combination of sociological and geographical concepts and tools would strengthen theoretical frameworks and empirical research. Furthermore, given that the few empirical attempts to include spatial context in this field mostly use geographical tools and methods in an oversimplified way, the technique for modeling accessibility in this study can be viewed as a suggestion for modeling spatial context in future research.

Second, most empirical studies, as well as theoretical frameworks, concentrate on developed regions. Especially in developing countries, however, where the average number of years of schooling is lower and the distribution of education is often more unequal, research in this field has a significant policy implications. Results from this study indicate that some processes and tendencies in educational decisions in El Salvador may differ from previous findings in developed regions. Conducting more research and acquiring more data in the region of Latin America are both highly important and necessary to address inequalities in access to education. In addition, research that not only focuses on the effects of certain characteristics on educational participation, but

also on the mechanisms behind these effects is highly desirable. All things considered, researchers in this field should shift their attention to developing regions for two main reasons: This contribution provided a further piece of evidence that mechanisms of educational decision-making may differ from those in developed countries. Furthermore, research in low- and middle-income countries can provide the scientific basis for an egalitarian development process.

Literature

ALTILLO. 2017. Universidades de El Salvador (Privadas y Públicas). *Universidades en El Salvador. Universidades Privadas y Públicas de El Salvador ordenadas por Departamento.* Available at: http://www.altillo.com/universidades/universidades_salvador.asp [Accessed July 29, 2017].

BASHIR, S. & LUQUE, J.A. 2012. *Equity in Tertiary Education in Central America: An Overview* (ID 2133815). Rochester, NY: Social Science Research Network.

BECKER, G.S. 1994. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. 3rd ed. Chicago: University of Chicago Press, 412 pp.

BEHRMAN, J.R. & BIRDSALL, N. 1983. The Quality of Schooling: Quantity Alone is Misleading. *The American Economic Review*, **73**, 928–946.

BOUDON, R. 1974. *Education, Opportunity, and Social Inequality: Changing Prospects in Western Society*. John Wiley & Sons Canada, 256 pp.

BREEN, R. & GOLDTHORPE, J.H. 1997. Explaining Educational Differentials: Towards a Formal Rational Action Theory. *Rationality and Society*, **9**, 275–305.

BREEN, R. & YAISH, M. 2006. Testing the Breen-Goldthorpe model of educational decision making. *Mobility and inequality : frontiers of research in sociology and economics*.

BRUNORI, P., FERREIRA, F.H.G. & PERAGINE, V. 2013. Inequality of opportunity, income inequality and economic mobility : some international comparisons (WPS6304). The World Bank. Available at: http://documents.worldbank.org/curated/en/864781468331737091/Inequality-of-opportunity-income-inequality-and-economic-mobility-some-international-comparisons [Accessed June 17, 2017].

BURDE, D. & LINDEN, L.L. 2013. Bringing Education to Afghan Girls: A Randomized Controlled Trial of Village-Based Schools. *American Economic Journal: Applied Economics*, **5**, 27–40.

CAMOU, M.M. 2015. Historical Patterns of Gender Inequality in Latin America: New Evidence. Programa de Historia Económica, FCS, Udelar. *38*.

CARNEIRO, P. & HECKMAN, J.J. 2002. The Evidence on Credit Constraints in Post-Secondary Schooling*. *The Economic Journal*, **112**, 705–734.

CASSASUS, J., CUSTATO, S., FROEMEL, J.E. & PALAFOX, J.C. 2002. *First International Comparative Study of Language, Mathematics and Associated Factors for Students in the Third and Fourth Years of Primary School*. Santiago: Latin American Laboratory for Assessment of Quality in Education.

CENTRO NACIONAL DE REGISTROS. 2015a. *Límites Político-Administrativos.* Available at: http://www.cnr.gob.sv/geoportal-cnr/ [Accessed July 29, 2017].

CENTRO NACIONAL DE REGISTROS. 2015b. *Transporte.* Available at: http://www.cnr.gob.sv/geoportal-cnr/ [Accessed July 29, 2017].

CHRISTOFIDES, L.N., CANADA, J.C.B. & HOY, M. 2001. Family Income and Postsecondary Education In Canada. *Canadian Journal of Higher Education*, **31**, 177–208.

CULLINAN, J., FLANNERY, D., WALSH, S. & MCCOY, S. 2013. Distance Effects, Social Class and the Decision to Participate in Higher Education in Ireland. *The Economic and Social Review*, **44**, 19–51.

DAUDE, C. & ROBANO, V. 2015. On intergenerational (im)mobility in Latin America. *Latin American Economic Review*, **24**, 1–29.

DAVIES, R., HEINESEN, E. & HOLM, A. 2002. The relative risk aversion hypothesis of educational choice. *Journal of Population Economics*, **15**, 683–713.

DEININGER, K. & SQUIRE, L. 1996. A New Data Set Measuring Income Inequality. *The World Bank Economic Review*, **10**, 565–591.

DIRECCIÓN GENERAL DE ESTADÍSTICA Y CENSOS (DIGESTYC) & MINISTERIO DE ECONOMÍA (MINEC). 2016. Encuesta de Hogares de Propósitos Múltiples 2015.

DURYEA, S., GALIANI, S., ÑOPO, H. & PIRAS, C. 2007. The Educational Gender Gap in Latin America and the Caribbean. Inter-American Development Bank, Research Department. *4510*.

ELIASSON, K. 2006. The Effects of Accessibility to University Education on Enrollment Decisions, Geographical Mobility, and Social Recruitment. Umeå University, Department of Economics. *690*.

ERIKSON, R. & JONSSON, J.O. 1996. *Can Education be Equalized?: The Swedish Case in Comparative Perspective*. Westview Press, 312 pp.

ESSER, H. 1999. *Soziologie. Spezielle Grundlagen: Band 1: Situationslogik und Handeln*. 1st ed. Frankfurt/Main ; New York: Campus Verlag, 532 pp.

EUROPEAN COMMISSION, JOINT RESEARCH CENTRE (JRC) & COLOMBIA UNIVERSITY, CENTER FOR INTERNATIONAL EARTH SCIENCE INFORMATION NETWORK - CIESIN. 2015. *GHS population grid, derived from GPW4, multitemporal (1975, 1990, 2000, 2015)*. European Commission, Joint Research Centre (JRC). Available at: http://data.europa.eu/89h/jrc-ghsl-ghs_pop_gpw4_globe_r2015a [Accessed July 29, 2017].

FARLEY ORDOVENSKY, J. 1995. Effects of institutional attributes on enrollment choice: Implications for postsecondary vocational education. *Economics of Education Review*, **14**, 335–350.

FLANNERY, D. & O'DONOGHUE, C. 2009. The determinants of higher education participation in Ireland: a micro analysis. *Economic & Social Research Institute*, **40**, 73–107.

