LUND UNIVERSITY
School of Economics and Management

Master in Economic Development and Growth

Evaluating the Impact of Catastrophic Health Payments on School Interruption: The case of Mexico

Maria Miguel Hernández Lagana
eut13mhe@student.lu.se

Abstract: Catastrophic health disbursements produce numerous injuries in the families experiencing them. Increasing the level of impoverishment and reducing the expenditure on basic needs such as education, are expected to be the most disastrous consequences in which uncovered families may be involved. Through the use of a matching procedure and a diff-in-diff method, the impact of out-of-pocket health expenditures on school interruption is measured for the case of Mexico. Consistent with the theory and recent empirical evidence, the findings expose that households experiencing an unexpected health shocks with disturbing effects on financial stability, have a negative impact on youth living within the affected family to finish schooling on time. Besides, it is found that initial endowments and parental background are determinant factors in the decision-making process when economic hardship is present.

Key words: catastrophic health payments, school interruption, initial endowments, Mexico.

EKHM52
Master thesis, second year (15 credits ECTS)
June 2014
Supervisor: Luciana Quaranta
Examiner: Kerstin Enflo

Acknowledgements

The development of this thesis was a challenging process impossible to be accomplished without the help and support from my family, friends, colleagues and professors. First of all I would like to thank my supervisor, Luciana Quaranta for all her help, comments, guidance and patience throughout the way.

I would like to express my special thanks to my family, Domenica, Miguel and Elvira for supporting and encouraging me in each decision I have taken in my life. Thank you for your love, trust and for giving me the opportunity to live this priceless adventure. Thank you for all your generosity, *un millón de gracias!* Thank you Oier for showing interest always in every single thing I do. For all your love, care and patience especially in the most difficult and stressful moments, I can only thank you immensely.

Thank you all MEDEG people, that more than friends and colleagues you became my family during all this period. Full of great moments, all of you made this master a true adventure. Special thanks to Nada Bedir for helping me improve this thesis with her bright comments and suggestions.

Finally, I would like to thank the Program, especially Kerstin Enflo and Jonas Ljungberg for helping and providing us support with the all administrative-related matters within the university and beyond.

Any mistakes or inaccuracies are my sole responsibility.

Table of Contents

List of Tables and Figures

Evaluating the Impact of Catastrophic Health Payments on School Interruption: The case of Mexico

1. Introduction

Health can be financed through three different mechanisms: (1) with taxes when health is publicly provided; (2) by a combination of public and private contributions, especially in the form of job-insurances; and (3) through private payments, either in the form of non job-related private insurances or in out-of pocket expenditures (Wagstaff & Van Doorslaer, 1998).

In developing countries, catastrophic expenditures or out-of-pocket payments are the main mechanisms in which low-income families afford health services (Van Doorslaer, *et al.*, 2006). The World Bank in its Voices of the Poor Study (2000) indicates that after presenting poor educational levels and scarce labor opportunities, expenditure in health care becomes the most important cause of poverty in the world. Approximately 150 million people are involved in catastrophic health expenditures worldwide, meaning that they are obliged to devote more than 30% of their non-food income on medical care. Moreover, after meeting these payments, the financial catastrophe instigates that around 100 million people fall below the poverty line (WHO, 2007).

A large proportion of households are forced to postpone or cancel medical attention for the patient due to economic hardship, deriving in the impediment to work particularly when the member concerned is economically active (CONEVAL, 2013). Furthermore, in the event of death due to the inability to access medical care, the health shock not only produces a human loss, but also an income injury within the family (WHO, 2007). Likewise, under a context where economic distress is present, most of the households are obliged to sell their limited durable assets to fund healthcare, which may potentially intensify their poverty levels (Knaul, *et al.,* 2006).

However, health shocks produce numerous repercussions other than worsening the health outcomes and the weakening of financial stability. To meet healthcare spending, people may be forced to reduce consumption of basic needs such as food and housing, clothing and education. In this context, interrupting education is a likely event to occur, particularly when

prioritizing healthcare is fundamental (Caneiro & Heckman, 2002). In developing countries, most of the households may diminish their consumption of education as a response of income volatility, predominantly when there are insufficient assets to trade with and the imminent need to pay for health services is present (Kremer, 2003). Consequently, the scarce financial protection for health in low and middle-income countries could be reflected not only in low health outcomes, but also in economic deterioration and low school attainment rates of children within families experiencing unexpected health shocks (Senne, 2010).

Research focused on the effect of health distresses on schooling decisions has been primarily conducted for industrialized countries such as Italy, Norway and the United States (Johnson and Reynolds, 2011). On it, the authors find that health-shocks within households lead to reductions in the probability of completing high school on time of the teenagers living in that household. Moreover, students living within the affected household are unlikely to complete a bachelor degree after the family experienced catastrophic health payments. Thus, the health shock not only damages the well-being of the person experiencing the distress and harms the household economy, but it also increases the probability of school interruption of young members (Johnson and Reynolds, 2011; Beegle, *et al*., 2007). Under these conditions, the reduction of the levels of human capital formation throughout the economy is expected to occur as a long-term consequence (Banerjee and Duflo, 2007).

Little research has been devoted to the analysis of health expenditure and schooling decisions, especially for developing countries. Consequently, the aim of this study is to test whether an income shock deriving into catastrophic expenditures on health within a household has a negative impact on schooling decisions in Mexico. In particular, the research intends to measure the effect of unexpected health distress producing catastrophic payments on school interruption. Doing an extensive literature review, this appears to be the first paper exploring out-of-pocket health payments on educational choices for the case of Mexico and for any other Latin American country.

The intuition behind the study is that conditioned on weak financial conditions in which families are placed after meeting catastrophic expenditures on health, households may be forced to reduce their spending on education. Namely, families reduce their consumption of education in order to reallocate the available endowments in other matters such as health.

This situation may therefore be reflected in the interruption of instruction of the children within the affected household after the health shock occurs.

Mexico becomes an interesting case to illustrate the impact of catastrophic health payments on schooling decisions. Firstly, notwithstanding of its recent improvement, the Mexican health system still presents serious coverage deficiencies. According to the World Health Organization (2012), only 47.3% of the population is covered by some job-related public insurance, while 52.7% covers health care by private expenditures. Furthermore, around 92% of the private health spending represents out-of-pocket expenditures; meaning that more than 48% of the Mexican population has suffered from catastrophic payments to afford healthcare in the last decade (WHO, 2012).

On the other hand, the country presents one of the lowest rates of secondary school attendance among the OECD countries and compared to other developing economies. Mexico reports the lowest secondary school enrollment rates compared to its OECD counterparts, where only 56% of the population between 15 and 19 years participates in the formal schooling. Under this context, as inferred, only 33% of young Mexicans complete their studies, attaining upper secondary education (OECD, 2013; World Bank, 2013).

By using the Mexican Family Life Survey (MxFLS) for 2002 and 2005, this study aims at establishing a causal relationship between catastrophic spending on health and school interruption. In order to conduct the analysis, the MxFLS is used since it is a longitudinal survey containing information about income and expenditure, health status, healthcare and diverse school indicators at household and individual levels.

The methodological procedure is centered in the combination of two methods: a Propensity Score Matching to select control and treatment units, and the estimation of a Difference-in-Difference model to identify the average effect of health shocks on school interruption. The average treatment effect is estimated for the individuals aged between 16 and 18, since they are more prone to interrupt their education and to enter the labor market instead.

Under this context and aiming at providing details of every part of the research, the present paper is organized as follows. The next section comprises the literature review enclosing the theoretical background and recent findings between out-of-pocket expenditures and schooling decisions. This section supports the hypothesis and motivation of the study by revealing the

theoretical discussion and recent evidence among the factors and mechanisms in which schooling decisions are determined. Section 3 presents the description of the database utilized to carry out this research. Section 4 describes the methodology developed to conduct the analysis, i.e. the PSM and the DiD procedures, as well as the explanation of the variables incorporated to perform the model. The fifth segment provides the descriptive statistics of the sample and summarizes main results obtained. Within section 6, different specifications of the model are proposed to conduct the sensitivity analysis, reflecting the robustness of the findings and results. Finally, section 7 presents the methodological and data limitations and caveats present in this research, whereas section 8 offers the main conclusions derived from this study and some policy implications the results suggest.

2. Literature review: Theoretical background and recent findings

Most of the studies on the effect of health status of an individual on cognitive abilities and school enrollment (Miguel & Kremer, 2004; Kremer, 2003) have excluded the potential effects that health shocks experienced within the household produce on school participation of other member. Studies dedicated to estimate the impact of economic shocks, as crisis and structural changes, on an extensive variety of outcomes at the household level, as poverty, labor market, health, demographic transition, and education have been actively performed in the past two decades (Belley & Lochner, 2007; Carneiro & Heckman; 2002; Thomas, *et al.*, 2003).

The effect of family wealth on educational choices, for instance, is concentrated on the role of family income to determine school participation. Belley and Lochner (2007) find a significant increasing importance of household income on defining college attendance rates. By using the U.S. National Longitudinal Survey of Youth Cohorts for 1979 and 1997, they show that the change in the family income importantly affects the levels of college attainment and completion. Through the development of a model including both borrowing constraints and the "consumption" value of schooling, they also find that income has significant effects on the quality of the school attended. Namely, households and students are more likely to choose to get enrolled in a two-year school rather than a four-year school when they are borrowing and income constrained. After controlling for ability and family background as parental education, they conclude that family income is still an important determinant of schooling decisions. Therefore, considering these findings it can be expected that wealthier families be

less likely to present high interruption rates among their children, since initial endowments perform to be crucial when making schooling decisions.

In the same line, Carneiro and Heckman (2002) find that lower attainment rates are often obtained within members of poorer families; circumstance they ensure can be attributed to initial low-income levels. Additionally, educational outcomes as a response of income variation are wider in the presence of income inequality across generations and especially, when there are differences between races and ethnic groups. The authors explain that lower attainment rates are closely linked to the credit constraints families face in the adolescent years, affecting the availability of resources to finance education. Under this context, the authors point out that around 8% of American youth are adversely affected by short-term liquidity constraints, delaying their schooling.

An alternative interpretation of the schooling decisions highlight the long run factors derived from initially higher family income is offered within the study mentioned. Children living in wealthier families are more likely to have high income throughout their lives. This wealth at early stages of life is related to higher quality of education and better environments, promoting the development of intellectual and non-cognitive skills. This second explanation is reinforced when ability is included in the Instrumental Variable estimation to study the case of the United States (Carneiro and Heckman, 2002). The main finding is a significant estimate reflecting that the family gap in enrollment is explained through the diversity of abilities needed to benefit from participation in post-secondary schooling. Meaning that long run factors contained in this variable become the major determinants on the relationship between family income and secondary schooling.

More evidence provided by Thomas, *et al*. (2003) showing the close link between the level of income and school attainment is revealed. Throughout the analysis, the case of Indonesia is studied during the crisis experienced in 1998. The authors find support about the causal relationship between economic crisis and school participation, where the relation highly depends on the siblings' composition though. The results indicate that under economic crisis, the poorest families invest more in education for the oldest children within the family, reducing significantly the amount spent for the youngest ones, since they can enter more rapidly the labor market. Despite the increase in the resources devoted to the oldest child, the overall effect observed in the economy is the decline in expenditure on education and the

reduction of the school participation rates. Under this concern, to explore the effects that income and siblings composition have on schooling decision is important, since it may influence the probability of interrupting education.

On the other hand, Fallon and Lucas (2002) illustrate the effect of the Southeast Asia and Mexican crisis on schooling and health. The poorest households were the most affected when trying to smooth their consumption as a response of declining incomes. Overall, the fraction of family budget spent on healthcare decreased as consequence of the rise of price of health care services, resulting in greater morbidity rates. For the Thailand experience, the situation also implied the faster spread of some epidemic diseases The presence of the crisis reduced public expenditure on health as a proportion of total public spending and the situation brought the decline in the use of health facilities in both regions.

Through the use of panel data, Fallon and Lucas (2002) noticed that children living in poor households, together with the reduction in health status, also presented a decline in school attainment when they were affected by the financial crisis. For the case of Indonesia, they find a decline in the overall enrollment rates and significant increase of the dropout incidence, especially among the poorest youth aged between 13 and 19 years. In Mexico something similar occurred, during the debt crisis in 1982 lower completion rates were observed in secondary school enrollment, while a positive and significant impact arose at the primary school level. Meaning that during the crisis, the youth enrolled in secondary schooling were the more affected by the crisis, while the children between 6 and 12 years were somehow benefited by this economic instability. This unusual circumstance may be explained through a possible trade-off between labor and education, predominantly noticed in teenagers. If individuals can enter the labor market instead of studying to bear the burden produced by the income shock, therefore there will be a competition between education and work (Senne, 2014). Summarizing, not only the weakening of financial stability can be a consequence of the crisis, but also lessened school participation and the increased labor supply can be observed outcomes when financial distress is faced.

