in fulfillment of the requirements
for the degree of Doctor in Economics and Management

Topics in Population and Family Economics

Keiti Kondi

Thesis committee:

Prof.	David de la Croix	Université Catholique de Louvain (co-promoter)
Prof.	Thomas Baudin	IÉSEG School of Management (co-promoter)
Prof.	Jean-François Maystadt	Université Catholique de Louvain
Prof.	Frédéric Docquier	Luxembourg Institute of Socio-Economic Research
Prof.	Luca Pensieroso	Université Catholique de Louvain

Louvain-la-Neuve, Belgium
September, 2022

To my dear ones

Preface

This dissertation is composed of three chapters that study topics in population and family economics. It places households at the center of the investigation, and it answers questions about their decisions on fertility, abortion, migration and labor market outcomes. This thesis elaborates on the importance of cultural norms in decision making and it studies how different constraints can affect these decisions. The first chapter studies a mechanism that can lower the sex ratio at birth by implementing policies that eliminate the sex bias in parents' preferences for children and the gender gap in intra-household bargaining. This study tackles the problem of sex-selective abortion in Albania by developing a parsimonious theoretical model which incorporates different utilities for boys and girls, the bargaining between family members and the decision about abortion dependent on its cost. The model is calibrated using data from the Demographic and Health Survey dataset on Albania for the year 2008. We find that equal investment in children and gender empowerment are not strong channels in bringing sex ratios at birth down to biological values. There is still a presence of a residual that considering historical backgrounds it can be accounted for as cultural norms.

The second chapter investigates how slow deterioration of soil, caused by climate change, affects internal migration and household resettlement. To examine this link, together with my co-author Stefanija Veljanoska, we instrument soil degradation by using distant climate shocks and controlling for recent weather conditions. We use the Integrated Household Survey in Malawi for the years 2010-2016. The underlying mechanism is that soil degradation is harmful for agricultural productivity, and therefore food security, which incentivizes households to seek for better opportunities through pushing their members to migrate.

The third chapter studies whether cultural norms in the origin country, measured at different times, affect fertility and labor working time of second-generation migrant women in France. Together with Thomas Baudin, we investigate empirically and follow an epidemiological approach to test whether the culture of origin affects people's behavior and decisions. We rely on the dataset TeO (Trajectoires et Origines) on population diversity in France in 2008. We find that not only culture is an important aspect in decision making but its power depends on the timing when the cultural norms are measured. Norms transmitted from peers can be stronger than those transmitted from parents. Moreover, the feeling of being French moderates the persistence of cultural norms differently for fertility and labor force participation, while the perceived feeling of being discriminated does not alter the persistence of the cultural norms.

Contents

1

Gender Gap, Intra Household Bargaining and Sex Selective Abortion in Albania [1]

[1] The author would like to thank David de la Croix for precious comments and suggestions, as well as Frédérick Docquier, Luca Pensieroso, Fabio Mariani, Bastien Chabé-Ferret, Paula Gobbi and members of ARC Workshop on Family Transformations: Incentives and Norms (Durbuy, BE) and participants at seminars and conferences at UCLouvain, University of Warwick, ASREC & Chapman University, EUI, University of Luxembourg, European Health Economics Association, University of Aix-Marseille, University of Cergy, University of Toulouse, German Society for Demography, Centro Studi Luca d'Agliano (Gargnano), CBS and British Society for Population Studies (Winchester) for great discussions.

Abstract

Among European countries, Albania has by far the highest sex ratio at birth with 1.12 boys per girl, compared to the European average of 1.058. Considering this disbalance, this study analyzes the relative importance of three underlying mechanisms: sex bias in parents' preferences for children, the gender gap in intra-household bargaining, and cultural norms. We develop a parsimonious model which incorporates different utilities for boys and girls, the bargaining between family members, and the decision about abortion dependent on its cost and including social stigma. We calibrate the model using data from the Demographic and Health Survey dataset on Albania for the year 2008 by measuring the preference for children of both sexes by the time invested in the child, education, violence, and women empowerment by how independent are women in taking their own decisions. We find that if we increase gender empowerment and equalize preference in children, the sex ratio decreases by 0.04 but it remains above its biological value. This residual could be interpreted as cultural norms affecting the decision-making and not allowing the above indicators in bringing sex ratios down to their biological values. To conclude we discuss different policies that can help in the decline of the sex ratio disparity while accounting for norms.

"150 euros if the fetus is 4 months, 230 euros if it is 5 months" - Abortion procedure prices in private clinics in Albania

1.1 Introduction

Sex-Selective abortion is the interruption of pregnancy because of the sex of the child. In Albania, the preference for male children has led to sex-selective procedures and termination of the pregnancy based on the sex of the fetus, not on other grounds. This is the case when the child happens to be female. The most common procedure is sex-selective abortion, which leads to sex imbalance. The sex of the child can be known only after twelve weeks of pregnancy and the mother can proceed with sex selection only after this period of time. There can be other reasons to abort and we assume no interaction between gender and these reasons. A couple's preference to have control over their fertility can be done medically through in vitro fertilization IVF and PGD treatments (Sills and Palermo (2002)), artificial insemination AI (Wert and Dondorp (2010)) and Ericsson method (Silverman et al. (2002)). They are not always effective, and their high cost makes them not affordable and not accessible everywhere. In many developing countries a cheap and successful way remains the sex-selective abortion. Abortion laws vary by country and it can be legal by request, but sex-selective abortion remains illegal worldwide. Nevertheless, altering the distribution of children is considered socially acceptable depending on the country's norms, so certain doctors undergo abortion even after the twelfth week of pregnancy. The biological sex ratio at birth is 104-105 boys per 100 girls (Pergament et al. (2002), Rubin (1976), Jacobsen et al. (1999), Ben-Porath and Welch (1976)). The sex ratio at birth (SRB) depends on two factors: the sex ratio at conception and the sex-specific survival rates of the fetus during gestation (Chahnazarian (1988)). The first one is called *the primary sex ratio* and the last one is the *secondary sex ratio*, considering medical complications and abortions. At birth, males outnumber females, but this phenomenon is partly offset during the early years of childhood since infant mortality is higher for males. Improvement in maternal health, nutrition, and medicine has lowered child mortality (World Health Organization). Since there is a decline in male mortality, the sex ratio of infants, has become more male-biased. Interruption of pregnancy due to sex selective abortion is particularly notable in South East Asia (Das-Gupta et al. (2003), Das-Gupta et al. (2009)). After the introduction of echography in the 90' in the Balkans, it became a phenomenon in Albania and the regions inhabited by the ethnic Albanian population. From the UNFPA Report 2011 on Sex Imbalances in Albania the ratio at birth is 112 boys per 100 girls. Data from the Institute of Statistics show an increase in the sex ratio after 1990, reaching its highest peak in 2009, followed by a downward trend in the coming years.