FRENETTE, M. 2004. Access to College and University: Does Distance to School Matter? *Canadian Public Policy*, **30**, 427–443.

FRENETTE, M. 2009. Do universities benefit local youth? Evidence from the creation of new universities. *Economics of Education Review*, **28**, 318–328.

FRENETTE, M. 2006. Too Far to Go On? Distance to School and University Participation. *Education Economics*, **14**, 31–58.

FULLER, W.C., MANSKI, C. & WISE, D. 1982. New Evidence on the Economic Determinants of Postsecondary Schooling Choices. *Journal of Human Resources*, **17**, 477–498.

GAMBETTA, D. 1987. *Were they Pushed or did they Jump?: Individual decision mechanisms in education*. Cambridge: Cambridge University Press.

GIANNELLI, G.C. & MONFARDINI, C. 2003. Joint decisions on household membership and human capital accumulation of youths. The role of expected earnings and local markets. *Journal of Population Economics*, **16**, 265–285.

GIBBONS, S. & VIGNOLES, A. 2012. Geography, choice and participation in higher education in England. *Regional Science and Urban Economics*, **42**, 98–113.

GILLIES, D. 2015. Human Capital Theory in Education. *In* Peters, M., *ed. Encyclopedia of Educational Philosophy and Theory*. Springer Singapore, 1–5.

HAGENAARS, A.J.M., VOS, K. DE, ZAIDI, M.A. & STATISTICAL OFFICE OF THE EUROPEAN COMMUNITIES. 1994. *Poverty statistics in the late 1980s: research based on micro-data*. Luxembourg: Office for Official Publications of the European Communities.

HERTZ, T., TAMARA, J., PIRAINO, P., SIBEL, S., NICOLE, S. & VERASHCHAGINA, A. 2008. The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends. *The B.E. Journal of Economic Analysis & Policy*, **7**, 1–48.

IGARAPÉ INSTITUTE. 2016. Homicide Monitor. Available at: https://homicide.igarape.org.br/ [Accessed July 26, 2017].

JHA, J., BAKSHI, S. & FARIA, E.M. 2012. Understanding and challenging boys' disadvantage in secondary education in developing countries. The United Nations Children's Fund (UNICEF). Available at: http://unesdoc.unesco.org/Ulis/cgi-bin/ulis.pl?catno=217868&set=0059568ED9_1_468&gp=&lin=1&ll=f [Accessed June 30, 2017].

KANE, T.J. 2006. Public Intervention in Post-Secondary Education. *In* Hanushek, E.A. & Finis, W., *eds. Handbook of the Economics of Education*. Elsevier, 1369–1401.

KAZEEM, A., JENSEN, L. & STOKES, C.S. 2010. School Attendance in Nigeria: Understanding the Impact and Intersection of Gender, Urban-Rural Residence and Socioeconomic Status. *Comparative education review*, **54**, 295–319.

MCCRACKEN, J.D. & BARCINAS, J.D.T. 1991. Differences between Rural and Urban Schools, Student Characteristics, and Student Aspirations in Ohio. *Journal of Research in Rural Education*, **7**, 29–40.

MICKLEWRIGHT, J., PEARSON, M. & SMITH, S. 1990. Unemployment and Early School Leaving. *The Economic Journal*, **100**, 163–169.

MINCER, J. 1958. Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*, **66**, 281–302.

MINISTERIO DE EDUCACIÓN EL SALVADOR (MINED). 2014. Instituciones de Educación Superior Legalmente establecidas. *Educación Superior.* Available at: https://www.mined.gob.sv/index.php/programas-educativos/educacion-superior [Accessed July 29, 2017].

MURAKAMI, Y. & BLOM, A. 2008. *Accessibility and Affordability of Tertiary Education in Brazil, Colombia, Mexico and Peru within a Global Context* (ID 1094224). Rochester, NY: Social Science Research Network.

OECD. 2008. *Latin American Economic Outlook 2009*. Paris: Organisation for Economic Co-operation and Development. Available at: http://www.oecd-ilibrary.org/content/book/leo-2009-en.

OPENSTREETMAP. 2014. *Roads*. MapCruzin. Available at: http://www.mapcruzin.com/free-el-salvador-country-city-place-gis-shapefiles.htm [Accessed July 29, 2017].

RAMA VITALE, C. 2007. Los nuevos estudiantes en circuitos diferenciados de educación. *In* Rodríguez Ortiz, G., *ed. Seminario Educación superior en América Latina y el Caribe: Sus estudiantes hoy*. Mexico: Unión de Universidades de América Latina y el Caribe, 13–42.

RIZZICA, L. 2013. Home or away? Gender differences in the effects of an expansion of tertiary education supply. Bank of Italy, Economic Research and International Relations Area. *181*.

SÁ, C., FLORAX, R. & RIETVELD, P. 2004. Determinants of the Regional Demand for Higher Education in The Netherlands: A Gravity Model Approach. *Regional Studies*, **38**, 375–392.

SCHMELKES, S. 2000. Education and Indian peoples in Mexico: An example of policy failure. *In* Reimers, F., *ed. Unequal schools, unequal chances challenges to equal opportunity in the Americas*. Cambridge: MA: Harvard University Press, 319–333.

SPIESS, K.C. & WROHLICH, K. 2010. Does distance determine who attends a university in Germany? *Economics of Education Review*, **29**, 470–479.

STOCKÉ, V. 2007. Explaining Educational Decision and Effects of Families' Social Class Position: An Empirical Test of the Breen Goldthorpe Model of Educational Attainment. *European Sociological Review*, **23**, 505–519.

STROMQUIST, N.P. 2004. Inequality as a way of life: education and social class in Latin America. *Pedagogy, Culture & Society*, **12**, 95–119.

TORCHE, F. 2014. Intergenerational Mobility and Inequality: The Latin American Case. *Annual Review of Sociology*, **40**, 619–642.

TURLEY, R.N.L. 2009. College Proximity: Mapping Access to Opportunity. *Sociology of Education*, **82**, 126–146.

ULUBAŞOĞLU, M.A. & CARDAK, B.A. 2007. International comparisons of rural–urban educational attainment: Data and determinants. *European Economic Review*, **51**, 1828–1857.

UNITED NATIONS CHILDREN'S FUND (UNICEF). 2003. The State of the World's Children 2004 - Girls, Education and Development. The United Nations Children's Fund (UNICEF). Available at: https://www.unicef.org/publications/index_18108.html [Accessed June 30, 2017].

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION (UNESCO) - IBE. 2010. Datos Mundiales de Educación. *7a edición.* Available at: http://www.ibe.unesco.org/fileadmin/user_upload/Publications/WDE/2010/pdf-versions/El_Salvador.pdf [Accessed July 29, 2017].

WESSLING, K.D. 2016. *The influence of socio-spatial contexts on transitions from school to vocational and academic training in Germany*. Dissertation, Universität Tübingen.