Krueger and Perri (2009) conducted a research to analyze the reaction of households as response of income shocks in the form of unemployment. They focused on the wealth loses that are likely to occur when the head of the household loses her/his job. Using the Italian Survey of Household Income and Wealth, *SHIW* (1987-2008) and the two waves of the Panel

Study of Income Dynamics in 2004 and 2006, they develop an OLS and Median Regression approach to estimate the effect of income shock on household's consumption and wealth. The diversity of variables measured included consumption expenditures on nondurable and durable goods, labor income after taxation and the asset position of financial and real wealth. Besides, controlling for differences in family size across households, Krueger and Perri (2009) find the causal link between income shocks and consumption decisions. Accordingly, the main findings are recapped as follows: (1) in the SHIW, the nondurable consumption response is stronger in the long run for those households that are affected by a income shock, i.e. in the next six years. Nevertheless, the wealth response is weaker in the presence of an income shock in the long term. (2) For households owning physical assets, as real estate or businesses, the durables and nondurables consumption response to income shocks is much smaller, whereas the wealth response is significantly larger.

The results insinuate that initially *richer* households tend to trade with their assets in order to absorb the shock in their income, while the *poorer* households rely on the reduction of expenditure and consumption to overcome financial shocks. Consequently, with the presence of physical and financial assets, households can make use of them in order to absorb the burden of the income shocks. Nevertheless, in the absence of these properties or funds, consumption bears all adjustment produced by these distresses (Ghez and Becker, 1975). Therefore, when households or individuals cannot go through short-term financial downturns, their ability to smooth consumption over time is dared. Families have to pool their resources in order to shift consumption from one good to another, and usually the poorest households face more restrictions to smooth their consumption over time.

A potential mechanism from which household resources, either time or income may be disturbed, is through unexpected transformations not only in wealth, but also in the health status of any of the members within the household. If astonishing changes in health produce the worsening of wellbeing together with the loss of income, thus the resources within households should be reallocated. As reviewed before, with the onset of the reduction of household's consumption, non-food expenditure may be affected as a result of income weakness. Within this categorical variable, the consumption of education is placed, being probably disturbed if economic hardship is confronted. Then, schooling decisions may be unfavorably affected when households agree diminishing investment in education. The situation is revealed in lower participation rates explained by the interruption of the formal

education (Dureya, Lam & Levison, 2003).

Thus, the harmful effects of income volatility could imply a trade-off between health spending or education consumption, being the probability to interrupt education increased especially for the youth. Namely, investing more in health and reducing the purchase of education becomes one of the mechanisms exploited by families to smooth consumption as a consequence of financial instability produced by healthcare spending (Caneiro & Heckman, 2002). In the short term, and in the long run if they persist, out-of-pocket health payments can derive into higher school interruption and lower completion rates, redirected into the overall reduction of education and productivity.

Finally, it is important to remark that most of this decline in school participation could have collateral effects as the increase in labor supply, especially for people aged between 14 and 18 years (Dureya, Lam & Levinson, 2003). Furthermore, the incidence of child labor also presents an expansion when there is a loss of wealth and assets given a financial shock within the household (Fallon & Lucas, 2002; Knaul, *et al.,* 2006). Accordingly, the interruption of schooling and the increased labor supply seem to be some recurrent behaviors performed by households to mitigate the economic hardship faced when unexpected external shocks occur. Overall, the increased labor supply among teenagers represents a disadvantage for the economy, particularly if they have interrupted their education, since this event reduces the human capital formation in the long-term (Baland & Robinson, 2000).

Little research has been developed to measure the effect of out-of-pocket expenditures on health care on educational decisions, and on school interruption in particular. For instance, Johnson and Reynolds (2011) studied the effect of household health shocks on the educational attainment of youth in the United States. By using data from the National Longitudinal Survey of Youth for 1997, they test whether health shocks have a negative effect on high school and college completion. Controlling for the size of the household, birth-order, youth health limitations, the AFQT score to proxy IQ and family income, the authors estimate a multinomial logistic regression to carry out their analysis.

The authors find that household health shocks lead to reductions in the probability of accomplishing high school on time. Moreover, youth are less likely to attend college and thus, completing a bachelor's degree. Additionally, the estimations show the existence of a quality reduction on education, reflected in switch-off from two-years schools to four-years

colleges. Besides, a reduced choice for colleges further away from home, being preferred the ones located closer in order to diminish housing and transportation expenditures accompanied the college trade-off. Contrasting Thomas, et al. (2003), Johnson and Reynolds (2011) find that the oldest sibling suffers from a higher reduction in educational attainment, while the youngest child is insulated from the health shock within the household. The conjecture is that given the oldest children present higher probabilities to enter the labor market, they are more likely to interrupt education; whereas the younger siblings, limited by age, are maintained enrolled in the education system.

As reviewed, income is one of the major determinants of schooling decisions (Fallon & Lucas, 2002; Thomas, *et al*., 2003), where the decision to interrupt schooling consequently, can be expressed as a response of the reduction in income and consumption (Caneiro & Heckman, 2002). Therefore it can be inferred that the financial instability produced by the health-distress of one of the members in the family may derive into the interruption of formal instruction of the youth living in the affected household.

Relaying on the existent evidence suggesting that the presence of unanticipated alterations in health and the catastrophic payments produced by them increase the probability of school interruption, the present paper aims at finding a potential relationship between the role of the family health shocks and educational choices. Hence, to test whether catastrophic health-related payments negatively affect schooling decisions, particularly school interruption, embodies the motivation of this research.

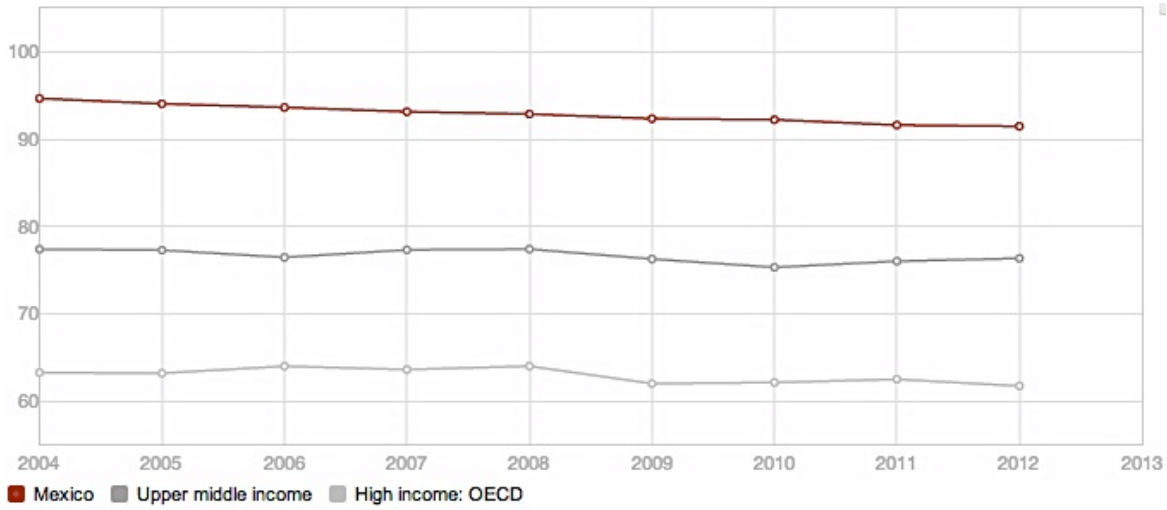
2.1 The case of Mexico

For the case of developing countries, the effect of financial catastrophe as a consequence of unexpected health shocks on schooling decisions has not been deeply reviewed. Mexico becomes an interesting case of study given that, in spite of the marked progress experienced in the last decade, the Mexican health system still presents serious coverage deficiencies. Besides, the country presents one of lowest investments in education among the middle-income economies, reflected in low enrollment, attainment and completion rates at secondary school level.

Mexico roughly invests 6% of GDP in health, representing the third lowest share among OECD countries after Turkey and Estonia. As reported by the WHO in 2012, only 47.3% of

the Mexican population is covered by some job-related public insurance, while the other 52.7% has to cover health care by private expenditures. Graphically illustrated in Figure 1, around 92% of these private health spending represents out-of-pocket expenditures (WHO, 2012). These conditions reflect that on average, more than 48% of the Mexican population has suffered from catastrophic payments to afford healthcare in the last decade. Thus, the universal access to healthcare, guaranteeing the insurance coverage for all the population, still a perceptible challenge for the Mexican government.

Figure 1. Out-of-pocket health expenditures as a share of private spending on health.



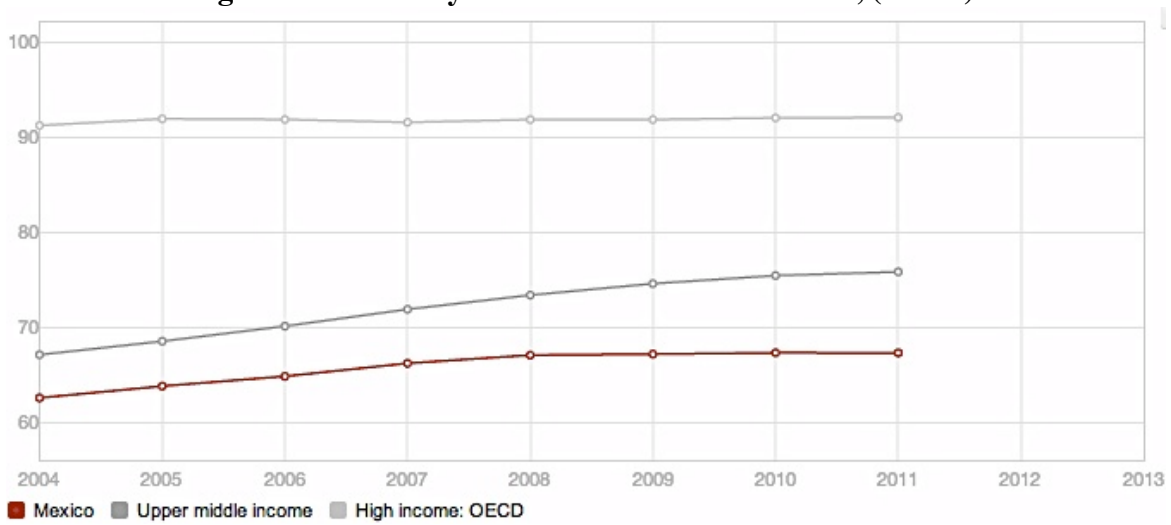
Data source: World Bank, 2014. The OOP expenditures include the cost of consults, drugs, therapeutic appliances, and other goods and services related to healthcare.

On the other hand, worried on educational choices, the country displays some of the lowest attainment rates in secondary schooling, having a negative repercussion on the graduation rates, which remain lagged at this level of education. Mexico presents one of the lowest secondary enrollment rates among the OECD countries and compared to other upper-middle economies as illustrated in Figure 2. The graph demonstrates that despite being compulsory, participation rates at this level of education continue low.

The country reports to have the lowest enrollment rates among youth with respect to the industrialized economies. Roughly 67% of the young population in 2012 was enrolled in the secondary school, while the OECD average is placed in 92% for the same year. Compared to the standards preformed by other Latin American countries such as Brazil (75%) and Chile (73%), Mexican educational system at secondary schooling remains lagged behind. Additionally, according to the OECD (2013), Mexico also presents the lowest secondary

education among adults compared to its OECD counterparts, being only 36% of them graduated from this educational level.

Figure 2. Secondary Enrollment Rates in Mexico, (% net).



Data source: World Bank, 2014. The graph represents the ratio of children in the official secondary school age actually enrolled with respect to the total population in the official secondary school age.

Despite of the increased participation rates at secondary level, the immediate consequence of the low enrollment rates is reflected in the diminished completion rates in the young population. According to the OECD (2013), roughly 33% of young Mexicans complete their studies, attaining upper secondary education. For the case of adults between 25 and 64 years, only 36% of them have obtained a grade equivalent to a high-school degree.

Many upper secondary schools are affiliated with large public universities, while others are state-controlled colleges or private schools. On average, in Mexico and among the OECD countries, the performance difference between public and private schools supports private schools (Hanusek, 2007; OECD, 2011). However, the cost of attending private schooling at upper secondary school level is placed around $150 and $350 USD a month (Ministry of Public Education, 2012). In addition to the monthly tuition, the existence of annual enrollment fees, the school uniform and school-materials expenditures notably increase the costs of private instruction. The expensive private option forces more than 82% of students to look for the public choice (World Bank, 2011). Nevertheless, even uncharged, the state-provided schools also imply costs in terms of school materials, transportation, uniforms and a diversity of fees (for instance school maintenance, festivals, among others) obliging teenagers

to reconsider whether to continue in the education system or to interrupt schooling if economic hardship is experienced.

As reviewed, a mechanism to confront the financial hardship and to bear the constrained income is to reduce school participation. The school interruption can be compensated with the rise in labor supply, especially for the youth that are legally allowed to work (Baland & Robinson, 2000). In these conditions, the prevalence of underage employment remains a major element, clearly perceptible in the Mexican labor market. The country presents one of the highest employment rates of underage compared to its OCDE partners, notwithstanding its *high* income per capita relative to other developing countries. The UNICEF (2007) describes that in Mexico around 12.5% of the population under 16 years is performing a job from which they directly contribute to the household income. Furthermore, 66% of the population aged between 5 and 17, i.e. 19,307,876 minors in the country, participate in some kind of housework. The numbers suggest that high rates of child and youth labor can be related to a certain competition with education, which could be directly produced for economic difficulties present within the household (Becker & Tomes, 1994).