Although abortion can take place for any number of reasons, sex-selective abortion happens for the desire to have a son over a daughter. This deliberate termination of a pregnancy due to the fetus' presumed sex leads to a sex ratio disbalance. Considering this practice and the uneven number of sons and daughters on an aggregate level, we consider an abortion decision model. This can map several dimensions of gender empowerment into sex-selective abortion of girls. We look at three mechanisms i) the sex-biased preferences for children and ii) the gender gap in intra-household bargaining power, iii) cultural norms. We develop a model where parents decide to continue or terminate a pregnancy based on these mechanisms. In the first channel, which is the preference of children, boys receive more weight than girls in the preferences of parents (Selenica (1927), Guilmoto and Duthe (2013)). We measure the sex bias in preference by looking at a time investment in children, education, and violence for children of both sexes. The second channel, gender empowerment, is measured by looking at women's independence in taking decisions in the household. Another point is also that we do not have data on abortion by gender of the child at the individual or regional level, hence we proceed with a structural model. The third channel is unobserved cultural norms that affect parents' preferences for the sex composition of their children.

What is new about this model is that it considers a utility cost of abortion, which enters the utility function of the family differently for the female and the male partners. The cost of abortion is incorporated in the utility of the family and parents decide on interrupting the pregnancy by comparing the utilities of keeping the child versus the expected utility of aborting. We consider families composed of two parents and two children. We analyze the first and the second pregnancy by considering the utility of having the child of a specific sex and also the cost of abortion. When during pregnancy the child is found to be male, the pregnancy will continue by assuming a male preference. When the child is found to belong to the female sex, the mother will continue the pregnancy only when the utility of having a female child is greater than the expected utility of aborting a child. After the abortion, there can be another pregnancy and we follow the same logic of events until the family terminates its fertility. We allow for only one abortion by considering that more than one can have serious health outcomes for the mother. This model is adequate for countries with low fertility rates below the replacement level. Many countries with high fertility, rely on the stopping rule for the sex composition of their children (for example they keep having children until they get a male child). Countries with low levels of TFR, rely on sex-selective techniques to compose the sex of their children, considering that low fertility acts as a strong constraint to the reproductive goals of parents.

The aim of this study is to evaluate the quantitative significance of preference in children and gender empowerment, on the sex ratio at birth, by doing a sensitivity analysis of the children's sex ratios with these two channels. The model is calibrated by using the Demographic and Health Survey data on Albania for the year 2008. Then we estimate the values of sex ratio for gender equal values of preference in children and women empowerment. Findings show that preferences in children and women's bargaining power can help explain cross-cluster variation in sex ratio but are not the main determinants. Removing the bias in preferences and bargaining power would lead to a sex ratio of 1.16, which is lower than the observed one 1.22, but still higher than the biological one.

The first contribution of this paper is to provide an overview of the problem with sex-selective abortion and why it is problematic. The second is in the creation of a model that accounts for fertility decisions under abortion cost constraints. The main contribution of this paper to the literature on fertility and intra-household decisions in family economics and gender studies. It moreover contributes to the literature on missing women and the determinants that help eliminate this phenomenon. Whereas the focus of the existing literature is on gender inequality and fertility decisions, this study introduces and accounts for the utility cost of abortion, which affects the decision of the parents during pregnancy. We also consider analyzing how cultural transmission of norms stands in this problematic. Taken together, the findings indicate that son preference is driven by deeply rooted norms and how people respond to these norms. The remainder of the paper is organized as follows. Section 2 reviews the literature for the most problematic countries concerning sex-selective abortion. We discuss theoretical and empirical evidence that suggests that sex-selective abortion is a problem. Section 3 describes the structural model. Section 4 explains data sources, variable construction, and provides summary statistics. The calibration, identification, over-identification, and counter-factual simulation are presented in Section 5. A policy recommendation is proposed in section 6. Section 7 concludes the findings of this study.

1.2 Literature Review

A strong preference for sons rather than daughters has been seen in many societies around the world. Figure 1 shows countries that have values of sex ratio at birth varying from 1.08-1.17. Azerbaijan, Armenia, and China record the highest levels of birth masculinity, but skewed SRB levels are observed in several other countries in Eastern Europe, South Asia, and South Asia. While each country is characterized by specific reasons, they all come down to son preference. The prevalence of dowries is known to be the main reason of sex-selective abortion in India. In China it is the one-child policy (Guilmoto (2010) Goodkind (2015)).[2] Therefore, different motives and a combination of several factors might affect the propensity to select the sex of future children. Apart from sex-selective abortion, other methods that can shape the sex composition of children. Among them is the stopping rule, postnatal sex-selection, and adoption. Prenatal sex-selection remains the most efficient way for preventing unwanted females. This has turned into a norm and habit for a section of the population, especially in low or middle-income countries (Guilmoto et al. (2018a)).

[2]Sex selection in China is more complex than one-child policy. Guilmoto (2010) links this problematic not only to the one-child policy, but also to cultural norms and old age dependency

In this section, we review the literature on sex imbalances, the masculinization of births, how they become problematic, and why it is important to account for them considering their consequences in the marriage market and the presence of violence. Moreover, we review papers on sex selection in South-East Asia and China, countries known for sex imbalance in their population. Then we continue with sex selection in Albania.