WINKLER, D.R. 1990. Higher education in Latin America : issues of efficiency and equity (WDP 77). The World Bank. Available at: http://documents.worldbank.org/curated/en/125181468772468327/Higher-education-in-Latin-America-issues-of-efficiency-and-equity [Accessed July 26, 2017].

WORLD BANK. 2016. *World Development Report 2018 : realizing the promise of education for development - concept note*. Washington, D.C.: World Bank Group. Available at: http://documents.worldbank.org/curated/en/667251485486049489/World-Development-Report-2018-realizing-the-promise-of-education-for-development-concept-note [Accessed July 26, 2017].

Appendices

Appendix 1: List of postsecondary institutions in El Salvador (table) 49

Appendix 2: Road network in El Salvador (map) 51

Appendix 3: Assigned speed to the different road types (table) 52

Appendix 4: Travel distance to the closest postsecondary institution (map). 52

Appendix 5: Sample comparison (table) 53

Appendix 6: Sample comparison (figure) 53

Appendix 7: Effects of additional institutions on study probability (table) 54

Appendix 8: Effects of additional study programs on study probability (table) 54

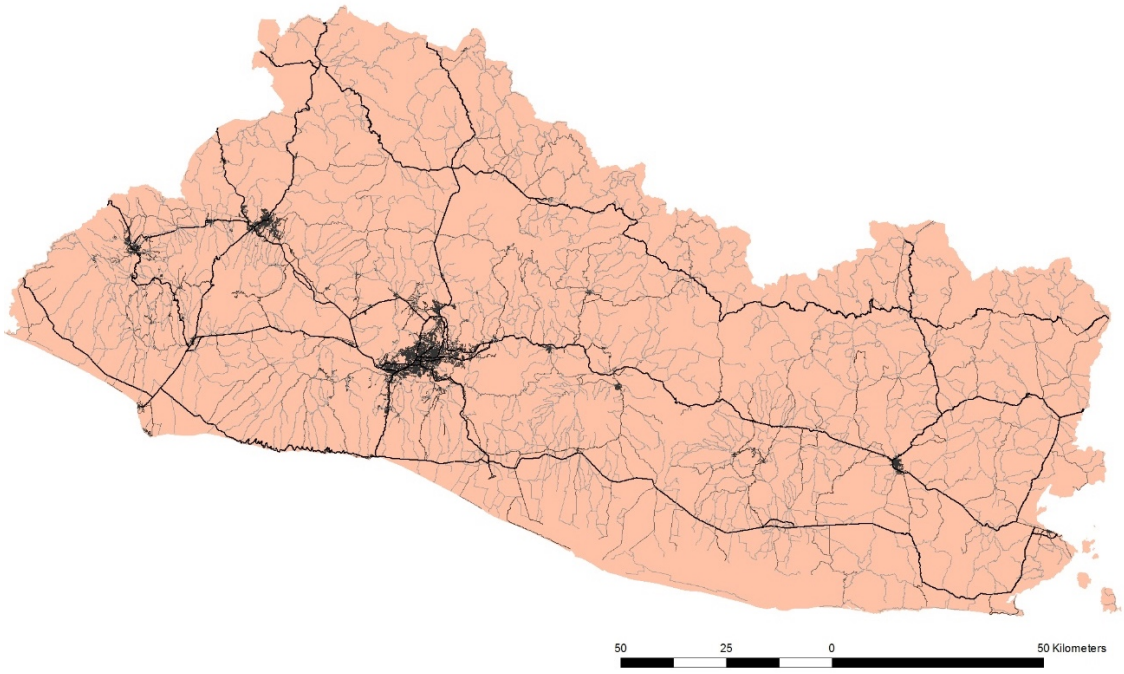
Appendix 9: Stata script 55

Name	Location	#programs	Public	X	Y	Fees p-a ($)	Homepage
Universidad de El Salvador (UES)	San Miguel	34	y	-88,1595580	13,4397200	63.30	www.fmoues.edu.sv
Universidad de El Salvador (UES)	San Salvador	59	y	-89,2038810	13,7182560	63.30	www.ues.edu.sv
Universidad de El Salvador (UES)	San Vicente	11	y	-88,7900290	13,6427320	63.30	www.fmp.ues.edu.sv
Universidad de El Salvador (UES)	Santa Ana	31	y	-89,5642330	13,9799460	63.30	www.uesocc.edu.sv
Universidad Panamericana (UPAN)	Ahuachapán	8	n	-89,8477600	13,9223540	930.00	www.upan.edu.sv
Universidad Monseñor Oscar Arnulfo Romero	Chalatenango	6	n	-89,1026970	14,1698220	720.00	www.umoar.edu.sv
Universidad Dr. Andrés Bello (UNAB)	Chalatenango	20	n	-88,9797720	14,0449640	NI	www.unab.edu.sv
Universidad Albert Einstein (UAE)	Antiguo Cuscatlán	9	n	-89,2352140	13,6745880	NI	www.uae.edu.sv
Universidad Centroamericana José Simeón Cañas (UCA)	Antiguo Cuscatlán	23	n	-89,2350070	13,6813260	915.00	www.uca.edu.sv
Universidad Dr. José Matías Delgado (UJMD)	Antiguo Cuscatlán	24	n	-89,2547940	13,6806850	NI	www.ujmd.edu.sv
Universidad Técnica Latinoamericana (UTLA)	Santa Tecla	8	n	-89,2849760	13,6778860	NI	www.utla.edu.sv
Instituto Tecnológico Centroamericano (ITCA)	Santa Tecla	19	n	-89,2786820	13,6732610	1500.00	www.itca.edu.sv
Escuela de Comunicación Mónica Herrera	Santa Tecla	3	n	-89,2879380	13,6773320	NI	www.monicaherrera.com