As evidenced, Mexico presents one of the lowest levels of investment among its OCDE partners and among other developing countries in the crucial determinants of human capital formation: education and health. With the absence of full health coverage, a large share of households are obliged to carry out catastrophic expenditures to afford medical care when unexpected health-shocks are faced. This situation may derive into the threatening of household's welfare, forcing them to make harsh decisions on consumption and expenditure to alleviate the financial constraints they confront.

Under the context presented, Mexico is a stimulating country to study, since it becomes a potential case to provide evidence relating financial shocks due to out-of-pocket expenditures on health and their influence on schooling decisions made within households. Chiefly, it could allow analyzing the increased probability of interrupting education derived from these unpredictable health events. Moreover, the case of Mexico is interesting since this is the first study driven for the country and for any other Latin American economy with the purpose of measuring the effect of catastrophic health payments on school interruption. Therefore, the results would contribute to the general knowledge to support previous theoretical discussions

and empirical evidence not only for the Mexico, but also for other similar economies such as the Latin American ones.

3. Data

The empirical analysis in this paper is conducted through the use of the Mexican Family Life Survey, MxFLS. The MxFLS is the unique longitudinal survey in Mexico measuring a large variety of socioeconomic, demographic and health indicators for the Mexican population.

The Mexican Statistical Agency, *Instituto Nacional de Estadísitca y Geografia (INEGI)* designed the baseline survey and the scheme of it, modified to be adequate to the characteristic of the Mexican families, follows the Indonesian Family Life Survey (IFLS). Besides, other Mexican surveys, as the *ENIGH* (National Survey of Income and Household Spending) and *ENOE* (National Employment Survey) were also considered for the design to allow for comparability among them at the national level. The sample units were selected considering the national, rural-urban and regional representativeness, and according the population projections made by the National Population Council (*CONAPO)*. The regions were selected according the National Development Strategy 2002-2006, a structure proposed by the National Government.

The MxFLS follows a probabilistic, stratified, multiphasic and by conglomerates design. The survey is probabilistic given that all the sample units have a known probability and different from zero to be chosen. It is stratified because the sample units with similar geographical and socioeconomic characteristics are grouped into strata. Since the last households designated were selected though a process following several stages, the survey is considered as multiphasic. Finally, the nature of conglomeration of the MxFLS is enlightened by the different sample units from which is constituted (INEGI, 2004).

In 2002, the baseline survey (MxFLS-1) was conducted collecting information for 8,440 representative Mexican households, comprising 35,677 individuals of 150 communities across 16 states of the country and the United States. The second wave of the survey, MxFLS-2, was guided between 2004 and 2005, with a 91% of participation of those households interviewed in 2002 (Rubalcava & Teruel, 2006). This second survey contains information for around 40 thousand individuals, explained by the demographic dynamic within the country.

Individuals who died or emigrated to the United States that were not surveyed in the second wave explain the remaining 9% of the missing sample. A limitation of this last situation is that a health shock producing financial weakness within the household may influence migration decisions in order to look for another source of income. Namely, when a health distress episode is present, not only the interruption of studies but migration could be a potential consequence of the income volatility produced. Besides, the death of a household member could also suggest an income shock for families. In addition to this missing sample, 13.5% of the families within the survey have experienced temporary or permanent migration, whereas 740 families (8.76% of households) have suffered the death of a family member between 2003 and 2005.

Since the incidence of migration and the death of a relative may produce parallel effects on the socioeconomic status of households, it is hard to isolate the effect of health-shock on schooling decisions. The intuition is that migration can be caused by a positive selection of the members that decide to migrate; i.e. the individuals that are physically and/or intellectually stronger are the ones more likely to migrate. The emigration of the high-skilled workers, basically promoted by the large income gap between sending and host countries, could have two main implications for the sending economy: the brain drain and the generation of remittances (Haque & Kim, 1995). One concern arises with the remittances effect, given that they increase the household income, producing the enlargement in the level of consumption and investment families make. In general, this can imply a reduction in poverty and increase human capital in the population, if such supplementary income is spent on education of those who remain (Agrawal *et al.*, 2011; McKenzie & Sasin, 2007).

On the other hand, the death of an integrant could also affect household income or time of the household members because of shortened labor supply, changing their labor market behavior (Johnson & Reynolds, 2011). Evidence has shown that adult mortality produces a negative influence on children education outcomes, being newly orphans the less likely to attend school the year following the death by 20% (Beegle, *et al.,* 2006; Senne, 2014). The findings suggest that in a context with scarcity of resources and absence of full health-coverage, households suffering unexpected shocks produce children schooling adjustments as an instantaneous strategy to absorb the burden. Moreover, the misfortune has negative consequences on human capital accumulation in the long term (Senne, 2014). These households may be the more likely to make harsh decisions with respect to schooling since

they can be initially weaker than the ones that, even with catastrophic payments can afford health and continue investing in education. Therefore, the exclusion of these subsamples is important since they could affect the results by overestimating the effect of health shocks on school interruption.

Aware of these concerns and the potential consequences that the omission of these variables might have on biasing the results, the study only focuses on households without migrants before and after the shock occurs and families without an income shock produced from a dead member. With the settled restrictions, 13.2% of the observations were dropped from the analysis. Through the reduction of the sample size, it can be expected that the power of the results be negatively affected in the same manner. Nevertheless, the exclusion of those households in the estimation process would produce more accurate estimates, since the potential collateral effects of migration and death are avoided. In addition, despite of the reduction in the sample, the size of the survey remains large enough to conduct the analysis comprising 6,568 families.

Datasets at household and individual levels of both waves enclose information about income, expenditure and consumption choices, savings, family owned assets and wealth. Levels of education, school participation and completion indicators are also variables included in these surveys. Besides, measurement and interpretation of health status, as well as health shocks suffered at individual and household level are characteristics accounted in the MxFLS to be exploited in the methodology procedure.

The multidimensional nature of the MxFLS therefore, facilitates the study of the interrelationship of different demographic and socioeconomic phenomena present in the dynamics of the welfare of the population. An example of these links is the relationship between the health shock experienced within a household and the schooling decisions latter made as a consequence of the shock. The advantage of having panel data following the same households and individuals throughout two points in time is determinant to infer how families make educational resolutions for their children in the short term in response to an unexpected income variation.

4. Methodology

The present research conducts the empirical analysis through the use of a combination of two methods, with the purpose to find the causal link between catastrophic health spending and school interruption. The first one is a Propensity Score Marching method, that based on a set of households' characteristics, will allow identifying control and treatment groups in the absence of a randomized design. The PSM balances both groups selecting households with similar probabilities to be treated and thus, making them statistically comparable. Following the PSM, a Difference-in-Differences procedure is performed, which would help to test the hypothesis of this research, i.e. whether out-of-pocket health expenditures have a direct relationship with school interruption in Mexico. The DiD approach will measure the change in the treatment and control groups before and after the treatment (health-shock) occurs.

4.1 Propensity Score Matching

The major methodological concern relies on the situation that households affected by a health shock could be initially different from those families that do not to experience health weaknesses. Moreover, the differences between both types of families may be correlated with the subsequent schooling decisions due to out-of-pocket health spending. For instance, poorer households performing initially worse in terms of health and income, are potentially more prone to experience a health distress episode, jeopardizing their socioeconomic status. Likewise, households suffering an unpredicted health event that impairs the member's ability to continue working, for instance the head of the household, could influence not only schooling decisions, but also negatively impact the household wealth. Thus, the correlation between income shocks produced by a health distress can be confounded with the wealth effect, generating an identification problem.

In the absence of a controlled randomized-trial, the method of excellence for avoiding selection bias problems, a non-experimental procedure is then proposed in order to establish a causal relation between health-shock and school interruption. A Propensity Score Matching method, as in a randomized procedure, will allow identifying the treatment and control groups to be studied, and with them, the PSM determines the probability to be treated (Becker & Ichino, 2002).

The treatment group is the one containing the units receiving the intervention. Namely, for this research the group is conformed by all the households that made out-of-pocket health

expenditures between 2002 and 2005. The variable is composed by the inclusion of a hospitalization event, which in turn instigates financial catastrophe, representing at least the 30% of household budget on health spending. Around 25.6% of the surveyed population disburses more than 30% of their non-food income to afford healthcare.

Units presenting the same pre-treatment characteristics as the treatment group will compose the comparison or control group. However, the relying difference between both units is that the comparison group does not experience an income shock due to health spending. Meaning that, control units are composed by those families with the similar average income level, initial health status and other household characteristics as the treatment units, that in spite of having the same probabilities of experiencing the health shock as the treatment group, they do not suffer from catastrophic health spending. This last group will allow measuring the unobserved counterfactual. In other words, the control group will answer what would have happened to the treatment group in the absence of the treatment (Glennerster & Takavarasha, 2013). For this research, the counterfactual is going to determine what would have occurred to the schooling decisions, particularly school interruption, of those young living in a household in the presence of a health shock, had they not experienced the financial catastrophe episode.

In the sample, more than 60% of the households spend 6% of their budget on health, while the mean health spending of the total population is placed in 10.3%. Since the vast majority of the Mexican population presents an annual expenditure on health around 10%, households spending 10% or less of their non-food income are selected to be the units of comparison.

The use of a propensity score matching (PSM) allows selecting individuals and households with similar probabilities to be treated, conditioned on the same pre-treatment characteristics (X). This statement is expressed as follows:

$$p(\mathrm{X}) = \Pr(\mathrm{D}=1|X) = \mathrm{E}(D|X) \qquad (1)$$

where D represents a dummy variable taking the value of one when the group experienced catastrophic health payments due to a random health event, and zero otherwise. Therefore, the PSM method undertakes that given P(x), the outcome distribution of the treated households will be the same as the one observed for the control units. This matching procedure, allows

reducing the potential problem of selection bias generated by unobserved confounding factors (Angrist & Pischke, 2009; Becker & Ichino, 2002).

Since treatment and comparison groups are designated on the base of a balancing score according, with the use of a rich and good-quality control variables, the propensity score reduces the bias generated of the arbitrary selection of the sample. Thus, the selection bias is then reduced when it can be considered that the units exposed to the treatment and the comparison groups are randomly assigned, since after matching they present the same values for the propensity score, i.e. the likelihood to be treated (Dehejia & Wahba 1999). Therefore, the PSM can be perceived as a tool to create artificial treatment and control groups when they are not initially defined through a randomized control trial.

Following Galiani, Gertle and Schargrodsky (2003), treatment and control observations on a common support can be identified as follows. First, control observations with propensity scores lower than the ones observed for the treatment households at the first percentile of the treatment propensity score distribution, are excluded from the study. Likewise, the observations in the treatment group presenting propensity scores higher than the ones corresponding of the control observations at the 99% of the control distribution are not accounted in the analysis. This allows ensuring that both comparison and treatment groups present the same probability of being treated.

The propensity scores are calculated conducting a logit model, according to the subsequent notation:

$$\Pr(D_i = 1|X_i) = \Phi\,\{h(X_i)\}$$

where Φ expresses the logistic cumulative density function (c.d.f.) and $h(X_i)$ is the initial specification including all the covariates in a linear form. The equation indicates that propensity score based on the logit estimation is conditioned on the observed characteristics previous to the occurrence of the treatment (Angrist & Pischke, 2009). Moreover, these control units are independent of assignment into treatment. Thus, the logit approach allows estimating the probability of undergoing any financial-health related shock, conditional on a set of household covariates. On it, two individuals with the same X, one treated and one non-treated are expected to have the same counterfactuals.

Summarizing, two basic assumptions have to be fulfilled in order to obtain a good matching. First, the conditional independence assumption mentioning that all variables that are relevant for jointly determining treatment and outcomes are observed and included in X. Thus, it is assumed that the potential outcomes are independent of treatment assignment, conditional on a vector of covariates X. Second, the overlap assumption, which considers that all participants should have a counterpart on the participants. This ensures that individuals with the same characteristics have a positive probability of being treated and non-treated, exposing the common support assumption (Caliendo & Kopeining, 2008).

For the study purposes, it is important to point out that the students should not be the ones directly affected by any disease or accident. The assumption will help to avoid that possible health injuries, disturbing cognitive and/or physical abilities, force the student to interrupt education. Despite these considerations allow insulating the effect of financial catastrophe on schooling decisions, they could produce the reduction of statistical power with the diminution of the sample size, and be a source of selection bias. However, the restrictions imply the exclusion of 64 families, and since the number only represents the 0.75% of the sample, this constraint does not affect the estimations performed. I.e. the small proportions of the omitted sample cannot determine the sign or the magnitude of the estimates biasing the results (Galiani, Gertle and Schargrodsky, 2003).

On the other hand, people that lose their job due to a health problem are also removed from the sample. The explanation relies on the fact that people that are not able to perform a job because they are physically or mentally injured would negatively affect the level of income within the family, especially if this is the case of the head of the household. Moreover, this event may oblige the youth to start working in order to compensate the income lost. Hence, school decisions would be confounded not only by the impact on health, but also by other variables such as the inability to work produced by the severity of the health shock. Under this concern and the bias derived from it, 325 observations were removed from the sample in order to avoid the true effect of household hospitalizations to be underestimated. The exclusion of this subsample does not represent a problem since it allows insulating the effects of health shocks on school interruption, not being they confounded by other factors.