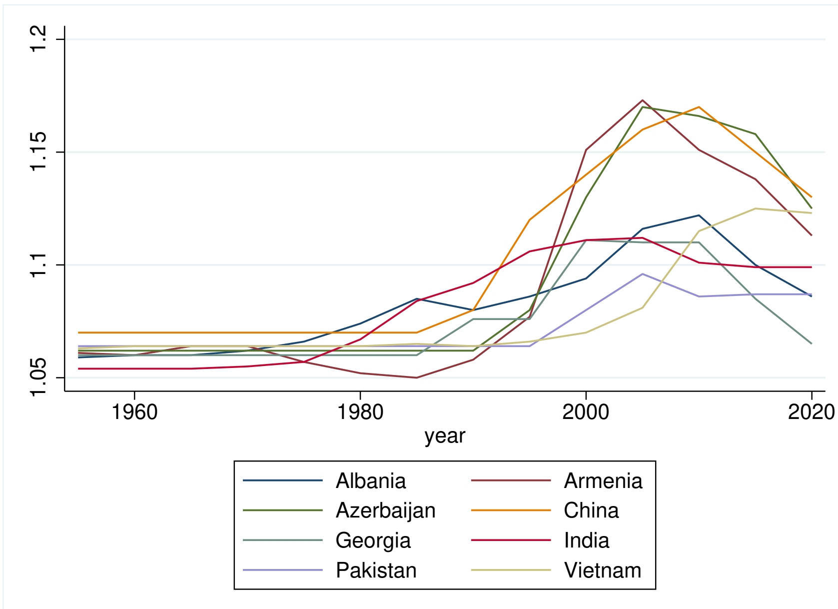


Figure 1.1. Sex Ratios at birth of most problematic countries *UNDP Data: World Population Prospects: The 2019 Revision*

1.2.1 Consequences of sex imbalance

Sex imbalance is problematic because of the consequences it can have on the population. Choosing the sex of the children creates a bias towards men. The high number of men in confront with females affects the marriage system by increasing the number of unmarried men in the marriage market, which delays their marriage. Men will have difficulty finding a marriage partner and starting a family of their own. This will lead to hyper-competition between men in the marriage market. Men belonging to a lower social class will be affected by remaining single (Becker (1974)).

Additional to problems in the marriage market, there is pressure regarding savings for couples that have a son, both cross-regional and household-level to improve their son's relative attractiveness for marriage (Zhang and Wei (2011)). As a result, job mobility and job participation may change. The family structure inhabits changes as well. In developing countries where unmarried men still live with their parents, they continue to do so until the time of the marriage and even after their marital status changes (Wang (2005)).

Other negative consequences can be in women's position in society. It can remain behind that of men and their lower social role may lead to gender-based violence, sex trafficking, prostitution, job discrimination, and less power in decision-making both at the household level and in society (UNFPA 2011). This surplus of men can also account for the increase in violence in society (Edlund et al. (2013)). There is an increase in the overall violence, committed by unmarried males who lack social bonds. In areas with more male-biased sex ratios both men and women continue to have more conservative attitudes towards women working, and women work fewer hours outside the home (Grossjean and Khatar (2018)). Fatherhood and marital status connected to social standing may affect the connection between males and crime activities (Hudson and den Boer (2004), Barber (2000)), which may as well lead to the development of an authoritarian political regime. Negative consequences for women especially violence towards them, coerced marriages, bride abduction, rape, and violent communities will as well increase at the population level.

Gender bias in the sex ratio can also lead to imbalances in education in the later life of children, leading to gender inequality (Wang (2005)). The marriage market and violence are interlinked with the education levels of youths and young adults, which later on affect the labor market outcomes. This is problematic because it reinforces the idea that one sex is better than the other, which can create an environment that is hostile or exclusionary to the members of the sex with fewer options.

The counter-factual of the above consequences would be that this abundance of men and this scarcity of females can be turned into an advantage for females. Can this scarcity increase the value of women in the marriage market? Considering Becker's theory on marriage markets for the smaller group, their weight in the matching process would be higher, if it was not for their consideration in society. Considering that this study is about sex-selective abortion in Albania and that there is high levels of gender inequality in Albanian society, we exclude this possibility. The advantage power of females would be in the selection of wealthier men in the marriage market but not in the bargaining power and decision making.

Moreover, the same sex can share resources better. Considering this, families that have a male first child, might be influenced in their decision for the sex of the second child, by accounting for shared costs (Himmelweit et al. (2013)). This increases the sex ratios for the consecutive children entering into a loop of male preference. Another circumstance is when parents want a male child at all costs, and they stop having children once they have a male child (Basu and De Jong (2010)).

1.2.2 Sex selection worldwide

In countries where the underlying context of son preference does not exist, the availability of techniques to determine sex does not necessarily lead to their use for sex-selection. Contrary, in countries with male preference, these techniques lead to sex-selective procedures and distorted sex ratios at birth. Sex-selective abortion has a presence in Eastern, Southern, and South-East Asia. China, India, Pakistan, Vietnam, Nepal, Armenia, and Azerbaijan are some of the countries with the highest sex ratios at birth due to sex-selective abortion. Other regions are the south Caucasus and some parts of Indonesia but not in the whole country. South Korea is an example of a country that struggled with this problematic but has now balanced the sex ratios at birth through awareness and campaigning (Das-Gupta et al. (2003)).

One of the drivers of sex-selective abortion in China is the one-child policy. This is particularly present in rural areas, where the son preference is strong. The 1.5-child loophole, where parents are allowed to have a second child in case the first one happened to be female lowers the value of having a daughter compared to a first child son. Moreover, there is evidence that also underreporting has affected distorted sex ratios at birth in China (Goodkind (2015)). The second more problematic country in terms of sex ratios at birth is India (Das-Gupta et al. (2003)). The factors behind this are not biological but this comes from the preference for sons (Das-Gupta (2005)). Considering that this problematic is strongest in rural areas, this could be linked also to ecological endowments. Hazarika et al. (2019) show that there are proportionately more missing women in countries whose ancestral ecological endowments were poorer.