Escuela Nac. de Agricultura Roberto Quiñonez (ENA)	Ciudad Arce	1	n	-89,4013220	13,8050400	NI	www.ena.edu.sv
Inst. Esp. de Educación Superior El Espíritu Santo (IEESES)	Santa Tecla	5	n	-89,2714100	13,6802910	NI	www.ieeses.edu.sv
Instituto Tecnológico Centroamericano (ITCA) (Megatec Zacatecoluca)	Zacatecoluca	4	n	-88,8664570	13,4975780	1500.00	www.itca.edu.sv
Instituto Tecnológico Centroamericano (ITCA) (Megatec La Unión)	La Unión	8	n	-87,8594300	13,3393170	1500.00	www.itca.edu.sv
Universidad Capitán General Gerardo Barrios (UCGB)	San Miguel	21	n	-88,1927580	13,4886140	880.00	www.ugb.edu.sv
Universidad Dr. Andrés Bello (UNAB)	San Miguel	15	n	-88,1767100	13,4767190	NI	www.unab.edu.sv
Instituto Tecnológico Centroamericano (ITCA)	San Miguel	3	n	-88,1631090	13,4997740	1500.00	www.itca.edu.sv
Universidad de Oriente (UNIVO)	San Miguel	22	n	-88,1836680	13,4814070	840.00	www.univo.edu.sv
Universidad Cristiana de las Asambleas de Dios (UCAD)	San Salvador	8	n	-89,1889300	13,7122530	NI	www.ucad.edu.sv
Universidad Don Bosco (UDB)	Soyapango	34	n	-89,2368120	13,6744080	782.00	www.udb.edu.sv
Universidad Dr. Andrés Bello (UNAB)	San Salvador	21	n	-89,2116200	13,7025550	NI	www.unab.edu.sv
Universidad Evangélica de El Salvador - UEES	San Salvador	20	n	-89,2390010	13,7170070	1360.00	www.uees.edu.sv
Universidad Francisco Gavidia (UFG)	San Salvador	35	n	-89,2202210	13,6998100	1260.00	www.ufg.edu.sv
Universidad Luterana Salvadoreña (ULS)	San Salvador	7	n	-89,1943220	13,6761950	NI	www.uls.edu.sv
Universidad Modular Abierta (UMA)	San Salvador	14	n	-89,2175820	13,7027030	750.00	www.uma.edu.sv
Universidad Panamericana (UPAN)	San Salvador	9	n	-89,2212450	13,7095870	1030.00	www.upan.edu.sv
Universidad Pedagógica de El Salvador (UPED)	San Salvador	22	n	-89,2035340	13,7077330	674.00	www.pedagogica.edu.sv
Universidad Politécnica de El Salvador (UPES)	San Salvador	10	n	-89,1945460	13,7089560	730.00	www.upes.edu.sv
Universidad Nueva San Salvador (UNSSA)	San Salvador	10	n	-89,2127160	13,6999530	740.00	www.unssa.edu.sv
Universidad Salvadoreña Alberto Masferrer (USAM)	San Salvador	15	n	-89,2005470	13,7031930	NI	www.usam.edu.sv
Universidad Tecnológica de El Salvador (UTEC)	San Salvador	38	n	-89,2016780	13,7006480	810.00	www.utec.edu.sv
Escuela Superior de Economía y Negocios (ESEN)	San Salvador	3	n	-89,2863090	13,6552780	5600.00	www.esen.edu.sv
Inst. de Ed. Sup. de Profesionales de la Salud (IEPROES)	San Salvador	4	n	-89,2109840	13,6959870	942.00	www.ieproes.edu.sv
Universidad Panamericana (UPAN)	San Vicente	7	n	-88,7854880	13,6459240	1030.00	www.upan.edu.sv

Universidad Autónoma de Santa Ana (UNASA)	Santa Ana	7	n	-89,5886660	13,9764800	NI	www.unasa.edu.sv
Universidad Francisco Gavidia (UFG) (CENTRO REGIONAL DE OCCIDENTE)	Santa Ana	12	n	-89,5653930	13,9925770	1260.00	www.ufg.edu.sv
Instituto Tecnológico Centroamericano (ITCA) (Regional Santa Ana)	Santa Ana	4	n	-89,5644270	13,9807280	1500.00	www.itca.edu.sv
Universidad Católica de El Salvador (UNICAES)	Santa Ana	36	n	-89,5476910	13,9823270	890.00	www.catolica.edu.sv
Universidad de Sonsonate (USO)	Sonsonate	13	n	-89,7188560	13,7307350	NI	www.usonsonate.edu.sv
Universidad Dr. Andrés Bello (UNAB)	Sonsonate	12	n	-89,7220360	13,7210360	NI	www.unab.edu.sv
Escuela Superior Franciscana Especializada/Agape (ESFE/AGAPE)	Sonzacate	3	n	-89,7087810	13,7368860	670.00	www.esfe.agape.edu.sv
Universidad Capitán General Gerardo Barrios (UCGB)	Usulután	16	n	-88,4183050	13,3417820	880.00	www.ugb.edu.sv
Instituto Tecnológico de Usulután (ITU)	Usulután	5	n	-88,4491410	13,3348860	NI	www.itu.edu.sv

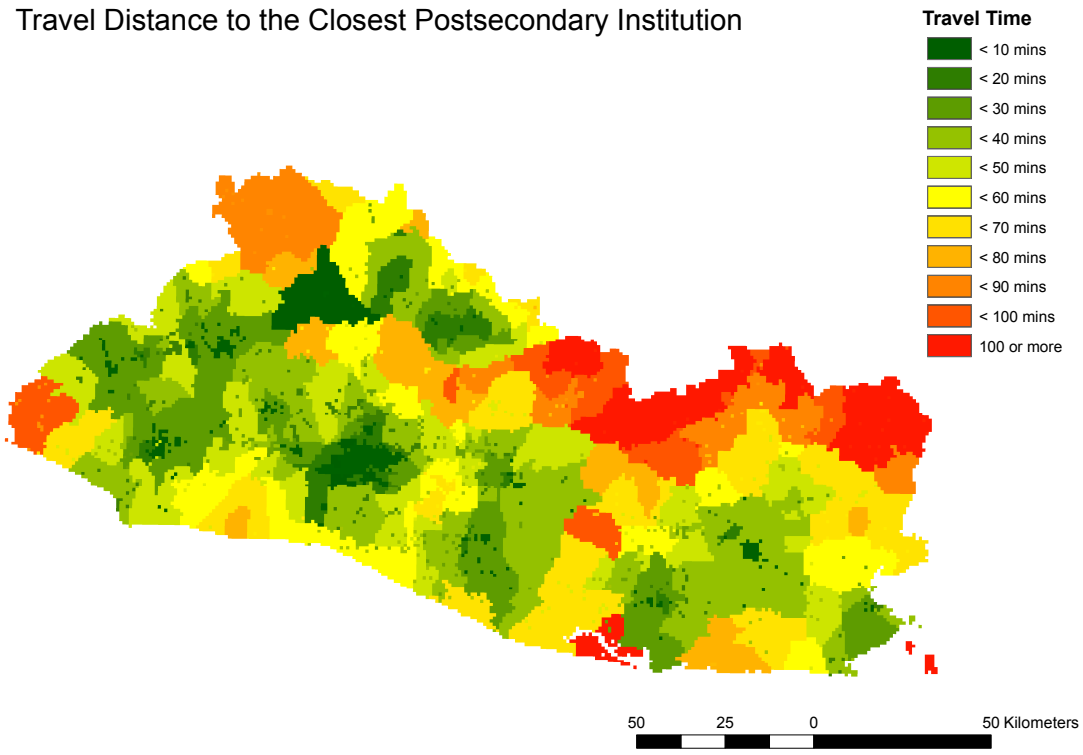
Appendix 1: List of postsecondary institutions in El Salvador. Institutions only offering postgraduate programs, military schools, and institutions that opened after 2014 were excluded. Data sources: websites of the institutions, MINED and ALTILLO.