Some essential characteristics of each household have to be included in order to obtain balanced groups. Being interested in the health status pronounces the importance to observe

at one of the sources of the health weakness that may produce catastrophic health payments: the initial health conditions. The Body-Mass Index, measuring the initial health status individuals, has some advantages to show information about health limitations of the family members (Dureya, Lam & Levinson, 2003). On one hand, the BMI shows the presence of overweight or underweight depending on the person's weight, height and age, reflecting the general health position of an individual. For instance, overweight can derive in supplementary diseases as hypertension or diabetes, making people spend more than average on health. Thus, the BMI can be related with the likelihood of being treated, predominantly with the latent probability of being candidate to present a chronic illness in case this index surpasses 25kg/m2 (WHO, 2012). This positions the height-weight index as a relevant covariate to be included. On the other hand, the BMI could present a relationship with learning outcomes, since the case of underweight has been proven to cause negative influence on cognitive abilities, limiting someone's school participation (Kremer & Miguel, 2004). Thus, this index, measured for youth and adults within households, helps to control for intergenerational relations between parents and children that are not captured by chronic diseases or hospitalization events (Johnson & Reynolds, 2012).

Accounting for households characteristics and the probability of experiencing a health shock, it is essential to control for the number of inhabitants living in a household; in particular for the share of children and aged people living within it. A high dependency ratio, i.e. the proportion of underage and elderly people with respect to the adult population (between 15 and 64 years) inside the same household[1], may reduce the availability of resources and the ability of saving within households. Moreover, since children and elderly are more prone to suffer from some illness and thus, the families incur in higher risks to experience catastrophic health payments (Kremer, 2004), an elevated dependency ratio could position families on weak conditions when economic difficulties are present. In the Mexican survey, the average dependency ratio is equivalent to 0.587; viz. there are around 58% of dependents in the sampled households. The numbers suggest the existence of relative less people in the working age supporting the household expenditures with respect to the number of dependents. Therefore, to include the dependency ratio to balance treatment and control groups and make them to present the same probabilities of being treated is essential for the study.

[1] The consideration of de dependency ratio also measures the implication of having less number of income-winners with respect to the income-dependents in the same household.

On the other hand, it is important to include some variables defining the location of the household in order to perform a better matching. According to Eberhardt and Pamuk (2004), rural populations are placed in a more disadvantaged situation on many dimensions of health compared with urbanized populations. Health indicators of individuals living in rural regions, despite its improvement, are worse off than the ones inhabiting more developed areas. Infant mortality and lower life expectancy are more visible characteristics within rural populations (WHO, 2012).

Some risk factors, such as obesity and smoking, are more commonly found among rural residents, which in turn are linked to higher adult mortality rates and prevalence of chronic health conditions in rural regions (Eberhardt and Pamuk, 2004). Besides, most of the rural areas are not fully covered by health insurances, given the scarce of formal jobs and other forms of health coverage promoted by the government, explaining some of the worse health performance presented in these regions (Banerjee and Duflo, 2007). Thus, the inclusion of the region is relevant to determine the probability of experiencing a health shock.

Finally, covariates showing the wealth position of the families, as the possession of assets and the housing conditions, are imperative to be included in the design of the PSM. The ownership of valuable durable assets as the property of the house or real state, and other *low-cost* assets as domestic devices (microwaves, stoves, or washing machines) shows the economic position in which families are placed on and the material means they have to trade with in case of financial hardship. Moreover, the inclusion of these characteristics reflects the ability that poor households have to save, since the possession of any of these goods implies expensive transactions for which low-income families need to set aside some remnant income (Banerjee and Duflo, 2007; Krueger and Perri, 2009). Thus, the ownership of these assets helps to proxy the overall economic security of families.

These wealth characteristics are considered as dummy variables taking the value of one if the family posses the different devices or the house where they are living in, and zero otherwise. Therefore, with the purpose to show the families' wealth conditions, it is constructed a dummy variable indicating the housing conditions of the family; i.e. whether the family owns the house or not. Lastly, dummies reflecting the ownership of assets as domestic devices, for instance microwaves and washing machines, were included to compute the PSM.

Once having the two groups balanced, one can expect that the any source of variation between them be explained by the treatment effect and not by any other confounding factors. Namely, it would show the effect of unanticipated health shocks on schooling resolutions, particularly in school interruption.

4.2 Differences-in-Difference Estimation

Several unobserved characteristics of households might imply an identification challenge especially for those that vary across families but not over time. Time- and space-invariant unobserved characteristics correlated with the probability of experiencing a hospitalization episode and schooling decisions, could be present in the model, confounding and biasing the results.

The previous problem can be overcome by using panel data and through the estimation of a Difference-in-Difference (DiD) model. The DiD approach is one mechanism that helps controlling for time invariant heterogeneity in the sample, especially if the source of this last is unobserved. In other words it is a way to remove fixed factors that might influence both treatment and control groups, which in turns affect the outcome variable (Angrist & Pischke, 2009).

To know the true effect on the treated, the parallel or the common-trend support should hold, and differences between the treatment and control groups should be fixed over time. However, this last assumption cannot be tested, meaning that it cannot be empirically known that both treatment and non-treatment groups present the same trends in the pre-treatment phase (Wooldridge, 2010). Under this concern, the PSM followed by the DiD approach become one of the best strategies to balance groups and obtain similar units to be treated and compared as if they were randomly selected, i.e. assuming they present the same trends before the treatment provision (Lundborg, 2013). Therefore, the use of both methods will allow obtaining the average treatment effect on the treated, where the difference between trends after the treatment can only be attributed to the effect of catastrophic health spending.

Relaying on these assertions, the DiD estimation will be held based on the selected observations within the common support established by the PSM, assuming that:

$$E(Y_{0ist}|s,t) = \gamma_s + \lambda_t \qquad (2)$$

where *s* denotes a dummy variable representing the catastrophic health payments. Besides, *t* expresses the period, before and after the shock occurs, i.e. after 2002 and before 2005. The dependent variable therefore, indicates whether school interruption is a consequence of meeting out-of-pocket expenditures due to a health-shock or not.

Equation 2 shows that in the absence of a health shock, schooling decisions will be given by the sum of a time-invariant region effect (γ_s) and a time effect (λ_t), both common across treatment and comparison groups. Thus, the additive region effect will control for these unobserved characteristics at household and individual levels.

As long as the unobserved differences across regions are fixed over time, and the common trend assumption is fulfilled, the DiD-estimator will report consistent estimates (Angrist & Pischke, 2009). Accordingly, the equation form of a DiD estimation relies on[2]:

$$Y_{ist} = \gamma_s + \lambda_t + \beta D_{st} + \alpha X + \varepsilon_{ist} \qquad (3)$$

where D_{st} is a dummy variable indicating with one the presence of catastrophic payments on health, and zero otherwise. The variable therefore will take the value 1 if households spend more than 30% of their total monthly non-food budget on health-related issues or present an unexpected health shock (an hospitalization event), and zero if this value is equal to or below 10%. Besides, Y_{ist} denotes the outcome variable measuring the school interruption, which takes the value of one when the student between 16 and 18 years interrupts his/her education and zero otherwise. The selected age range is conducted considering one element: notwithstanding youth are prone to start or being enrolled in secondary education, with 16 years Mexican teenagers are legally allowed to enter the labor market. Therefore under economic hardship, this sector will be the more likely to interrupt its education given that a 16-year-old person is unconstrained to look for a job instead. In addition, school interruption is not an isolated event occurring in Mexico, where approximately 2,163 youth comprised in the 30% of the households in the survey, have interrupted their studies between 2002 and 2005. Therefore, to study whether catastrophic spending is a determinant of school interruption is relevant for the case of Mexico.

Finally, X represents a set of control variables capturing individual and household characteristics influencing educational choices. For the analysis purposes and based on the

[2] Notation based on Angrist and Pischke (2009).

economic theory and recent findings, X contains variables such as family income, health status of the household's members proxied by the incidence of chronic diseases, the type of school the youth is attending, parental background, the dependency ratio and the composition of siblings.

Household income incorporated as explanatory variable is explained by a major aim. One might expect that greater household income imply more family resources and higher prior investments in health care, which in turn reduces the probability of suffering a health shock (Dureya, Lam and Levinson, 2003). Despite the presence of an unanticipated distress in health, it is expected in the results that high initial family income reduce the probability of stopping education due to financial hardship produced by catastrophic health-payments.

The construction of this variable is done through the inclusion of labor and non-labor income of all family members present in the household, including the income perceived by children. Income received from social programs as *Oportunidades* or *Seguro Popular* is also included, since they additionally provide health coverage to poor people. In addition, an interaction term between income and catastrophic spending is included to capture the effect of income on school interruption given the occurrence of a health shock.

Continuing with the household characteristics, the importance of including the size of the family is vital since it might capture some unobserved factors behind the schooling decisions that larger families have to make. The change in family size over time, given the presence of newborns, deaths or the foundation of new families, may produce variations in spending and schooling pronouncements within households (Banerjee and Duflo, 2007; Black, Devereux and Salvanes, 2005).

Based on the quality-quantity theory, the number of family members could be a powerful determinant of children outcomes in the future (Becker and Tomes, 1976). The hypothesis stating the persistent trade-off between child quantity and quality and the general belief that greater family size produces negative outcomes on children's performance, is an important consideration when measuring the effect of financial hardship on school interruption. This quantity-quality theory suggests that the larger the family size, the more restricted the budget is to invest in children, leading to the increase in the marginal costs of quality with respect to family size (Becker, 1960; Becker and Tomes, 1976).

Contrasting this theory, Black, Devereux and Salvanes (2005) find that even if there is some evidence suggesting causality between the size of the families and the average children's education, there is not sufficient support for the existence of a causal effect between quantity and the "quality" of the children. Moreover, it could be the case that families who have a larger number of children are initially different from those having fewer children; and as consequence, their children tend to present lower schooling despite of the size of the household. For the case of Norway, the authors identify that family size and birth-order have a negative influence on the educational completion of children. Nevertheless when indicators for siblings' composition and twin births were includes as instrumental variables; the effects of household size become insignificant. Under these contradictory results, it becomes crucial to control for the size of the family in order to reduce the potential heterogeneity between groups and identify the effect of the household size on school interruption in Mexico.

It is important to observe that the presence of large or extended families living together could also have advantageous effects when the health of any of the members is threatened. Having more relatives in the family allows spreading the fixed costs over a larger number of people when financial catastrophe is experienced (Banerjee and Duflo, 2007). Therefore, even presenting higher probabilities of suffering health shocks, they can also spread the financial burden into a larger number of persons not being deeply forced to make harsh decisions.

When experiencing financial catastrophe, low-income households could face other decisions not only on what to spend the scarcity of resources, but also they have to decide to whom they should direct their restricted endowments. That is, households are forced to resolve how to reallocate the untaken reserves among their members. Distinctive assessments have to be made when the families have children aged between 16 and 18, i.e. while the minor is prone to start or to finish the secondary school level in Mexico. Moreover, when turning 16 years, they are legally allowed to enter the labor market.

According to Johnson and Reynolds (2011), the birth-order is determinant when financial hardship is faced within households, particularly true when the shocks are experienced in large and/or extended families. Therefore, one can expect to find a positive relationship between being the oldest sibling and the likelihood of discontinuing education. Inferring that the younger siblings are isolated in a substantial amount from the negative effects of a hospitalization event, whereas the oldest child generally assumes the burden of the

unexpected shock. For this reason, additionally to the family size, the inclusion of a dummy indicating whether the child is the oldest among his siblings is essential.

As it can be thought, families either small or large, presenting chronic illnesses are more likely to incur into catastrophic payments. Since more than 35% of the surveyed population is affected by a chronic disease, controlling for it through the inclusion of a dummy variable in the model becomes relevant. It can be inferred that a chronic illness increases the probability of a household to experience an income shock due to a hospitalization incident; even this last one can be more predictable. This situation could force family members to make intra-household resources allocation to meet all their needs in terms of health and education. The study only accounts for the relatives and not the student himself presenting any of the existent chronic diseases to avoid that cognitive or physical injuries affect the schooling decisions of the individual.

Being interested on educational outcomes, other covariate involved is the average years of schooling of parents. That is, the parental education measured as the total years of formal instruction reached at maximum level of education completed. Family background characteristics, providing a higher economic value to education, may induce individuals to stay in school longer (Altonji and Dunn, 1996). Hence, one can infer that higher parental education positively influence children's schooling assortments.

However, the potential correlation between parental education and income arises. Namely, wealthier families with higher educational levels tend to invest more in schooling (Banerjee and Duflo, 2007). In order to reduce the latent relationship between income and education for this case, it is included a variable determining the type of school the student is attending. According to Hanusehek (2002), if parents have larger interest in education and want to provide more motivation or a better learning environment for their children, they will tend simultaneously to pick private schooling. Students who attend a more selective college, i.e. the private option, may have higher initial endowments additionally to the higher interest of their parents to receive a *better quality* of education. Especially when these students are compared to the pupils attending private schools in developing countries (Lindahl and Regnér, 2005).