Although distorted sex ratios at birth come down to son preference, drivers of this phenomenon vary by country. Different from India and China, Guilmoto et al. (2010) argues that the high sex ratio at birth in Vietnam is linked to various factors such as access to modern health care, number of prenatal visits, level of higher education and employment status, young age, province of residence and prenatal sex determination. In another study Guilmoto et al. (2018b), reveal geographical and socioeconomic disparities and show that after a period of increase it now appears to be leveling off, signaling a possible future trend reversal, which is a similarity it shares with India and China.

Another driver leading to sex-selective abortion can be economic uncertainty and poor institutions. Das-Gupta (2015) argues that sex selection manifested itself in Azerbaijan only after the sudden economic and governance meltdown following the dissolution of the Soviet Union in 1991 where jobs, basic services, and social protection mechanisms unraveled. Son preference in Azerbaijan was considered as a coping mechanism, where sons offer the traditional form of support under uncertainty. Basic services, pensions, and safety nets have been rebuilt, but the process involved years of policy changes.

Recently, this problematic is common not only in the east but also in the western and more developed countries with presence of Asian diaspora (Dubuc and Koleman (2007)), North America, in USA and Canada (Edlund and Almond (2008); Abrevaya (2009), Almond et al. (2013)), Italy (Ambrosetti et al. (2015)), Spain (Gonzalez (2014)) and Norway (Singh et al. (2010)). These papers show that when migrants move to another country, they bring with them cultural norms of son preference.

Culturally speaking, from psychological and historical constraints regarding judgment and sensibility differences, the solution has been to consider women as either deviant or deficient in their development (Gilligan (1982)). A factor that reinforces this is the family line continuation. It influences the decision of the parent to have a male child because the son takes the surname of the family (Xing and Hesketh (2006)). Son preference is not just about cultural norms, it can be also interpreted economically. One possible explanation is the old age dependency care. Parent's worry about old age dependency, can explain male preference as an insurance (Chen (2008)) or they usually get pressured by their families as it is the case in the study of Varma (2002) where urban, educated, and middle-class women undergo sex detection techniques. In countries with a poor institutional framework regarding pension schemes, old people rely on their children for financial support.

Some types of children's sex discrimination can be not only pre-natal (between unborn children), but also post-partum (after the children are born). Parents internalize certain norms that lead them to give better care to their sons compare to their daughters. This consists in taking care less of a female child compared to the male counterparts. It also contributes to a scenario where daughters are of a lower value than sons to their parents (Das-Gupta (1987)).

The problematic of this phenomenon is not only the negative aspects it can have on gender equality but also its persistence over time. The consequences of uneven sex ratios on cultural attitudes, labor supply decisions, and occupational choices can persist in the long run, well after sex ratios are back to the natural (biological) rate (Grossjean and Khatar (2018)). An example of this is this is distorted sex ratios during world war II in Russia and male scarcity, which lead to lower rates of marriage and fertility, higher non-marital births, and reduced bargaining power within marriage for women most affected by war deaths. The impact of sex ratio imbalance on marriage and family persisted for years after the war's end. It was likely magnified by policies that promoted non-marital births and discouraged divorce (Brainerd (2017)). During world war II men use their advantageous position in the marriage market to "marry up" and marry women of a higher social class. (Abramitzky et al. (2011)). Another addition to this hypothesis is from Bethman and Kvasnicka (2013) that analyze how the loss of men in World War II affected out-of-wedlock births in Bavaria where regions with greater loss of male life experienced significantly higher shares of out-of-wedlock births after the war.

1.2.3 Sex Selection in Albania

Albanian legislature allows abortion by request up to the 12th week. Nevertheless, sex-selective abortion remains illegal. The procedure can be done in private clinics around the country illegally, without proof or report of it to authorities (UNFPA (2011)). Considering that these statistics in Albania are not tracked and we cannot distinguish which part of total abortions can be sex-selective abortions, we decide to not rely on this data and focus on sex ratio at birth. The four reasons behind historical distorted sex ratios are under-registration, its patrilineal society, agricultural and financial support: i) Sex ratios have been historically distorted due to the under registration of female children. Apart from sex-selective abortion, there is evidence of preference for boys since 1920. Male children were being registered in the official registries after birth, while their female siblings were not. Another practice that reinforces son preference in 1920 is that women were following traditional medicine to understand the sex of the baby like the shape and the position of belly during pregnancy (Selenica (1927)).[3] ii) The second reason that has influenced son preference is Albania's patrilineal society during Kanun times (Kaser et al. (1996), Kaser (2008)), which gave rise to the phenomenon of becoming a sworn virgin and getting independence and full rights as a man. Cultural norms that push women to not give birth to females, are reasons that girls are considered a burden to their families and that they are unable to perpetuate the family lineage (Kaser et al. (1996)). iii) Another historical reason might be traced to agriculture. Hazarika et al. (2019) argue that patriarchal cultural norms may be traced to man's

[3]During the 60' the dictator Enver Hoxha, started campaigns for the modernization and uprooting of traditions considered retrograde.

transformation from hunter-gatherer to farmer activities. The authors' findings show that gender roles are more unequal in countries with longer histories of agriculture, which is also the case for Albania. In rural agricultural areas, the male sex can defend the territories, providing returns for the family. As a result, parents bias their investments toward sons. iv) Son preference might have its roots during the Kanun of Lek Dukagjin, a set of laws, where the daughter could not be part of the inheritance. The country still suffers a patriarchal society and a male is considered to bring luck, power, and financial support to the family. Considering that parents rely on their male children for financial support and old age dependency this affects the decision-making of women when it comes to the choice of continuing the pregnancy with a female child or terminating it (Kaser (2008)). Albanian culture continues to be patriarchal in terms of gender equality (UNFPA (2011)).[4]

Another driver of biased sex ratios is migration. In the Albanian context, migration of men has done the opposite effect. It stabilized male/female sex ratios. As a result marriage markets were not distorted and the sex bias phenomenon became less problematic in terms of aggregate sex ratios of the population (this is different for sex ratios at birth). Considering that no action was taken against this problematic, it led to the continuation of the preference for sons (Guilmoto et al. (2018a), UNFPA (2016)). As shown in Figure 2, ever since 1990, the national Sex Ratio at Birth (SRB) estimate has been around 110 and it reached its peak 111.5 in 2008. The rise of the ratio in the 90' corresponds to the introduction of sex detection technology, the lowering of fertility rates, and difficult economic conditions in the country. The demographic indicators in Albania show that the Total Fertility Rate (TFR) rate has decreased from above 6 in 1960 to 3.2 in 1990 and to 1.5 in 2015. Guilmoto et al. (2018a) argues that sex selection techniques are more common in countries showing a low TFR. Low fertility acts as a constraint to the reproductive targets of parents, where biologically without sex selection techniques around 1/4 of parents with two children would remain without a son. Additionally, a strong gender bias