Appendix 2: Road network in El Salvador. Data: Centro Nacional de Registros (2015b) and OpenStreetMap (2014).

Category	Description	Paving (MAJ)	Assigned speed
0	*Especial*	Paved	60 km/h
1	*Primaria*	Paved	50 km/h
2	*Secundaria*	Paved	40 km/h
3	*Terciaria*	Not paved	25 km/h
4	*Terciaria modificada*	Paved	30 km/h
5	*Rural*	Not paved	20 km/h
6	*Rural modificada*	Paved	25 km/h
7	No street	-	4 km/h

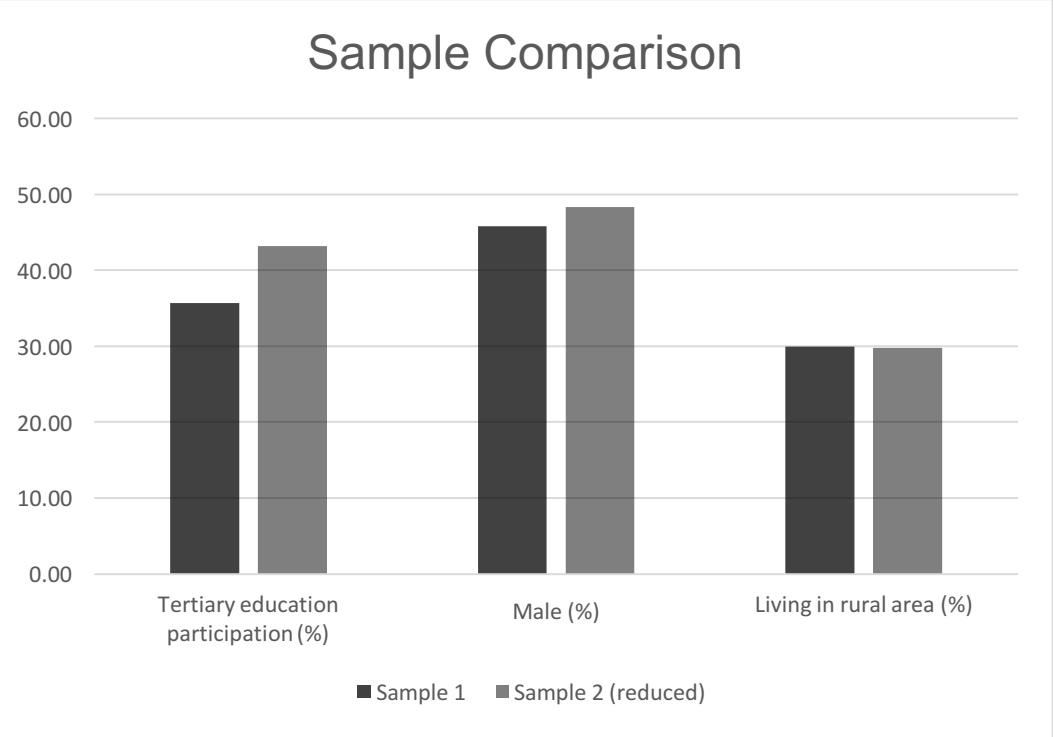
Appendix 3: Assigned speed to the different road types.



Appendix 4: Travel distance to the closest postsecondary institution in El Salvador. Spatial resolution: Municipalities, divided into rural/urban.

	Sample 1	Sample 2 (reduced)
Tertiary education participation (%)	35.71	43.16
Male (%)	45.81	48.31
Living in rural area (%)	29.98	29.77
Hours to the closest institution	0.56	0.56
Age (mean)	23.50	22.90
Sample size	**8,526**	**4,951**

Appendix 5: Comparison of individual and spatial characteristics of the used samples. Sample 1 (n=8,526) consists of individuals between 18 and 30 years old who graduated from secondary school. Sample 2 (reduced) (n=4,951) consists of individuals between 18 and 30 years old who graduated from secondary school and are still living in the parental home (coded as son/daughter in a household).



Appendix 6: Comparison of individual and spatial characteristics of the used samples, presented as a graphic. Sample 1 (n=8,526) consists of individuals between 18 and 30 years old who graduate from secondary school. Sample 2 (reduced) (n=4,951) consists of individuals between 18 and 30 years old who graduated from secondary school and are still living in the parental home (coded as son/daughter in a household).

	30 min	45 min	60 min	90 min	120 min	150 min	180 min
Supply of postsecondary	**0.0067*** **	**0.0045*** **	**0.0023***	**0.0031*** **	**0.0021****	0.00035	-0.00059
educational opportunities	(0.0012)	(0.0011)	(0.00096)	(0.00083)	(0.00078)	(0.00082)	(0.00082)
N=individuals	8,526	8,526	8,526	8,526	8,526	8,526	8,526

*Appendix 7: Effects of additional institutions on the probability of participating in tertiary education within different travel time radii, presented as AMEs; significance level: *p<0.05 **p<0.005, ***p<0.001, based on the complete sample, additional control variables: age, sex, area (rural/urban).*

	30 min	45 min	60 min	90 min	120 min	180 min
Supply of postsecondary	**0.000367*** **	**0.000250*** **	**0.000131***	**0.000170*** **	**0.000130****	0.0000921
study programs	(0.000066)	(0.000061)	(0.000055)	(0.000049)	(0.000048)	(0.000050)
N=individuals	8,526	8,526	8,526	8,526	8,526	8,526