When exposing quality of education, private schools seem to perform better. Quality measured as smaller class' size, cultural and sportive activities, books and contents, is a key

characteristic often observed in the private education system. Substantially more expensive than the public choice, evidence has shown that private schools present higher educational quality, reflected in higher graduation rates in developing countries (Banerjee, et al., 2005; (Belley and Lochner, 2007). Thus, to determine whether the individual is attending to public or private education with the inclusion of a binary variable would help to capture some unobserved differences in educational choices and opportunities within households. Especially when school participation is sensible and subject to the costs of education and the disposable income to invest in it. Meaning that, this variable allows controlling for unobserved household characteristics that probably affect educational choices.

In addition to the above-described covariates, the error term of the model ε_{ist} composed by unobserved time and space factors, is included in the equation. These factors are expected not to be correlated with unobservable elements contained in this error term, expressed as: $E(\varepsilon_{ist}|s,t)$=0.

Assuming the previous statement as true, one can then consider that:

$$E(Y_{ist}|s = U, t = 2005) - E(Y_{ist}|s = U, t = 2002) = \lambda_{2005} - \lambda_{2004} + \beta$$
$$E(Y_{ist}|s = R, t = 2005) - E(Y_{ist}|s = R, t = 2002) = \lambda_{2005} - \lambda_{2004} \qquad (4)$$

Derived form the difference between expected value of Y_{ist} after and before being exposed to the treatment; the effect of experiencing a financial catastrophe or a health distress episode on school interruption, is given by:

$$E(Y_{ist}|s = R, t = 2005) - E(Y_{ist}|s = R, t = 2002)$$
$$- E(Y_{ist}|s = U, t = 2005) - E(Y_{ist}|s = U, t = 2002) = \beta \qquad (5)$$

where β in equation 5 is the causal effect of interest. Thus, the DiD estimator compares the change in the schooling decisions made by individuals before and after the catastrophic health payments occur to the change in the same decisions of teenagers that did not suffer from this burden.

5. Results

Table 1 presents the summary statistics before the balancing is carried out. Using sampling weights and expansion factors to state the national representativeness of the sample, an

average of 25.7% houses of the surveyed population have experienced out-of-pocket expenditures on health. That is, around 2,959 of the sampled Mexican families spent more than 30% of their budget on healthcare between 2002 and 2005. These private payments are significantly higher compared to the ones observed for the vast majority of families in the remaining sample, which present a mean expenditure on health by less than 10.5% annually.

Table 1. Summary statistics for the complete sample at household level, Mexico 2002 and 2005.

Variable	Observations	Mean	Std. Dev.
BMI Youth	6,568	23.3685	7.3171
BMI Adults	6,568	27.2894	5.7470
Chronic Disease	6,568	0.3501	0.1430
Hospitalization Event	6,568	0.2173	0.1430
School Interruption	6,568	0.3083	0.2459
Parental Education	6,568	9.5545	4.3063
Household Income	6,568	56,106.61	62,396.53
Household Size	6,568	4.7688	2.0631
Number of Siblings	6,568	2.6436	1.7501
Dependency Ratio	6,568	0.5871	0.1892
Rural Area	6,568	0.0993	0.2991
House Ownership	6,568	0.8587	0.3483
Durable Assets	6,568	0.8807	0.3242
Electronic Devices	6,568	0.8659	0.3407

Note: Summary statistics are performed with sampling weights for the full sample. BMI for youth is considering population aged between 16 and 18. Durable assets dummy indicates the possession washing machines or stoves. Electronic devices include microwaves and T.V. Income is reported in Mexican Pesos (MXN).

Accounting for the initial health position of the individuals in the full sample, the BMI is placed at 23.36 kg/m2 for the youth, while adults present a height-weight index of 27.28. The index reflect that the young Mexicans in the sample present normal levels of weight according their age and height; whereas the adults present overweight, since the BMI is above the normal weight standards. Meaning that, on average and according to the World Health Organization (2006) classification, the adults present overweight and being in the edge to be obese class 1, with the increasing risk to develop a chronic disease. In this context, approximately 35% of the total sample presents at least one chronic disease, while 21% of the sample has experienced an unpredicted hospitalization event. The presence of a member with a long-term disease within the family may threaten the financial stability of the household due to constant health spending or an income shock derived from an unexpected hospitalization.

Measuring educational outcomes, roughly 31% individuals interrupt or abandon their studies when they are aged between 16 and 18. The fact reflects high interruption rates at secondary level, being consistent with low graduation rates among youth displayed by the OECD. Accordingly and related to educational outcomes, parental education remain low in this sample of the Mexican population. An average of 9.6 years of schooling represents the maximum level that parents achieve. This number of formal years of schooling reflects that on average, parents are roughly completing lower-secondary education, while upper-secondary schooling even started is not completed in most of the cases.

The absence of higher levels of education can be reflected in low productivity and low income. The average income of the included Mexican families oscillates around 56,100 Mexican Pesos (6600 USD in 2002), being notably less than the achieved for the North American families in 2002, equivalent to 36,764 USD (OECD, 2012). The initial low-income levels are likely to produce catastrophic payments when a health shock occurs, particularly when families are uncovered from any health insurance. Thus, in the propensity matching it is essential to control for income in order to homogenize the families selected to be part of treated or non-treated units.

Accounting for the characteristics of the households, the average family size in the sample comprises 4.76 members, with a median of five persons in each family. Besides, the dependency ratio is found at 58.71% indicating the large proportion of children and aged people within the families with respect to the breadwinners. Large household size either explained for the quantity of children or for the presence of extended families, is more likely to be found in rural areas, representing these regions approximately the 10% of the sample[3].

It is important to point out that these averages are below the national levels, which according to the World Bank, 24.64% of the population was living in rural areas in 2002. This fact remarks the absence of complete representativeness of the survey stated before. The smaller proportion of residents living in rural counties in the sample may diminish the probability of the individuals presenting catastrophic health-payments. This situation could lead to the underestimation of the effect of financial shocks on educational choices. However, with the use of the Propensity Score Matching it can be ensured that both groups present the same

[3] According to the classification of the National Institute of Statistics and Geography in Mexico, rural areas those with 2500 inhabitants or less.

characteristics, and thus, the same probabilities of being treated regardless of the place where people live. This will allow reducing the selection bias created though the initial design of the survey.

Housing position and the possession of durable assets, a measure of wealth, are crucial covariates to be included into the study. In the sample, 85.9% of the families own a house, reflecting that despite the income level, families are able to buy or to construct a house by their own. Nevertheless, it is important to remark that most of these houses are low quality structures, with shortages of water and electricity. According to Rivera, *et al*. (2004), around 47% of houses require improvement actions, are of very poor quality and needs to be replaced.

Likewise, when analyzing the low-cost durable assets, more than 87% families own electronic devices such as microwaves and possess washing machines. Thus, the vast majority of the surveyed Mexican households are able to save some income in order to invest it or pool from this residual when financial hardship is present.

5.1 Matching Groups

In expectations, both control and treatment groups should show similar outcomes if the treated group had not experienced a health-shock. Therefore, it should be expected that the means of the covariates to be the same for both groups in the pre-intervention period, which in turn suggests a high probability that outcomes in both units are similar in the absence of the shock. If some covariates are very different between groups, they are ignored for the estimation of the PSM for this to be valid. Testing the means of covariates, for both treatment and comparison groups, before the hospitalization event occurs, would help to estimate this unobserved counterfactual.

Accordingly, Table 2 encloses the tested means provided for the propensity score matching for the treated and comparison groups. Panel A in Table 2 reports the means before and after matching to validate the hypothesis that no statistical significance between groups is present after performing the matching procedure. Panel B report the means and the t-test for them for other covariates, that despite not being included in the construction of the PSM, it is important to assess their balance to pursue the following DiD estimation.

Covariates as dependency ratio, house ownership, the possession of low-value durable assets, electro domestic devises, initial health status and being located in a rural area are variables involved in the construction of the PSM. The inclusion of these observed characteristics will determine then, the symmetry of both groups.

Table 2. Control and treatment groups: Mean differences before and after matching

Variable		Mean		t	P
	Panel A				
		Treated	**Control**		
Dependency-ratio	Before	0.58	0.53	2.01	0.045**
	After	0.58	0.57	0.47	0.6410
Initial Health Status: Youth	Before	23.63	23.16	3.35	0.001***
	After	23.63	23.57	1.08	0.278
Initial Health Status: Adults	Before	31.35	31.16	1.13	0.2570
	After	31.35	31.23	0.61	0.5390
Number of Siblings	Before	4.74	4.90	2.05	0.040**
	After	4.74	4.90	1.54	0.1230
Rural	Before	0.10	0.10	0.75	0.454
	After	0.11	0.10	0.54	0.587
House Ownership	Before	0.83	0.81	1.91	0.056*
	After	0.83	0.83	0.01	0.992
Durable Assets	Before	0.89	0.89	-0.10	0.923
	After	0.89	0.89	-0.46	0.646
Electronic Devices	Before	0.88	0.87	1.55	0.120
	After	0.88	0.88	0.84	0.403
	Panel B				
School Interruption	Before	0.39	0.40	-0.51	0.607
	After	0.39	0.40	-0.35	0.723
Parental Education	Before	8.82	9.27	-4.08	0.000***
	After	8.82	8.90	-0.56	0.578
HH Income	Before	56,922	62,351	-1.43	0.154
	After	56,922	58,532	-1.45	0.1480
Observations		**2,163**	**4,327**		

Notes: Matching method uses logistic regression with sampling weights. Durable assets indicate de possession washing machines or stoves. Electronic devices include microwaves and T.V. Income is reported in Mexican Pesos (MXN). Mean difference: *** statistical significance at 1% level; ** statistical significance at 5% level and * at 10% level.

In the first two columns of Table 2, the average dependency ratio is reported for treated and control groups respectively. The non-treated group presents on average smaller number of dependents than the treated group. Around 53% of dependents with respect to breadwinners are found in the comparison households, whereas the average dependency ratio for the treatment group is equivalent to 58% before matching. The present difference between groups

is significant at 5% level, with a t-statistic of 2.01. Thus, the null hypothesis of no significant difference between groups is rejected. The reason behind the disparity can be attributed to the close relation between dependency ratio and the probability of being affected by a health shock stated in the theory. Thus, it has to be ensured that at least both groups in the selected sample present the same average number of dependents and consequently, similar probabilities of being treated.

After performing the PSM, the new control units selected reflect an increased dependency ratio being balanced with the treatment sample, showing an average of .57. It is important to notice that both groups present a dependency ratio consistent with the one observed for the full sample. With the matching procedure, the slight difference between groups is not statistically significant at conventional levels, being then comparable and having the same probabilities of being assigned into the treatment.

Using the Body Mass Index (BMI) for adults and youth, the initial health status is measured. In the baseline survey, the mean BMI for the youth in treatment and control groups oscillates around 23 kg/m2. This situation reflects that overall, people in both groups are healthy consistent with the complete sample. The difference in both groups is statistically significant at 1% level, though. The difference could be accredited to the fact that the selected treated units before matching account for some individuals with overweight and obesity present in the full sample.

On the other hand, adults present a BMI equivalent to more than 31 kg/m2; thus, they are placed in the obesity level. The treated group presents a higher "height-weight" index than the non-treated group before performing the PSM, being this difference equal to 0.1890 points, but it is not significant at any confidence level. With the disparities observed between groups, the PSM becomes essential to balance both units to present similar means and trends, which in turn ensure that all the individuals have the same probability of being treated. Moreover, the PSM homogenizes both groups to make them comparable and representative of the complete sample. Once the matching is executed, selected agents of comparison report a BMI equivalent to 23.57 units, being placed in a "healthy weight" level. With the use of the PSM, the difference in means between groups reduces up to 0.06 points for youth and it is negligible at 10% level. For the case of adults, after matching, the group increases its mean up to 31.23 kg/m2 and the null hypothesis is still rejected at conventional levels.

Other included covariate to perform the PSM is a dummy variable determining whether a family lives in a rural area or not. Table 2 illustrates that approximately 10% of the population, i.e. 2,517 families live in a rural area. For this case, a larger proportion of the treated group belongs to a rural community, 10.5%, while 10% of the comparison group is placed in this group. The t-statistic for the 0.5% difference between groups is equivalent to 0.54, turning this slight difference in means is not statistically significant at conventional levels, neither before nor after matching.

Nonetheless, even this sample remains below the 20% of rural population estimated by the World Bank; the reported share is consistent with the national sample included in the survey. Additionally, since the null hypothesis cannot be rejected for the difference between treated and non-treated groups, together with the inclusion of a random event in the model, such as the hospitalization episode, the DiD approach later performed may address some of the bias caused by this relative small rural sample.

Finally, the inclusion of covariates indicating the socioeconomic status of the family and the wealth position are shown in Table 2. Within the sample for treated and comparison groups, more than 81% of the population owns a house. The reported means are higher for the treated groups, suggesting that on average they possess more assets to commercialize if financial catastrophe is experienced. Thus, in general terms, treated units are better-off to face unexpected financial shocks than the comparison agents, being this pre-matching difference statistically significant at 10%. After performing the matching procedure, the difference in groups is diminished, going from 1.5% to 0.01%. This indicates that the non-treated group is balanced with the treatment group in the post-matching phase. The slight variance between groups cannot reject the null at 10% level.