[4]Gender equality and gender violence continue to be at the core of the UNDP programs in the country.

and not only introduction and presence but also greater access to sex-selection techniques, have influenced in bringing sex ratios to high levels. Changes in TFR levels could also pressure the reproductive choices of couples. For instance, having one child less on average, family size doubles the probability of having no sons.[5]

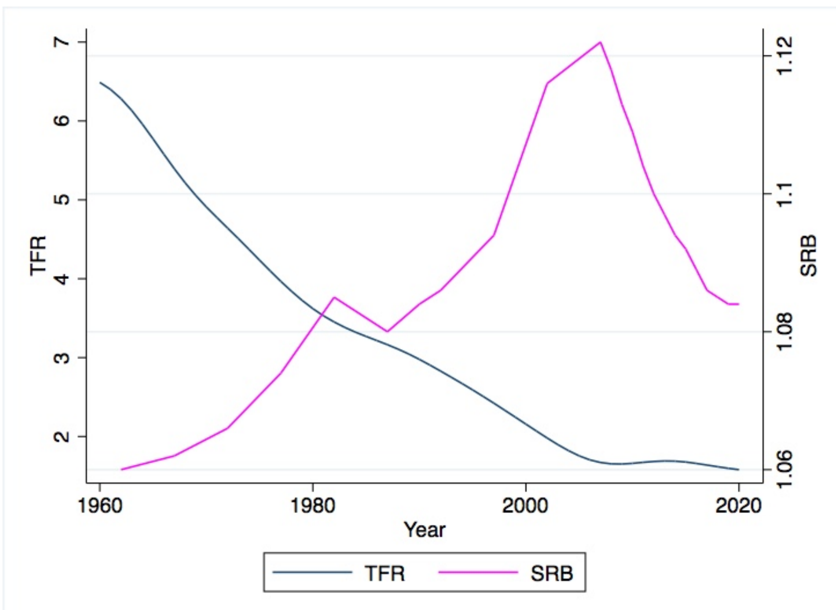


Figure 1.2. Sex Ratios at birth and TFR in Albania. *Source: World Bank*

[5]Such fertility decrease in less than five years has been observed in countries such as Brazil, Iran, Zimbabwe, and China (Guilmoto et al. (2018a))

1.3 Model

We consider a structural parsimonious partial equilibrium model for countries with low exogenous fertility. Low fertility acts as a constraint to the reproductive goals of parents. This constraint is even stronger in presence of son preference. Parents take abortion decisions trying to bias the sex of their children that maximize their utility. Considering the possible drivers of son preference, the model we construct will serve as an instrument to measure the sensitivity of abortion to gender gap policies and preferences in children. Our theoretical framework considers a partnership with a family planning of 2 children. We chose specifically 2 children since it is a representation of the average number of children in the Albanian society.[6] During pregnancy, the mother can give birth to a boy with a probability p or a girl with a probability $1-p$. Fertility is exogenously interrupted after a first child with probability q and continues fertility with probability $1-q$. Reasons behind a parent's decision to stop fertility can be the ideal number of children, financial constraints, age of parents, and a difficult abortion. In the data, we find that only 10% of parents stop having children after the first child. For simplification in our model fertility is not endogenous and mothers do not abort male children.

Utility of the family accounting for the cost of abortion

The utility of each parent is defined over personal consumption of parents c_M nd c_F, number of boys and girls as children n_b and n_g and their respective preference towards the child, which are parametrized by x_b and x_g. Mothers and fathers face the same preferences towards male children. The same goes for female children. The male child is preferred to the female child which is expressed by $x_b > x_g$. We write $g = \frac{x_g}{x_b}$. These preferences enter the utility function through γ that represents the weight at which children enter parent's utility function. The abortion cost enters the utility function for each parent differently considering different medical and financial situations they have to

[6]:In 2000 TFR is 2.2, in 2008 it is 1.7 and in 2022 it is 1.56.

face and undergo.

$$U_{Mother} = \ln(c_M) + \gamma \ln(n_b x_b + n_g x_g) - DA_M$$
$$U_{Father} = \ln(c_F) + \gamma \ln(n_b x_b + n_g x_g) - DA_F$$

The abortion cost faced by the mother is A_M and the one faced by the father is A_F. By D we express an abortion dummy that takes the value 1 when pregnancy is interrupted by sex selective abortion and the value 0 otherwise. The utility of the family is the sum of the utilities of both parents considering different bargaining power θ for each of the partners.

$$U = \theta U_M + (1 - \theta) U_F$$

After substituting for U_M and U_F the U of the family is expressed as following:

$$U = \theta \ln(c_M) + (1 - \theta) \ln(c_F) + \gamma ln(n_b x_b + n_g x_g) - A$$

where A is be the sum of the costs of abortion that each parent faces. For simplicity we allow one abortion possibility. This is also in line with the medical literature stating that an increase in the number of abortions increases health complications for the mother. We assume that the cost of abortion is the same before and after having a child. This means that the cost of abortion in the first pregnancy is the same with the one in the second pregnancy in case they did not undergo an abortion procedure in the first one.

$$A = \theta A_M + (1 - \theta) A_F$$

Considering that the mother is the one that undergoes the medical abortion procedure, we write that the cost of abortion of the father is only part of the cost of the abortion of the mother. This is expressed by $r > 1$.

$$A_{Mother} = r A_{Father}$$

Family faces the following budget constraint, where y is income, c_M and c_F represent the consumption of each partner, P shows the price of having children and n_b together with n_g are the respective number of male and female children. For simplicity P is not gender specific. P can be also considered as cost of children.

$$y = c_M + c_F + P(n_b + n_g)$$

In the first channel of this theoretical framework, the sex-biased preferences for children are captured by x_b and x_g and it captures the preferences of parents post-partum. This represents an investment in children after they have been born. The second channel of gender empowerment is captured by θ. The third channel of cultural norms is different from the first one as it shows the pre-natal preferences of children. This preference is important in the decision making if to continue or terminate a pregnancy considering the sex of the child.