*Appendix 8: Effects of additional study programs on the probability of participating in tertiary education within different travel time radii, presented as AMEs; significance level: *p<0.05 **p<0.005, ***p<0.001, based on the complete sample (n=8,526), additional control variables: age, sex, area (rural/urban).*

```
1  ******************************************************************
2  ** DO FILE *******************************************************
3  ** BSC THESIS -- LIVIA JAKOB *************************************
4  ******************************************************************
5
6  **INITIAL SETTINGS**********************************
7  *set working directory
8  cd /Applications/Stata/ba_files
9
10 *create log-file
11 log using ba.log, append
12
13 **close logfile
14 *log close
15
16 **DATA SETTINGS & IMPORTS*****************************
17
18 ** load dataset
19 use "EHPM 2015.DTA"
20
21 ** save dataset
22 save "EHPM 2015.DTA", replace
23
24 ** import traveltimes
25 import delim using traveltimes.csv, delim(",")
26 rename cod_mun4 r005
27 rename urb_rur area
28 save traveltimes.dta, replace
29
30 import delim using traveltimes2.csv, delim(",")
31 rename cod_mun4 R005
32 rename urb_rur AREA
33 save traveltimes2.dta, replace
34
35 import delim using traveltimes3.csv, delim(",")
36 rename cod_mun4 R005
37 rename urb_rur AREA
```

```
save traveltimes3.dta, replace

** merge traveltime with ehpm dataset
merge m:1 r005 area using traveltimes.dta
drop if _merge==2
save "EHPM 2015.DTA", replace

merge m:1 R005 AREA using traveltimes2.dta
drop if _merge==2
save "EHPM 2015.DTA", replace

merge m:1 R005 AREA using traveltimes3.dta
drop if _merge==2
save "EHPM 2015.DTA", replace

** Upper case variables
foreach v of varlist _all {
    capture rename `v' `=upper("`v'")'
}


*************************************************************
** SURVEY SETTINGS **********************************************

*VAR -- Generate household IDs ****

gen HH_temp=0
sort IDBOLETA
replace HH_temp=HH_temp+1 if IDBOLETA!=IDBOLETA[_n-1]
gen HH_ID=sum(HH_temp)
drop HH_temp

gen PERS_ID =_n

*VAR -- fpc1 = Number of PSUs (in population)****
sum LOTE // According to EHPM Publication 2012: 12423
gen fpc1=12423
```

```
 75
 76 *VAR -- fpc2 = number of households (in population)****
 77 sum HH [w=FAC]
 78 // Sample: 88 184 people,  23670 households;  //Population: 6 459 911 people
 79 sum R021A if R101==1 [w=FAC00]
 80 sum MIEMH  if R101==1 [w=FAC00] // 3.666712 (average # of household members)
 81 dis 6459911/3.66647
 82 gen fpc2= 1761887 // estimated number of households in population
 83
 84 drop if LOTE==3254 // Error in data!
 85
 86 *SVY -- set survey structure****
 87  svyset  LOTE  [pweight=FAC00],  strata(ESTRATOAREA)  fpc(fpc1)  ||  HH_ID,
fpc(fpc2)
 88
 89 svyset LOTE [pweight=FAC00], strata(ESTRATOAREA)|| HH_ID
 90 // without finite population correction
 91
 92
 93 ************************************************************************
 94 ** GENERATE VARIABLES (VAR) ***************************************
 95
 96 clonevar age = R106
 97 clonevar sex = R104
 98 clonevar fam_income = INGFA
 99 clonevar studying = R203 // 1=yes ;; 2=no
100
101 *VAR -- highest level of education
102 clonevar educ_level = R217A // not recorded when studying
103 tab educ_l
104 tab educ_level, nolab mis
105 recode educ_level (1=0) (2=1) (3=2) (4=3) (5=3) (6=.) (7=.) (8=0)
106 label define educ_level 0 "No Education/Kindergarten" 1"Basic Education" ///
107 2"Secondary Education" 3"Tertiary Education"
108 label values educ_level educ_level
109
110 *VAR -- actual level of education
111 clonevar study_what_ = R204
```

```
clonevar study_what = R204 // does not recorded when not studying at the moment
recode study_what (1=0) (2=1) (3=2) (4=3) (5=3) (6=.)
label values study_what educ_level //label
tab study_what

clonevar level_when_study = study_what
tab level_when
tab level_when, nolab
recode level_when (1=0) (2=1) (3=2) (4=2) //code the highest achieved level
tab level_when

*VAR -- merge highest level of educ for studying and finished studies****
replace educ_level = level_when if level_when!=.


tab educ_level
clonevar ever_studied = R215
tab ever, nolab
replace educ_level = 0 if ever_studied==2
tab educ_level if age>4, mis

*VAR -- University****
tab R204, nolab // 4 = uni
gen uni_studying = R204==4
gen uni_level = R217A==4
gen uni_both = uni_level + uni_study

*VAR -- Tertiary****
tab R204, nolab // 4 = uni, 5 =tecnico
gen tert_studying = R204==4 | R204==5
gen tert_level = R217A==4 | R217A==5
gen tert_both = tert_level + tert_study //finished tertiary or still studying

*VAR -- Secondary****
tab R204
tab R204, nolab // 3 = High school
gen secondary_study = R204==3
```

```
149 gen secondary_level = educ_level>=2
150 gen secondary_both = secondary_study + secondary_level
151
152 *VAR -- Primary****
153 gen primary_study= R204==2
154 gen primary_level = educ_level>=1
155 gen primary_both = primary_study + primary_level
156
157
158 *VAR -- Status: Parents highest educational attainment ****
159
160 gen individual = R103 == 3 //R103: 3=son/daughter ;; individual=1 when hijo/a
161 gen jefe = .
162 replace jefe = educ_level if R103==1 // R103: 1=jefe(head)
163 sort HH_ID
164 egen level_jefe = max(jefe), by(HH_ID)
165 tab level_jefe
166
167 gen esposa = . //wife or husband (note: also stepparents allowed)
168 replace esposa = educ_level if R103==2
169 sort HH_ID
170 egen level_esposa = max(esposa), by(HH_ID)
171 tab level_esposa
172
173 gen father = .
174 replace father = level_jefe if R103==1 & sex==1
175 egen level_father = max(father), by(HH_ID)
176 tab level_father
177
178 gen mother = .
179 replace mother = level_jefe if R103==1 & sex==2
180 replace mother = level_esposa if R103==2
181 egen level_mother = max(mother), by(HH_ID)
182
183 tab level_mother
184 label values level_m level_f level_e level_j educ_level
185
```

```
svy: tab level_mother if individual==1
svy: tab level_father if individual==1

egen level_both_max = rowmax(level_mother level_father) //max of both