For other durables included, as washing machine/stoves, and additional electro devices such as microwaves, the large majority of the households own them. More than 88% of the families in the sample possess at least one of these devices, being the non-treated group roughly better off than the treatment group in general terms. The slight differences between comparison and treatment households are not significant, signaling that groups are well balanced in terms of wealth and the material means they can pool-of if a health and financial shock is faced.

Generally, the wealth position presented for full sample and the subsamples selected to be control and treatment units, confirm that Mexican households have a strong ability to save. Namely, they have the faculty to keep some of their available income to afford some expensive durable assets. Moreover, the households' ability to save can be determinant to purchase health services, specifically when these savings prevent from financial catastrophe.

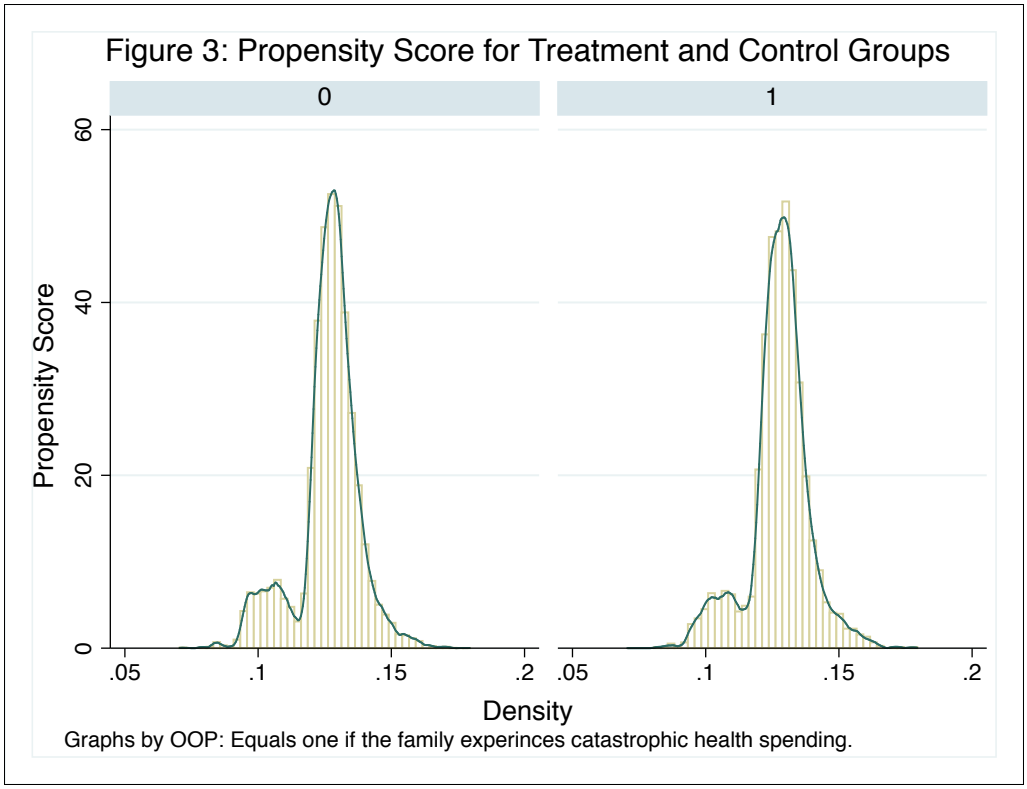
Under the assumption of common support, where a treated unit has a perfect match with a control unit, 28,884 observations were deleted since they could be placed neither in the control, nor in the treatment group. Meaning that, with pronounced disparities between groups and missing information for some individuals (age, education, expenditure on health and income) the absence of a point in common with the potential control entities in the dataset makes not feasible to find a partner for each treated unit in the sample. In additions, most of the omitted observations are part from the households not incurring in catastrophic spending to afford health. Nonetheless, the exclusion of these observations can cause a selection bias problem and reduce the power of the results. Moreover, it could limit the generalization of the results obtained. However, in order to account for the excluded individuals, the robustness check for the model proposes different specifications with the inclusion of these observations in order to measure the magnitude of the potential selection bias.

The established constraints produce a final sample with 2,163 treatment households and 4,327 families belonging to the control group. The larger (and preferable) comparison group ensures that each of treatment agent have at least one control unit with similar characteristics to be compared with.

Since the common-trend assumption cannot be tested, a t-test is performed to corroborate the overall balancing of the sample selected on the common support basis. The t-test on the hypothesis establishes that the mean value of each variable is the same in the treatment group and in the control units. If $p>0.1$, the null hypothesis cannot be rejected at the 10% significance level. For all the variables, with a p-value of 0.129 the null hypothesis cannot be rejected, i.e. the mean variables of the two groups do not differ after matching. Furthermore, the bias of the sample reduced from 3.1 in the raw survey to 1.8% in the final matched sample. Meaning that both groups are well balanced after proceeding with the PSM and the bias was significantly reduced. As a way of visually testing the symmetry between groups, in

addition to the t-statistic, a histogram for the propensity score when the treatment variable equals one and zero is showed in Figure 3.

The figure illustrates that with the nearest neighbor approach for each treated unit (random draw version), both the treatment and comparison individuals present the same probability of being treated. Therefore, after matching it can be observed that both groups are balanced, and there are not simultaneous differences among groups that may affect the diff-in-diff estimation.



Finally, it can be assumed that through the use of the propensity-score nearest-neighbor matching, it was possible to artificially create a control group similar enough to the treatment group. Similar groups in expectations can be used then to estimate the average treatment effect on the treated through the Diff-in-diff procedure.

5.2 The effect of health-shocks on schooling decisions in Mexico

Table 3 reports the results derived from the Difference in Differences estimation measuring the effect of health shocks on school interruption. The outcome is based on the observations obtained with the principle of common support. The estimates for treated and control groups, in addition to the differences between both groups are reported for the base line (2002) and

for the follow up (2005) observations.

For the case of the base line estimates, the difference between control and treatment groups is negative. Thus, during the pre-treatment phase, there was a negative trend on the school interruption behavior, where the treatment group appears to have a slight advantage reflecting less probabilities of interrupting education than the comparison units. However, the difference between both groups is not statistically significant at conventional levels.

For the follow up estimate, the difference between groups is equivalent to 0.168 in 2005, and statistically different from zero at 1% level. The treatment group presents an exponential change in its coefficient, revealing that after presenting catastrophic health spending, the student is more likely to interrupt his instruction.

Finally, the DIFF-IN-DIFF column shows the parameter of interest, i.e. the average effect of health shocks on schooling decision, specifically school interruption. The estimate is the result of the difference between the control and the treatment groups before and after the health-shock occurs. The parameter is positive and strongly significant at 1%, meaning that the financial instability produced for unexpected health shock does increase the probability of interrupting formal education by 17.8% in Mexico.

Table 3. Difference in Differences Estimation: The effect of Health-Shocks on School Interruption in Mexico.

			Base line			Follow up	
Outcome Variable	**Control**	**Treated**	**Difference**	**Control**	**Treated**	**Difference**	**DIFF-IN-DIFF**
School interruption	0.0270	0.0170	-0.0100	0.0230	0.1910	0.1680	**0.1780**
Std. Error	0.0050	0.0100	0.0100	0.0050	0.0500	0.0500	0.0510
P- value	0.0000	0.0860	0.3090	0.0000	0.0000	0.001***	0.000***

Note: The Base line shows the estimates obtained for the treatment and control groups in 2002, whereas the Follow up estimates are the corresponding results for 2005. DIFF-IN-DIFF is the parameter of interest. Standard Errors are reported within brackets. Significance *** p<0.01; ** p<0.05; * p<0.1.

With the percentage change in the control group equivalent to -14.8% between 2002 and 2005, the control group showing the counterfactual reflects that not being affected by catastrophic health payments reduces the likelihood of interrupting education, confirming the negative trend observed before the shock. Thus, if the treatment group had not disbursed more than 30% of its income to fund medical care, treated youth would have been less likely to interrupt their education.

However, the results obtained in this first equation can be biased due to endogeneity problems. That is, the omission of variables correlated with schooling decisions would

probably create an upward bias in the estimation. Since some household characteristics are excluded from the model and might be contained in the error term, the true effect of the catastrophic health spending on schooling decisions is overestimated. Moreover, if time invariant health characteristics or cognitive abilities of the individuals that affect the outcome are present, a potential fixed-effect is biasing the results. Thus, the existence of characteristics constant through time may be an important source of time invariant heterogeneity correlated with the explanatory variables.

The use of panel data and the inclusion of a larger set of covariates allow adjusting for the potential endogeneity concerns. The second and third columns in Table 4, therefore reports the results obtained considering the previous understandings. In the first column the coefficient obtained through the simplest version of the DiD estimation is included only for comparison purposes.

Table 4. The effect of Health-Shocks on School Interruption

Variables	[1]		[2]		[3]	
OOP Health Payments	0.178	***	0.1370	***	0.121	**
	(0.051)		(0.014)		(0.059)	
log (HH Income)			-0.126	***	-0.089	***
			(0.003)		(0.017)	
Oldest Sibling			0.025	***	0.026	***
			(0.005)		(0.007)	
HH Size			0.087	***	0.079	***
			(0.001)		(0.004)	
Chronic Disease			-0.052	***	-0.055	***
			(0.006)		(0.011)	
Parental Education			-0.016	***	-0.018	***
			(0.001)		(0.002)	
Rural Area			0.024	***	0.022	*
			(0.008)		(0.014)	
OOP*log(income)					-0.130	***
					(0.002)	
OOP*HH Size					0.011	
					(0.010)	
Public Education					0.004	
					(0.008)	
Observations	6,474		4,931		4,931	

Note: Model 1 represents the estimation for the parameter of interest without controls. Model 2 reports the estimates controlling for other variables that have a potential effect of schooling decisions. Model 3 includes interaction terms between OOP and income, and OOP and HH size. Sample size is smaller for Models 2 and 3 since some information is not available for the whole sample. Standard Errors comprised within brackets. Significance levels *** p<0.01; ** p<0.05; * p<0.1.

Model 2, reported in the second column within Table 4, controls for the household income in logarithms, the family size, the incidence of a chronic disease in the affected household, and for the average years of parental education. It also includes a dummy variable exposing whether the potential deserter is the oldest sibling in the family, as well as one binary determining if the household is positioned in a rural area. Model 3 in Table 4 controls for the same covariates as the previous model, with the difference that it also includes the type of school the youth is attending. Additionally, it comprises an interaction term between out-of-pocket health payments and the log(income); and catastrophic health payments and household size.

The coefficient in the first raw, the effect of health-shock in school interruption reports a positive sign and a statistically significant estimate. Meaning that, there is a direct relationship between a financial catastrophe caused by a hospitalization event, and the decision to interrupt education. The probability of interrupt education is equal to 13.7% in model 2, while its counterpart in model 3 is equivalent to 12.1%. The decrease in the magnitude of the estimates confirms that the existence of time invariant characteristic does affect schooling decisions. Besides, the fixed components comprised within the error term were biasing upward the results reported in the first column of Table 4. Moreover, the result suggests that schooling decisions are not only driven by health-income shocks.

Despite of the reduction of the magnitude of the parameter, the inclusion of more variables in the estimation preserved the statistical significance of the estimate in model 2. Namely, the reduction of the bias produced by the presence of omitted variables, increased the power of the estimation. For the case of the third estimation, the coefficient is statistically significant at 5% with larger variance than the previous specification. The inclusion of interaction terms and the type of school in the third equation reduced the significance of the parameter of interest; i.e. the health shock. The reduction of the magnitude suggests a possible correlation between schooling decisions and public education, where being enrolled in the public education system reduces the likelihood of interrupt instruction. Moreover, studying in a public entity could be a consequence of a health shock in the past. Besides, lower income or the family size can be potentially correlated with the amount spent on health.

Nonetheless, considering the health shock as a random and unexpected and accounting only for households spending more than 30% of their budget on health as a consequence of these

unexpected shocks, might reduce the mentioned source of endogeneity. Since health shocks are unexpected events threatening the financial stability of a family, it can be assumed that the effects of them on school interruption are isolated from other confounding factors. Moreover, the effect is not correlated with potential omitted variables. Besides, the interaction terms and the type of school are not affecting the direction in which catastrophic payments affects the probability of interrupting education.

The role of the household income is important in the decision-making process, especially when it is related to schooling resolutions. Using the observations on common support, the corresponding estimator for the income available for each household reports a negative and statistically significant result in equations two and three in Table 4. Meaning that, the higher the family income, the lower the probability of delaying education. The second column illustrates that the higher the initial income, there exist 12.6% lower probabilities of interrupting education; while roughly -9% is the case for the third column. The findings insinuate that when the family is wealthier in monetary terms, they are less likely to force their children to reduce their participation in school. Moreover, the interaction term in model 3 shows that higher household income is determinant for make schooling decisions when out-of-pocket payments are faced. For the Mexican case, youth belonging to the affected families have less probabilities of interrupting their education when presenting health and financial catastrophes.

However, households with low initial levels of income can be forced to make decisions not only in what to spend, but also on whom to spend the available income. The fourth row in Table 4 reports the estimate for the household size, being positive and statistically significant at 1% level. As expected, the larger number inhabitants within the household produces adverse variations in schooling decisions. Meaning that, the family size closely related to the dependency ratio, does have a negative influence on schooling decisions, increasing the probability of interrupting education.