Abortion decisions accounting for the sex of the child
During pregnancy, female and male children are born with different probabilities. There is a probability of $p = 0.52$ for the child to be born male and $p = 0.48$ for the child to be born female. After 12 weeks of pregnancy, the sex of the child can be verified through ecography. For simplification, if the child happens to be a male, the mother will continue the pregnancy and not consider an abortion. If the child happens to be a female, the family needs to take the decision if to continue the pregnancy or interrupt it by aborting.
They will choose the situation that gives them the highest utility. As a result they will abort when the $EU_A \geq U_G$ and continue the pregnancy when ($U_G > EU_A$). In the expected utility of abortion, parents also consider the probabilistic utility of a second pregnancy that leads to a male or a female child.

Considering our model, where parents take decisions based on the sex of their children, we provide possible scenarios of the pregnancy, conditional on parents having a higher preference for male children than for females $x_g < x_b$.

1. **When the first pregnancy, at time t leads to a boy**. This happens with probability $p = 0.52$. Consider that during the pregnancy, the chances for the child to be born a boy are with probability p and for a girl are $1 - p$.

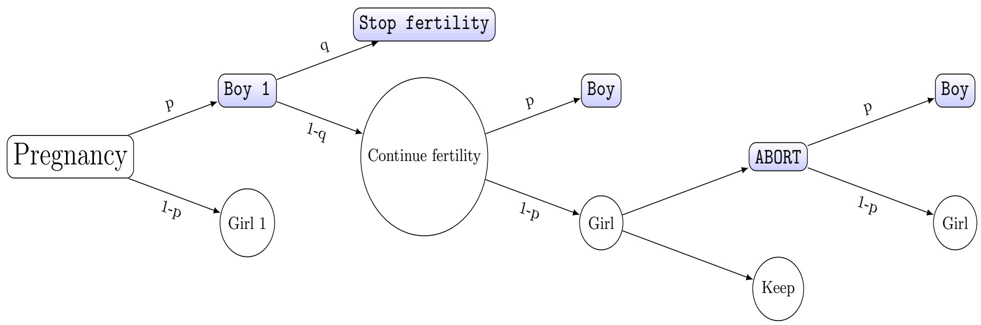


Figure 1.3. First child is a boy

The parents will always keep the first child when it happens to be a boy. At a second stage, $t+1$, after the first male child, the family has the possibility to stop fertility, which happens with probability q or to continue another pregnancy with probability $1-q$.
In case they continue with a second pregnancy, they will have a boy with probability p and a girl with probability $1-p$. They will always keep the second child if it happens to be a boy. Contrary, if the child happens to be female, the family will take the decision if to abort or continue the pregnancy with this second child according to their respective utilities.

In order to take this decision, the family will compare the utility of having a girl with the expected utility of abortion, considering a third pregnancy after abortion, at $t+2$. In this third pregnancy the probability of giving birth to a boy or a girl are again p and $1-p$ respectively.

Figure 1 shows a tree diagram with the decisions that parents can take when the first child is a boy and they are thinking of aborting or not the second child in case it happens to be a girl. The diagram shows also the respective probabilities of the parents' decisions and the pregnancy outcomes. By comparing the utility of keeping the girl with the expected utility of abortion, calculations of which can be found in the Appendix A.1, we can state that the pregnancy will be interrupted only when the cost of abortion is below the following threshold *(Proof 1 in the Appendix A.1)*. We see that the cost of abortion A depends on p, γ, x_b and x_g. Additionally, the choice to abort or not has no impact on the ultimate number of children and consumption. The purpose of the model is to compare "sub-utility" alternatives with exogenous fertility decisions.

$$EU(Abortion) > U(female)$$

$$\ln(c) + \gamma \ln(x_b + x_g) < pU(male) + (1 - p)U(female)$$

$$A < \bar{A} = p\gamma \ln(\frac{2x_b}{x_b + x_g})$$

2. **When the first child is a girl**, we analyze two possible scenarios. The first one is when the daughter is the first child, and the second one is about the second daughter. We analyze parents decisions regarding abortion in both these pregnancies.

   - When during pregnancy the child happens to be a girl, at time t parents have the choice to continue or interrupt the pregnancy. If they continue the pregnancy and give birth, at time $t + 1$ they can stop having children with probability q or continue with a second pregnancy with probability $1 - q$, which will lead to a boy or girl with their respective probabilities p and $1 - p$.

     In case they abort the girl, they can have another pregnancy at $t+1$ which can lead to a boy with probability p or a girl with probability $1 - p$. We assume one abortion possibility in this model, and as result, in the second pregnancy, in case the child happens to be a girl they are not anymore able to abort her. Moreover, the family can either stop fertility with probability q or continue fertility with a third pregnancy at $t + 2$, which will lead to a boy or girl with their respective probabilities p and $1 - p$.

     Figure 2 shows a tree diagram with the decisions that parents take and their respective probabilities in the case when the first child is a girl and they need to take a decision if to keep her or not. In order to take the decision if to abort or not the girl child, parents compare the utility of keeping the girl with the expected utility of abortion, considering a second pregnancy after aborting the first one. After comparing the two utilities, the pregnancy will be interrupted only when the cost of abortion is below the following threshold *(Proof 2 in the Appendix A.1)*.

$$EU(Abortion) > U(female)$$

$$q[\ln(c)+\gamma\ln(x_g)]+(1-q)[p(\ln(c)+\gamma\ln(x_b+x_g)+(1-p)[\ln(c)+\gamma\ln(2x_g)] < pU(B)+(1-p)U(G)$$

$$A < \hat{A} = p[-q\gamma\ln(\frac{x_b}{x_g})-(1-q)p\gamma\ln(\frac{2}{1+\frac{x_b}{x_g}})-(1-q)(1-p)\gamma\ln(\frac{1+\frac{x_b}{x_g}}{2})]$$