label define levels_educ_ 0 "No education" 1"Primary" ///
2"Secondary" 3"Tertiary"
label values level_both_max levels_educ_

*VAR -- Household equivalence income****
tab MIEMH
gen plus14 = age>=14
sort HH_ID
egen miem14plus = sum(plus14), by(HH_ID)
gen zeroto13 = age <14
egen miem_under14 = sum(zeroto13 ), by(HH_ID)
gen memb_weight = 0.5 + 0.5*miem14plus + 0.3*miem_under14
//OECD definition: adult 1 is counted double: 2*0.5=1
tab memb_w

gen household_equivalence_income = fam_income/memb_w
gen log_household_equivalence_income =ln(household_equivalence_income)

*VAR -- Household equivalence income without own income****
gen household_income2 = fam_income-INGRE
gen equivalence_income2 = household_income2/memb_w
gen log_equivalence_income2 =ln(equivalence_income2)
label variable log_equivalence_income2 "Family income"

*VAR -- square root household income****
gen sqrt_income = fam_income/sqrt(MIEMH)

*VAR Income Quartiles
xtile income_quartile=household_equivalence_income [w=FAC], n(4)
xtile income_decile=household_equivalence_income [w=FAC], n(10)

*VAR -- children secondary_level, where famliy characteristics are known
gen child_secondary = (secondary_level == 1 & individual==1 ///
```

```
223 & age <30 & age >18)
224
225 *VAR -- secondary level, where famliy characteristics are not known
226 gen notchild_secondary = (secondary_level == 1 & individual==0 ///
227 & age <30 & age >18)
228
229 *VAR -- Hours to closest tertiary institution
230 gen minhours= MINDIST/60
231 label variable minhours `" "Hours to the"  "closest institution" "'
232
233 *VAR -- Hours to closest public university
234 gen pub_minhours= min_publ/60
235 label variable pub_minhours  "Hours to closest public university"
236
237 *VAR -- Income: Half, Tertile, Quartile
238 ssc install egenmore
239 sort log_equivalence_income2
240 egen rich = xtile(log_equivalence_income2), nq(2)
241
242 xtile income_third=log_equivalence_income2 [w=FAC], n(3)
243 label define tertile_label 1 "First Tertile" 2"Second Tertile" ///
244 3"Third Tertile"
245 label values income_third tertile_label
246
247 xtile income_quartile2=log_equivalence_income2 [w=FAC], n(4)
248 label define quartile_label 1 "First Quartile" 2"Second Quartile" ///
249 3"Third Quartile" 4"Fourth Quartile"
250 label values income_quartile2 quartile_label
251
252 egen income_third_whithInd = xtile(log_equivalence_income), nq(3)
253 label values income_third_whithInd tertile_label
254
255 *VAR -- centered income
256 ssc install center
257 center log_equivalence_income2, generate(c_income)
258
259 *VAR -- Labels
```

```
label define area_label 0 "Rural" 1"Urban"
label values AREA area_label
label define gender_label 1 "Male" 2"Female"
label values sex gender_label

**LAYOUT*************************************************************
set scheme s2mono


**********************************************************************
*CALCULATIONS **************

*CALC -- MODEL 1: Time Radii with universities ***********************
svy: logit tert_both UNI30 i.AREA i.sex age if age <30 ///
& age >18 & secondary_level==1
margins , dydx(*) post
estimates store Radius1

svy: logit tert_both UNI45 i.AREA i.sex age if age <30 ///
& age >18 & secondary_level==1
margins , dydx(*) post
estimates store Radius2

svy: logit tert_both UNI60 i.AREA i.sex age if age <30 ///
& age >18 & secondary_level==1
margins , dydx(*) post
estimates store Radius3

svy: logit tert_both UNI90 i.AREA i.sex age if age <30 ///
& age >18 & secondary_level==1
margins , dydx(*) post
estimates store Radius4

svy: logit tert_both UNI120 i.AREA i.sex age if age <30 ///
& age >18 & secondary_level==1
margins , dydx(*) post
estimates store Radius5
```

```
297
298 svy: logit tert_both uni150 i.AREA i.sex age if age <30 ///
299 & age >18 & secondary_level==1
300 margins , dydx(*) post
301 estimates store Radius6 //not significant
302
303 svy: logit tert_both UNI180 i.AREA i.sex age if age <30 ///
304 & age >18 & secondary_level==1
305 margins , dydx(*) post
306 estimates store Radius7 //not significant
307
308 *Table*
309 esttab Radius1 Radius2 Radius3 Radius4 Radius5 Radius6 Radius7, ar2 se ///
310 varlabels( ///
311 2.sex "Gender (Female)" 0.AREA "Area (Ref. Rural)" 1.AREA "Urban" ///
312 age "Age" ) starlevels(* 0.05 ** 0.01 *** 0.001)
313
314
315 *CALC -- MODEL 1. Time Radii with Carreras (study programs)
316 svy: logit tert_both CARR30 i.AREA i.sex age if age <30 ///
317 & age >18 & secondary_level==1
318 margins , dydx(*) post
319 estimates store Carr1
320
321 svy: logit tert_both CARR45 i.AREA i.sex age if age <30 ///
322 & age >18 & secondary_level==1
323 margins , dydx(*) post
324 estimates store Carr2
325
326 svy: logit tert_both CARR60 i.AREA i.sex age if age <30 ///
327 & age >18 & secondary_level==1
328 margins , dydx(*) post
329 estimates store Carr3
330
331 svy: logit tert_both CARR90 i.AREA i.sex age if age <30 & ///
332 age >18 & secondary_level==1
333 margins , dydx(*) post
```

```
estimates store Carr4

svy: logit tert_both CARR120 i.AREA i.sex age if age <30 & ///
age >18 & secondary_level==1
margins , dydx(*) post
estimates store Carr5

svy: logit tert_both CARR180 i.AREA i.sex age if age <30 & ///
age >18 & secondary_level==1
margins , dydx(*) post
estimates store Carr6 //not significant

*Table*
esttab Carr1 Carr2 Carr3 Carr4 Carr5 Carr6, ar2 se ///
varlabels( ///
2.sex "Gender (Female)" 0.AREA "Area (Ref. Rural)" ///
1.AREA "Urban" age "Age") starlevels(* 0.05 ** 0.01 *** 0.001)
//similar effects like universities


*CALC -- MODEL 1: Time Radii with Carreras (study programs)
svy: logit tert_both CARR30 UNI30 i.AREA i.sex age if age <30 ///
& age >18 & secondary_level==1
margins , dydx(*) post
estimates store Carr11

svy: logit tert_both CARR45 UNI45 i.AREA i.sex age if age <30 ///
& age >18 & secondary_level==1
margins , dydx(*) post
estimates store Carr22

svy: logit tert_both CARR60 UNI60 i.AREA i.sex age if age <30 ///
& age >18 & secondary_level==1
margins , dydx(*) post
estimates store Carr33