To measure the effect of living in larger families when experiencing out-of-pocket health expenditure on school interruption, the third column comprises an interaction term between both variables (OOP*HH Size). Regardless the expected positive sign of the coefficient, the interaction term does not produce statistically significant results at conventional levels.

Therefore, it cannot be concluded that being part of a large family when the health shock occurs positively influences the probability of interrupting education.

A dummy indicating the prospect to interrupt education if the pupil is the oldest child is included in models 2 and 3. The estimations expose that the probability of interrupting education increases around 2.5% when the child is the primogenital among his/her siblings. Under this context, the result confirms for the case of Mexico that families pool their resources through the lessening of spending on education in the oldest children. The situation can be related with the age of the teenager, since older children (from 16 to 18 years) can chose to enter the labor market and interrupt their instruction, while their younger siblings limited by age are more likely to be maintained enrolled in school.

The variable showing the health status of the family members is incorporated in models two and three in Table 4. A dummy indicating the incidence of chronic illnesses in the household but not in the student is included as control. The result illustrates in both equations that presenting a chronic disease have an inverse relation with school interruption in more than 5%. This negative correlation insinuates that the incidence of a chronic illness reduces the probability of the student to stop his education. Even counterintuitive, this may suggest that having a chronic illness produces recurrent health spending, but not out-of-pocket payments that adversely affects schooling decisions. Thus, the majority of people suffering them are not incurring in catastrophic health payments to afford medical care. The coefficient is statistically different from zero at 1% levels in both models.

Parental education, as the theory suggests, is another determinant of schooling performance and educational decisions. The estimate obtained with the DiD approximation, reflects an inverse relationship between parental background and the decision of interrupt education. Despite the small magnitude of the estimates (-1.6% and -1.8% for model 2 and 3 respectively), the result is statistically different from zero at the 0.01 level. That is, more educated parents reduces the probability of interrupting education. Meaning that, parents with higher educational background will be more interested in maintaining their children enrolled in the educational system, rather than oblige them to interrupt their studies. Nevertheless, the small coefficient could reflect that for Mexican families, parental education only influences the schooling decisions of their children in a minimum amount, not being economically

significant for this case. Therefore, parental background is not the major determinant of educational resolutions in Mexico.

As mentioned, it is essential to include characteristics of the type of school the youth is attending to, since this controls for the costs of schooling. As inferred, the estimate obtained in equation 3 in Table 4 is positive, showing that being enrolled in a public school increases the probability of school interruption in 0.4%. Nevertheless, the magnitude of the estimate is minuscule and it is not statistically significant at conventional levels. Namely, participating in public education does not influence schooling resolutions in any direction. Moreover, the results suggests that paying higher quotas to access to the education system does not guarantee a good quality of schooling either, much less ensures school completion. Therefore, the underlying differences in quality for private schools respect to the public ones, still being an ambiguous benefit to be drawn for any policy purpose for the case of Mexico.

6. Robustness check

The Kernel density-weighting procedure is another mechanism to obtain the generalized DiD matching estimator and find the average treatment effect on the treated. The Kernel matching utilizes weighted averages of all individuals comprised in the comparison group to construct the counterfactual outcome. The sample used in this approach accounts for the individuals that did not present a perfect match during the PSM procedure, but comprises similar characteristics to the treatment units. One benefit of the method is that the inclusion of individuals found in the borderline may produce better results, able to be generalized for a larger sample, and thus have stronger support for policy implications.

The Kernel method is an approximation to the "Randomization at the bubble" approach, where participants that are not completely eligible are randomized to be part of the treatment or control units (Glennerster & Takavarasha, 2013). This method allows determining the intention to treat effect and it is especially helpful if one program plans to be extended to a larger proportion of the population than the one that was initially targeted. Since more information is included to perform the DiD estimation, the variance reduces significantly becoming the main advantage of this type of matching (DiNardo and Tobias, 2001). Consequently, the Kernel density-weighting method also allows addressing the potential serial correlation created in DiD models.

The first column in Table 5 indicates the estimate obtained through the Kernel Density Matching method. On it, only the coefficient of interest derived from the DiD estimation, i.e. the effect of health shock on schooling decisions in 2005 is reported. The health shock, as predicted, brings a positive and strongly significant estimate, representing the direct effect it produces on finishing education on time.

The estimate obtained with the Kernel-matching increases in magnitude with respect to the previous DiD estimations, but is smaller than the one in the baseline equation. After an episode of health weakness causing economic instability within a family, the student is more likely to postpone or abandon his/her education as a response to this situation by 13.8%. The standard errors are equivalent to 0.010 being the result strongly statistically significant different form zero at 1%.

Table 5. Sensitivity Analysis:
Kernel Matching and Health Spending's Thresholds

	[1]		[2]		[3]	
	Kernel		**11%-19%**		**20%-29%**	
OOP Health Payments	0.138	***	0.008		0.009	
	(0.010)		(0.039)		(0.086)	
log (HH Income)			-0.131	***	-0.121	***
			(0.009)		(0.002)	
HH Size			0.088	***	0.068	***
			(0.004)		(0.001)	
OOP*log(Income)			-0.141	**	-0.194	***
			(0.030)		(0.008)	
Oldest Sibling			0.018	***	0.067	***
			(0.008)		(0.004)	
Chronic Disease			-0.049	***	-0.012	***
			(0.011)		(0.004)	
Parental Education			-0.016	***	-0.004	***
			(0.002)		(0.000)	
Rural Area			0.017		0.043	***
			(0.014)		(0.003)	
Observations	5,122		7,896		5,946	

Note: Model 1 represents the estimation for the parameter of interest with a Kernel Density Approach. Model 2 reports the estimates for the effect of households spending between 11% and 19% of their non-food income on health. Model 3 shows the estimates considering the families spending between 20% and 29% of their non-food income on health. Standard Errors comprised within brackets. Significance levels *** p<0.01; ** p<0.05; * p<0.1.

Second and third columns in Table 5 comprises the families initially excluded from study; i.e. the household spending between 11% and 29% of their available income on healthcare. The importance of these specifications is to determine whether the excluded families and individuals placed in the borderline have a special behavior in the schooling decision-making process. Moreover, it will allow identifying if being positioned bellow the limit (30%) is determinant when experiencing an unexpected health shock but not incurring in catastrophic payments.

The second column in Table 5 proposes a different specification of the basic equation, where families spending between 11% and 19% of their non-food income on health are the treatment group[4]. The purpose is to analyze whether a small amount of money spent on health, regardless the fact that these households also experienced a hospitalization event, influences schooling decisions. Model 2 reflects a positive estimate, although it is not statistically significant at conventional levels. Thus, with an annual expenditure on health above 10% and below 19% of the income, teenagers and families are not forced to decide whether to invest in education or health.

The estimates in Models 3, considering families spending between 20% and 29% of their income in health as treatment units, are slightly positive and statistically insignificant. The positive sign reinforces the direct relationship that health expenditure and schooling decisions maintain within households. In addition, since the null cannot be rejected, the results suggest that families not running into catastrophic payments to afford healthcare do not influence their children to interrupt their education. Consequently, it can be inferred that for the case of Mexico, out-of-pocket health expenditures have a negative influence on school completion.

Additionally, variables as household income and household size are some of the major determinants of schooling resolutions in the families, being consistent and supportive with previous research (Fallon & Lucas, 2002; Banerjee & Duflo, 2007). Considering households spending between 11% and 19%, household income reduces the probability of interrupting education by 13%, while 12% is the case for the families with annual health expenditure between 20% and 29%. Moreover, when households experience catastrophic payments on health, initial wealthier families are less likely to oblige their children to interrupt education.

[4] Treatment groups in models 2 and 3were also selected based on a score matching but considering only households spending less than 30% of their income on health but above 10%.

Hence, initial income is a determinant factor when making educational choices, especially when the family and the youth are confronting economic hardship derived from unexpected health events.

7. Limitations and Caveats

The model presented above and the nature of the data used have some boundaries. Firstly, as analyzed, the Mexican Family Life Survey is not complete representative of the Mexican population. The small proportion of rural households may report biased estimates, based on the intuition that rural families are more prone to experience health shocks, and their observed initial low-income levels force them to incur in out-of-pocket expenditures. Thus, the findings can only be set for the local sample, incapable to be generalized for the all the countries in the developing world.

Additionally, even with the selection of one of the best strategies to find the average treatment effect on the treated in the absence of a randomization control trial, the results cannot isolate the non-random nature of the experiment. Thus, it should be recognized that there may exist some time invariant and varying omitted factors, which cannot be overcome neither with the use of the PSM nor with the DiD method. For instance, even controlling for the level of income and the amount of money spent on health, it is important to notice that catastrophic expenditures on healthcare may be underreported. Meaning that, families that do not experience income shocks do not necessary reflect that they are not suffering unexpected health-shocks. I.e. people that do not have enough endowments to afford healthcare and forced to interrupt medical care due to the high cost it implies, may not be presenting these OOP-health payments. Thus, the potential underrepresentation of this population might bias the estimates obtained.

The main drawback of this research relies on the missing 9% of the sample that was not surveyed due to migration and death. The missing sample may affect the results since migration or death could be direct consequences of health shocks and the economic instability derived from them. For instance, migration can be a response of financial instability within households when unexpected health events occur, and if it is the case, the omission of the variable may be biasing upward the outcomes. On the other hand, the case of the forgone income with the loss of the job is likely to be a drawback in this study as well. The youth living in a distressed family, which it is also affected by the loss of wealth derived from the

loss of the job, might increase the probability of interrupting education. Therefore, the potential correlation between these covariates and the probability of school interruption may produce overestimated results. Moreover, not accounting for this sample limits the generalization of the outcomes derived from the analysis.

Consequently, further research to study the alternative mechanisms in which schooling decisions are made when economic hardship is experienced should be conducted. Besides, more research should be driven considering these subsamples in order to measure the effect of health shocks on migration or death, or to estimate the potential reverse causality between both variables

Finally, despite of the robustness of the results, the main shortcoming present in the Kernel-matching procedure is that observations that can be considered as bad matches are possibly accounted to compute the final estimation. Meaning that, the robustness check even reporting strong significant results, also comprises for individuals that may not be directly comparable with the treatment units.

8. Conclusions

Medical care expenditures represent one of the major concerns in developed and developing countries since it is a direct mechanism to contribute to the human capital accumulation (Wagstaff & Van Doorslaer, 1998). In the absence of health coverage from the Government or from other entities, private payments become the unique instrument in which people are able to afford healthcare. If these private payments exceed the 30% of the non-food income in a year, they are considered as catastrophic expenditures, which may threat the economic stability of a family (Senne, 2014).

Evidence for the developed and developing world has shown that the income distress faced after experiencing unanticipated health shocks has on average, serious repercussions on the expenditure behavior within households (Beegle, *et al.,* 2006; Fallon & Lucas, 2002). The reduction in resources directed in basic needs such as food, housing, clothing and education are the most outstanding endangered aspects that families and individuals have to sacrifice in order to fund healthcare (Dureya, Lam & Levinson, 2003).

This study hypothesized that experiencing a health shock, and the consequent out-of-pocket spending derived from it, forces households to reduce expenditure in education reflected in

the increased school interruption. Using a combination of methodologies, such as the PSM and the DiD, the case of Mexico revealed that catastrophic health spending maintain a close relationship with school interruption.

Controlling for initial income, health status, household size and siblings composition, parental education and the type of school the student is attending, the out-of-pocket health spending shows a direct and strongly significant relationship with the probability of interrupting education. Namely, experiencing catastrophic payments on health increases the probability of the student to interrupt his/her education being consistent to the theory and recent findings (Dureya, Lam & Levinson, 2003; Johnson & Reynolds, 2011).

Several factors lead to consider that the link between out-of-pocket health expenditures and school interruption is causal. First, treatment and comparison groups showed similar trends in the pre-shock period, thus they were balanced and can be considered to be comparable. Second, health shocks affect school interruption from income-related issues produced by the health distress, but it showed no effect when different specifications of health expenditures where performed. Third, the effect of catastrophic health payments was largest for families with the lowest income, meaning that financial endowments are determinant when schooling resolutions are made.

This last result consists in one of the main findings obtained in this research, supporting the theoretical discussion where initial income maintains a strong influence with schooling decisions, especially when financial catastrophe occurs. Meaning that, initially well-endowed families are more likely to encourage their children to continue their education if health shocks are experienced. The statistically significant results reveal that having higher family income when disturbed by a health shock reduces the likelihood of school interruption and coherent with the theory (Caneiro & Heckman, 2002), the initial level of income is determinant in the decision making process within households. Moreover, initial endowments do determine children's educational outcomes in the future, and wealthier families appear to be more protected against financial instability.

This pronouncement is important because of the policy implications that can be derived from it. Considering households' endowments as determinant for the schooling decisions when a financial distress occurs, low-income families might be trapped into poverty (Subramanian & Deaton, 1996). Explicitly, when facing economic instability, poor families cannot continue

investing in education, which in turns imply a reduction in their human capital accumulation. Overall, low investment in human capital, in both health and education, derives into lower income when accessing the labor market, conforming a vicious cycle making poverty to prevail (Thomas, *et al*., 2006). Meaning that, worse endowed families are more likely to be forced to interrupt their children's education, implying the reduction of human capital formation and a lower income in the future. This may produce poor families to be deprived to scape from poverty if health shocks producing the disbursement of more than 30% of their income are experienced (Dasgupta, 1997).