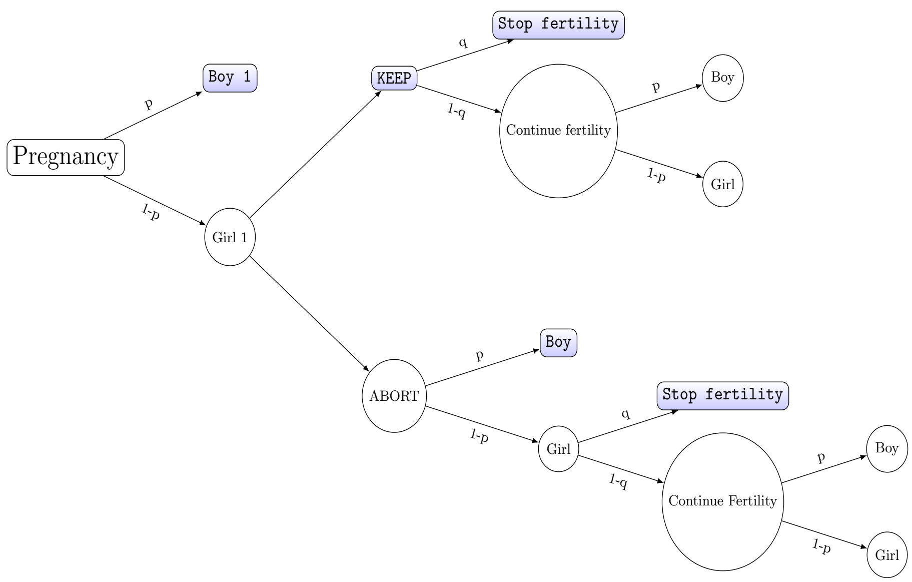


Figure 1.4. First child is a girl, abortion decision on the first child

- In the scenario where parents do not interrupt the pregnancy of their first girl at t and they continue with a second pregnancy at $t+1$. This second pregnancy will give a boy with probability p or a second girl with probability $1-p$. Parents are still allowed one abortion possibility. They face another choice, which is about keeping the second girl or terminating the pregnancy. Termination allows for a third pregnancy at $t+2$ that can lead again to a boy or girl with the usual probabilities of conception.

  In order to take this decision, parents compare the utility of having a second girl with the expected utility of abortion, considering a future pregnancy after aborting. Figure 3 shows a tree diagram with the decisions that the parents can take, the respective probabilities of their decisions and pregnancy outcomes. By comparing the utility of having a girl with the expected utility of abortion, we can derive that the pregnancy will be terminated when the cost of abortion is below the following threshold, calculations of which can be found in Appendix A.1.

$$EU(Abortion) > U(female)$$

$$\ln(c) + \gamma \ln(2x_g) < pU(male) + (1-p)U(female)$$

$$A < \tilde{A} = p\gamma \ln(\frac{x_b + x_g}{2x_g})$$

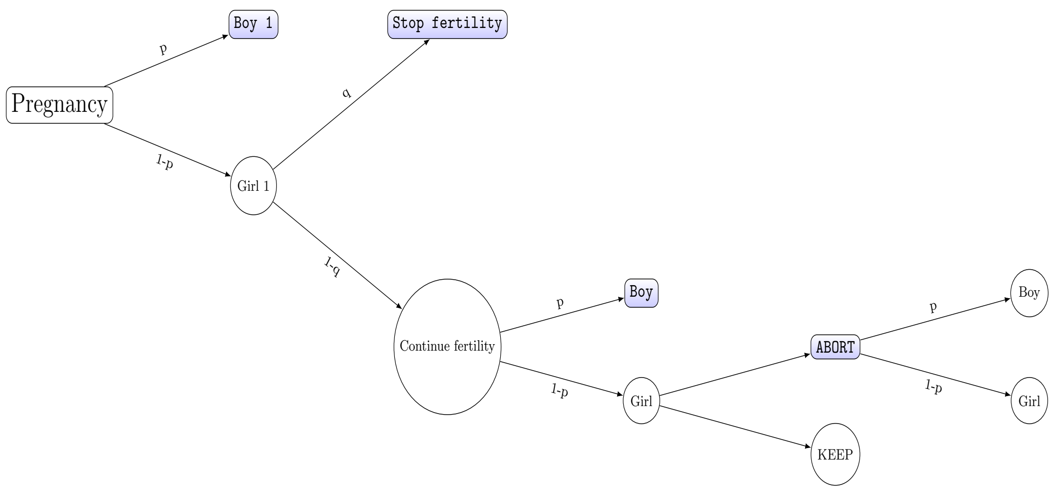


Figure 1.5. First child is a girl, abortion decision on the second child

We consider three thresholds for the cost of abortion for which parents consider when taking the decision if to terminate or continue a pregnancy. We write these thresholds as $\bar{A}$, $\hat{A}$ and $\tilde{A}$. By comparing them we can derive $\bar{A} < \hat{A} < \tilde{A}$. Considering these intervals, we can write the following propositions:

Proposition 1: $\tilde{A} > \bar{A}$
The abortion threshold for the decision to abort a second girl child when the first one is also a girl, is greater than the abortion threshold for which parents abort second girl child, after the first one is a boy. The ratio $\frac{\bar{A}}{\tilde{A}} < 1$. This means that a second child who is female is more often aborted when first child is female iff boys are preferred to girls ($x_b > x_g$, i.e. $g = \frac{x_g}{x_b} < 1$). This is also in line with the literature of having a boy at all costs. *(Proof of Proposition 1 in Appendix A.1.)*

Proposition 2: $\hat{A} < \tilde{A}$
The abortion threshold for the decision to abort a second girl child when the first one is a girl, is smaller than the abortion threshold to for which parents take the decision if to abort their first child which is a girl. If people have always two children, they abort more easily a first girl than

a second girl. Ratio $\frac{\hat{A}}{\tilde{A}} < 1$, which means $\tilde{A}$ is larger than $\hat{A}$, so first girls are more often aborted than second girls. *(Proof of Proposition 2 in Appendix A.1.)*

First girls are aborted more often than the second girls because the threshold for the cost of abortion for aborting the first girl child is lower than the threshold of aborting the second girl child. This is counter intuitive with the son preference from society pressure, once you have son, you have no more pressure for what child comes next.