```

```
svy: logit tert_both CARR90 UNI90 i.AREA i.sex age if age <30 ///
& age >18 & secondary_level==1
margins , dydx(*) post
estimates store Carr44

svy: logit tert_both CARR120 UNI120 i.AREA i.sex age if age <30 ///
& age >18 & secondary_level==1
margins , dydx(*) post
estimates store Carr55

svy: logit tert_both CARR180 UNI120 i.AREA i.sex age if age <30 ///
& age >18 & secondary_level==1
margins , dydx(*) post
estimates store Carr66

esttab Carr11 Carr22 Carr33 Carr44 Carr55 Carr66, ar2 se ///
varlabels(2.sex "Gender (Female)" 0.AREA "Area (Ref. Rural)" ///
1.AREA "Urban" age "Age") starlevels(* 0.05 ** 0.01 *** 0.001)
//study programs have no significant effect anymore


****CALC MODEL 2 -- Without parents attributes
svy: logit tert_both c.minhours i.AREA i.sex age if age <30 ///
& age >18 & secondary_level==1
margins , dydx(*) post
estimates store Margins_Superior_NoP
coefplot Margins_Superior_NoP, drop(_cons age) ///
omitted baselevels xline(0) ///
coeflabels(minhours=`" "Hours to the "  "closest institution" "' ///
1.sex="Male (ref.)" 0.AREA="Rural (ref.)", labgap(5)) ysize(2) xsize(3) ///
headings(1.sex= "{bf:Gender}" 0.AREA= "{bf:Area}" ///
minhours = "{bf:Travel distance}" , labgap(5)) ///
title("Model 2: Individual and Spatial Characteristics", span) ///
name(superior_bac2, replace) xtitle(AMEs on study probability)

****CALC MODEL 3 -- With parents attributes (only hijos/hijas)
svy: logit tert_both c.log_equivalence_income2 i.level_both_max ///
```

```
i.AREA i.sex c.age c.minhours ///
if age <30 & age >18 & secondary_level==1 & individual==1
margins , dydx(*) post
estimates store Margins_Superior_Parents
*GRAPH
coefplot Margins_Superior_Parents, ///
drop(_cons age minhours 0.AREA 1.AREA 1.sex 2.sex) ///
omitted baselevels xline(0) ///
coeflabels(0.level_both_max="No education (ref.)" ///
log_equivalence_income2="Equivalence income (log)", labgap(5))  ///
headings(0.level_both_max="{bf:Parents education}" ///
log_equivalence_income2= "{bf:Family income}" , labgap(5)) ///
title("Model 3: Family Income and Background", span) ///
name(superior_bac2, replace) xtitle(AMEs on study probability)

****CALC -- INTERACTIONS ***************************

*CALC INTERACTION 1 -- Interactions Income xx Distance
gen incomeXminhours = c_income*minhours

svy: logit tert_both c.c_income c.minhours ///
c.incomeXminhours i.sex i.AREA c.age ///
if age <30 & age >18 & secondary_level==1 & individual==1
margins , dydx(*) post
estimates store Interaction1
*GRAPH
coefplot Margins_Superior_Interact1, ///
drop(_cons age) omitted baselevels xline(0) ///
coeflabels(0.level_both_max="No education (ref.)" , labgap(5))  ///
headings(0.level_both_max=`"{bf:Parents education}"' ///
0.log_equivalence_income2= "{bf:Family income}" ///
1.sex= "{bf:Gender}" 0.AREA= "{bf:Area}" , labgap(5)) ///
title("Logit Model", span) name(superior_bac2, replace) ///
xtitle(AMEs on Study Probability)


*CALC INTERACTION 2 -- Interactions Parents education xx Distance
```

```
gen lev = level_both_max ==0 if level_both_max !=.
gen levXminhours = lev * minhours
gen lev1 = level_both_max ==1 if level_both_max !=.
gen lev1Xminhours = lev1 * minhours
gen lev2 = level_both_max ==2 if level_both_max !=.
gen lev2Xminhours = lev2 * minhours
gen lev3 = level_both_max ==3 if level_both_max !=.
gen lev3Xminhours = lev3 * minhours

svy: logit tert_both c.minhours i.level_both_max ///
c.lev1X c.lev2X c.lev3X i.AREA i.sex c.age ///
if age <30 & age >18 & secondary_level==1 & individual==1
margins , dydx(*) post
estimates store Interaction2
*GRAPH
coefplot Margins_Superior_Interact, ///
drop(_cons age) omitted baselevels xline(0) ///
coeflabels( , labgap(5))  ///
headings(0.level_both_max=`"{bf:Parents education}"' ///
0.log_equivalence_income2= "{bf:Family income}" ///
1.sex= "{bf:Gender}" 0.AREA= "{bf:Area}" , labgap(5)) ///
title("Interaction Model: Parental Education and Distance", span) ///
name(superior_bac2, replace) ///
xtitle(AMEs on Study Probability)

****CALC INTERACTION 3 -- Interactions Gender xx Distance //not significant

gen men = sex ==1 if sex !=.
gen menXminhours = men * minhours
gen women = sex ==2 if sex !=.
gen womenXminhours = women * minhours

svy: logit tert_both i.sex c.minhours c.womenXminhours i.AREA c.age ///
if age <30 & age >18 & secondary_level==1
margins , dydx(*) post
estimates store Interaction3
*GRAPH
```

```
coefplot Margins_Superior_Interact4, drop(_cons age) ///
omitted baselevels xline(0) ///
coeflabels( , labgap(5))  ///
headings(0.level_both_max=`"{bf:Parents education}"' ///
0.log_equivalence_income2= "{bf:Family income}" ///
1.sex= "{bf:Gender}" 0.AREA= "{bf:Area}" , labgap(5)) ///
title("Interaction Model: PGender and Distance", span) ///
name(superior_bac2, replace) ///
xtitle(AMEs on Study Probability)


**TABLE INTERACTION 1 - 3 ****
esttab Interaction1 Interaction2 Interaction3, ar2 se ///
varlabels(1.sex "Gender (Ref. Male)" ///
2.sex "Female" 0.AREA "Area (Ref. Rural)" 1.AREA "Urban" age "Age" ///
c_income "Family Income (centered)" ///
minhours "Hours to Closest Institution" ///
incomeXmin "Interaction Income and Distance" ///
) starlevels(* 0.05 ** 0.01 *** 0.001)


**********************************************************************
**COMPARING SAMPLES

**SAMPLE 1 without family attributes
tab tert_both if age<30 & age>18 & secondary_level==1
tab sex if age<30 & age>18 & secondary_level==1
tab AREA if age<30 & age>18 & secondary_level==1
mean age if age<30 & age>18 & secondary_level==1
mean log_equivalence_income2 if age<30 & age>18 & secondary_level==1
mean minhours if age<30 & age>18 & secondary_level==1

**SAMPLE 2 for family attributes
tab tert_both if age<30 & age>18 & secondary_level==1 & individual==1
tab sex if age<30 & age>18 & secondary_level==1 & individual==1
tab AREA if age<30 & age>18 & secondary_level==1 & individual==1
mean age if age<30 & age>18 & secondary_level==1 & individual==1
```

```
519 mean log_equivalence_income2 if age<30 & age>18 ///
520 & secondary_level==1 & individual==1
521 mean minhours if age<30 & age>18 & secondary_level==1 & individual==1
```

Appendix 9: Stata script. Tested with Stata 13.0.

Erklärung

gemäss Art. 28 Abs. 2 RSL 05

Name/Vorname:	Livia Lea Jakob
Matrikelnummer:	13-131-578
Studiengang:	Geography
	Bachelor ☑ Master ☐ Dissertation ☐
Titel der Arbeit:	Inequality in Access to Tertiary Education – Evidence from El Salvador
LeiterIn der Arbeit:	Dr. Andreas Heinimann

Ich erkläre hiermit, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Alle Stellen, die wörtlich oder sinngemäss aus Quellen entnommen wurden, habe ich als solche gekennzeichnet. Mir ist bekannt, dass andernfalls der Senat gemäss Artikel 36 Absatz 1 Buchstabe r des Gesetzes vom 5. September 1996 über die Universität zum Entzug des auf Grund dieser Arbeit verliehenen Titels berechtigt ist. Ich gewähre hiermit Einsicht in diese Arbeit.

Bern, 07.08.2017

Ort/Datum

Unterschrift