Additionally as expected, if these wealthier families are headed by more educated parents, the children present lower expectations to delay their instruction. For the case of Mexico, parental education presents a negative relationship with school interruption. Results are coherent with previous findings and the theoretical discussion, where the education of parents emerges to be a strong determinant of children performance (Altonji and Dunn, 1996). Nonetheless, even with a strong significance of the estimate, the small magnitude of it cannot be considered as the main factor discouraging school interruption. Therefore, the positive aspect derived from the result is that even with low parental education, young Mexicans can still be encouraged to get enrolled in the scholar system.

Other interesting outcome obtained is that youth belonging to larger families are more likely to interrupt their education. Consistent with the quality-quantity theory proposed by Becker (1960), the number of children and the number of dependents within a household do influence children performance in their later years. The results show that the size of the family is negatively correlated with school interruption; i.e. families with a small number of members are less likely to force the school interruption of one of their members, regardless the presence of economic hardship, and vice versa.

This last statement is supported with the direction of the interaction term between the size of the family and catastrophic payments. However, for the case of Mexico, given the statistically insignificance of the result, it cannot be expressly stated a direct effect of catastrophic spending on school interruption. Following Black, Devereux and Salavanes (2005), it could mean that Mexican families are possibly absorbing some of the burden coming from a health shock, spreading the fixed costs derived from catastrophic payments among a larger number of members.

On the other hand, the results obtained also suggest that teenagers are the most prone to abandon or delay their education, especially if they are the oldest among their siblings. The finding reinforces the study of Krueger and Perri (2009), where the birth-order clearly affects the schooling decisions within households, being the oldest ones the most exposed. But opposite to them, the magnitude of the parameter and the fact that this estimation does not control for the exact composition of siblings and their ages, does not allow stating that not being the oldest sibling is an automatic barrier to be insulated from the burden.

This context suggests that more research should be conducted to test whether the birth-order affect schooling decisions and resource allocation within families, particularly for the case of developing countries given that they tend to present larger families and a higher dependency ratio. In addition, since the results insinuate that some of the Mexican youth are interrupting their education, probably because they have found in the labor market a place to be, the sibling's composition is relevant to be studied in the future.

The results reported are statistically significant in their majority, being robust when different specifications are proposed. In the sensitivity analysis, through the enlargement of the sample with the inclusion of imperfect matches, the Kernel-matching procedure confirms the negative relationship between catastrophic health payments and educational choices. Moreover, it gives the possibility to expand the findings of this research to a larger proportion of population, since it accounts for those individuals that are placed in the border of being either control or treatment groups. Therefore, presenting similar characteristics, the new units of treatment and comparison are the ones that can be selected to determine what would happen if an existent program to cover against out-of-pocket health payments was intended to be widespread to a larger portion of the population.

Additionally, different specifications of health spending were conducted to measure the effect of spending less than 30% of the non-food income on health on school interruption. The estimations reflect that non-catastrophic payments do not have any effect on the probability of interrupting education. Moreover, the results support the previous findings for the control variables used in the benchmark equation. Therefore, it can be concluded that out-of-pocket health expenditures increases the likelihood of school interruption of the Mexican youth. Summarizing, it can be expressed that catastrophic health payments does derive into harsh

decisions within families, and households suffering from financial catastrophe are forced to reallocate their available endowments, being education one of the main disturbed areas.

After doing an extensive literature review and to the best of my knowledge, this is the first paper analyzing the effect of catastrophic health spending on school interruption in Mexico and for any other Latin American country. Therefore, the study of Mexico is relevant to contribute to the general knowledge and theoretical discussion of the potential effects that meeting out-of-pocket payments have on school interruption. Moreover, it may further motivate the analysis of other determinants of education, not only for the case of Mexico, but also for other developing countries such as the Latin American ones.

Finally, the present study can derive into one main policy implications for Mexico and for other similar emerging countries. The need of universal health coverage for families is crucial, being especially important for the households placed in the lowest quintiles of the population, since they appear to suffer more for the burden derived from catastrophic health spending. Consequently, health coverage for the poorest families will certainly allow improving the health performance, but also would produce higher educational and completion rates among children and teenagers. Altogether, the conjunctions of both factors are able to increase the levels of human capital in the country, enhancing the population welfare if it is reflected in the enlarged income. As a long-term as reviewed, people can escape from poverty if they are provided with more healthcare services and more education.

References

Agrawal, A., Kapur, D., McHale, J., & Oettl, A. (2011). Brain drain or brain bank? The impact of skilled emigration on poor-country innovation. *Journal of Urban Economics*, *69*(1), 43-55.

Altonji, J. G., & Dunn, T. A. (1996). The effects of family characteristics on the return to education. *The Review of Economics and Statistics*, *78*(4), 692-704.

Anderson, G., & Hussey, P. S. (2001). Comparing health system performance in OECD countries. *Health Affairs*, *20*(3), 219-232.

Angrist, J. D., & Pischke, J. S. (2009). Mostly harmless econometrics: An empiricist Vs companion. *Princeton Univ Pr*.

Baland, J. M., & Robinson, J. A. (2000). Is child labor inefficient?. *Journal of Political Economy*, *108*(4), 663-679.

Banerjee, A. V., & Duflo, E. (2007). The economic lives of the poor. *The journal of economic perspectives: a journal of the American Economic Association*,*21*(1), 141.

Becker, G. S. (1960). An economic analysis of fertility. In *Demographic and economic change in developed countries*. Columbia University Press, 209-240.

Becker, S. & Ichino, A. (2002). Estimation of average treatment effects bassed on propensity scores. *The Stata Journal,* 2(4), 358-377.

Becker, G. & Tomes, N. (1976). Child Endowments and the Quantity and Quality of Children. *Journal of Political Economy*, 84(4), 143-162.

_____ (1994). Human capital and the rise and fall of families. In *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education (3rd Edition)* (pp. 257-298). The University of Chicago Press.

Beegle, K., De Weerdt, J., & Dercon, S. (2006). Orphanhood and the long-run impact on children. *American Journal of Agricultural Economics*, *88*(5), 1266-1272.

Belley, P. & Lochner L. (2007). The Changing Role of Family Income and Ability in Determining Educational Achievement. *Journal of Human Capital*, University of Chicago Press, 1(1), 37-89.

Black, S., Paul J. Devereux, P. & Salvanes, K. (2005). The More the Merrier? The Effect of Family Size and Birth Order on Children's Education. The Quarterly Journal of Economics , 120 (2): 669-700.

Caliendo M. & Kopeinig S. (2008), Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys*, 22: 31-72.

Carneiro, P. & Heckman J. (2002). The Evidence on Credit Constraints in Post-Secondary Schooling. *Economic Journal*, 112(482), 705-734.

Consejo Nacional de Evaluación de la Política de Desarrollo Social (CONEVAL) (2013). Uso de una encuesta panel para evaluaciones de impacto: ensayo metodológico con la ENNViH 2002-2005, México, 145.

Dasgupta, P. (1997). Nutritional status, the capacity for work, and poverty traps. *Journal of Econometrics*, *77*(1), 5-37.

Dehejia, R. H., & Wahba, S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American statistical Association*, *94*(448), 1053-1062.

DiNardo, J., & Tobias, J. L. (2001). Nonparametric density and regression estimation. *Journal of Economic Perspectives*, *15*(4), 11-28.

Eberhardt, M. S., & Pamuk, E. R. (2004). The importance of place of residence: examining health in rural and nonrural areas. *American Journal of Public Health*, *94*(10), 1682.

Encuesta Nacional de Niveles de Vida de los Hogares (2005). Mexican Family Life Survey,2002-2005. Retrieved from: http://www.ennvihmxfls.org/es/ennvih.php.

Fallon, P. R., & Lucas, R. E. (2002). The impact of financial crises on labor markets, household incomes, and poverty: A review of evidence. *The World Bank Research Observer*, *17*(1), 21-45.

Galiani, S., Gertler, P., & Schargrodsky, E. (2005). Water for life: The impact of the privatization of water services on child mortality. *Journal of political economy*, *113*(1), 83-120.

Ghez, G., & Becker, G. (1975). The allocation of time and goods over the life cycle. *NBER Books*.

Glennerster, R., & Takavarasha, K. (2013). *Running Randomized Evaluations: A Practical Guide*. Princeton University Press, 472.

Hanushek, E. (2002). Publicly Provided Education. *Stanford University Publications*. NBER, 2046-2125.

Haque, N. U., & Kim, S. J. (1995). "Human capital flight": Impact of migration on income and growth. *IMF Staff Papers*, *42*(3), 577-607.

INEGI (2004). "Diseño muestral". Descripción de la muestra de la línea basal de la ENNViH. Documento de trabajo. CIDE y UIA. Retrieved from: www.inegi.org.mx

Johnson, E. & Reynolds, L. (2011). The Effect of Household Health Shocks on the Educational Attainment of Youth. Kent State University, 1-42.

Knaul, F. M., Arreola-Ornelas, H., Méndez-Carniado, O., Bryson-Cahn, C., Barofsky, J., Maguire, R., & Sesma, S. (2006). Evidence is good for your health system: policy reform to remedy catastrophic and impoverishing health spending in Mexico. *The Lancet*, *368*(9549), 1828-1841.

Kremer, M. (2003). Randomized Evaluations of Educational Programs in Developing Countries: Some Lessons. The American Economic Review , 93 (2), 102-106.

Kremer, M. and Miguel E. (2004). Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities, *Econometrica* 72(1), 159-217.

Kremer, Michael, Edward Miguel, and Rebecca Thornton, January (2009). Incentives to Learn. Mimeo Harvard University , 91(3), 437-456.

Krueger, D., & Perri, F. (2009). How do households respond to income shocks?. *University of Pennsylvania, Mimeo*.

Lundborg, P. (2013). The health returns to schooling—what can we learn from twins?. *Journal of population economics*, *26*(2), 673-701.

McKenzie, D. J., & Sasin, M. J. (2007). *Migration, remittances, poverty, and human capital: conceptual and empirical challenges* (Vol. 4272). World Bank Publications.

Narayan, D., Schafft, K., Rademacher, A., & Koch-Schulte, S. (2000). Voices of the Poor: *Can anyone hear us?*. New York: Oxford University Press, (1), 271.

OECD (2013). Economic Survey of Mexico 2013: How Does Mexico Compare. Retrieved from: www.oecd.org/eco/surveys/mexico-2013.htm

_____(2013). How is Life? At a Glande. *How's Life? 2013: Measuring well-being,* OECD Publishing. Retrieved from: http://dx.doi.org/10.1787/how_life-2013-6-en.

Rivera, J. A., Barquera, S., González-Cossío, T., Olaiz, G., & Sepulveda, J. (2004). Nutrition transition in Mexico and in other Latin American countries.*Nutrition reviews*, *62*(s2), S149-S157.

Rubalcava, Luis y Teruel, Graciela (2006). Guía de Usuario de la Encuesta Nacional sobre Niveles de Vida de los Hogares Primera Ronda".

Rubalcava, Luis y Teruel, Graciela (2007). Primera ronda de la Encuesta Nacional Sobre Niveles de Vida de los Hogares. Retrieved form: www.ennvih-mxfls.org.

Rubalcava, Luis y Teruel, Graciela (2008). Guía de Usuario de la Encuesta Nacional sobre Niveles de Vida de los Hogares Segunda Ronda.

Senne, J. N. (2014). Death and schooling decisions over the short and long run in rural Madagascar. *Journal of Population Economics*, *27*(2), 497-528.

Subramanian, S., & Deaton, A. (1996). The demand for food and calories.*Journal of political economy*, *104*(1), 133.

Thomas, D., Beegle, K., Frankenberg, E., Sikoki, B., Strauss, J., & Teruel, G. (2004). Education in a Crisis. *Journal of Development Economics*, *74*(1), 53-85.

Thomas, D., Frankenberg, E., Friedman, J., Habicht, J. P., Hakimi, M., Ingwersen, N., ... & Wilopo, S. (2006). Causal effect of health on labor market outcomes: Experimental evidence.

UNICEF (2007). Análisis preliminary de los Resultados del Módulo de Trabajo Infancil 2007. (Encuesta Nacional de Ocupación y Empleo, 2007). Retrieved from: http://www.unicef.org/mexico/spanish/proteccion_6928.htm

Van Doorslaer, E., O'Donnell, O., Rannan-Eliya, R. P., Somanathan, A., Adhikari, S. R., Garg, C. & Zhao, Y. (2006). Effect of payments for health care on poverty estimates in 11 countries in Asia: an analysis of household survey data. *The Lancet*, *368*(9544), 1357-1364.

Wagstaff, A., Doorslaer, E. V., & Paci, P. (1989). Equity in the finance and delivery of health care: some tentative cross-country comparisons. *Oxford Review of Economic Policy*, *5*(1), 89-112.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT Press, 954.

World Health Organization (2012). WHO Region of the Americas: Mexico statistics summary (2002-present). Retrieved from: http://apps.who.int/gho/data/node.country.country-MEX.

_____ (2006). BMI Classification: Global Database on Body Mass Index. Retrieved from <://apps.who.int/bmi/index.jsp>