1.3.1 Sex ratio and the cost of abortion

We want to compute the aggregate sex ratios and for this we consider the sex ratio for both the first and the second child. In order to bring the model to data, we assume that the cost of abortion A is distributed following a Logistic Distribution with the subsequent probability density function $f(x; m, b)$ and the cumulative density function: $G(x; m, b)$, where m and b are the location and scale parameters of the distribution:

$$f(x; m, b) = \frac{e^{-\frac{x-m}{b}}}{b(1 + e^{-\frac{x-m}{b}})^2}$$

$$G(x; m, b) = \frac{1}{1 + e^{-\frac{x-m}{b}}}$$

Some parents face negative cost of abortion. This means that their utility increases by aborting. An interpretation of this can be that cultural norms are strong enough to affect people's decisions regarding the sex of their children. When the cost of abortion is zero we can write:

$$G(0; m, b) = \frac{1}{1 + e^{\frac{m}{b}}}$$

Sex ratio of the children, for the groups of parents that do not abort will be the biological value of the number of boy children over the number of girl children, that have not been aborted. In probability terms we can write that sex ratio for parents that do not abort is the probability p to have a boy child over the probability $1-p$ to have a girl child. The sum of the two probabilities equals one: $p+(1-p)=1$. The sex ratio when parents do not abort will be the biological one and it can be expressed as: $SR_{na}=\frac{p}{1-p}$.

Sex ratio of the children, whose parents have aborted girls, will be greater than the biological one. In probability terms we can express it as: $SR_a == \frac{p+(1-p)p}{(1-p)-(1-p)p} = \frac{p(2-p)}{(1-p)^2}$, where the numerator $p+(1-p)p$ represents the number of boys, probability p to have a boy, and $(1-p)p$ is probability of having a boy after aborting a girl. The denominator $(1-p)-(1-p)p$ shows the probability $(1-p)$ of having a girl and $(1-p)p$ is what is left of girls after parents abort. Considering that $p=0.52$ and $1-p=0.48$, substituting these values in SR_{na} and SR_a we get: in case of non-abortion, sex ratio at birth will be $SR_{na}=1.05$ while in case of abortion, it will be $SR_a=3.2$. This last holds true for the first child and for the second child, irrespectively of the sex of the first child.

Sex ratio for the first child is the composition of children from parents that have aborted girls and the children from parents that have not aborted. As a result, we write that the sex ratio for the first child is the sum of the distribution of people that have aborted for the cost of abortion $\hat{A}$ and the distribution of people that have not aborted for the same cost of abortion. Figure 6 show schematically the two different groups. We write the sex ratio for the first child $s_i^1(m,b)$ in the following equation:

$$s_i^1(m,b)=G(\hat{A})SR_a+(1-G(\hat{A}))SR_{na}$$

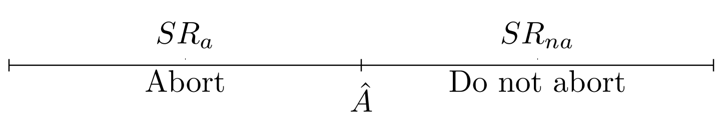


Figure 1.6. Sex ratio of children from parents that choose to abort and not abort the first child

Substituting for the values of $SR_a = \frac{p(2-p)}{(1-p)^2}$, $SR_{na} = \frac{p}{1-p}$ and the cumulative density function of the logistic distribution, we can rewrite

$s_i^1(m,b)$ as:

$$s_i^1(m,b) = \int_0^{\hat{A}} \frac{p(2-p)}{(1-p)^2} f(\hat{A})dA + \int_{\hat{A}}^{\infty} \frac{p}{1-p} f(\hat{A})dA$$

Sex ratio for the second child is composed by children of parents that have already aborted once and those that have not done so at their first pregnancy. Those that have aborted, have lost their one possibility of abortion. Figure 7 shows schematically the group of parents that cannot abort and their corresponding sex ratio. The share of these children is $1-p$ and the sex ratio $s_i^2(na)$ would be written as the following:

$$s_i^2(na) = (1-p)G(\hat{A})$$

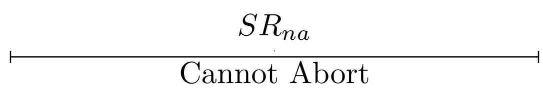


Figure 1.7. Sex ratio of children from parents that have already aborted the first child

When parents still have the possibility to abort, they would do so when the cost of abortion is up to the interval $\tilde{A}$ when the first child was a girl and up to the interval $\bar{A}$ when the first child was a boy. This is represented schematically in Figure 6 and 7.

Sex ratio of children from parents that choose to abort or not abort the first child.

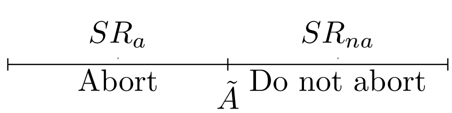


Figure 1.8. Case when first child was a boy

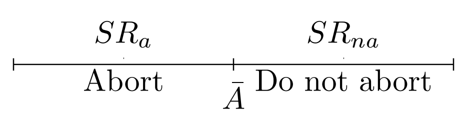


Figure 1.9. Case when first child was a girl

Putting together the cases when parents have already aborted and when they still have one abortion opportunity, we can derive the sex ratio for the second child:

$$s_i^2(m,b) = (1-p)G(\hat{A})SR_{na} + [1-(1-p)G(\hat{A})][(1-p)G(\tilde{A})SR_a + (1-p)(1-G(\tilde{A}))SR_{na} + pG(\bar{A})SR_a + p(1-G(\bar{A}))SR_{na}]$$

Substituting for the distribution of the cost of abortion, we can express $s_i^2(m,b)$ as:

$$s_i^2(m,b) = (1-p)\int_0^{\hat{A}} \frac{p}{1-p} f(\hat{A})dA + [1-(1-p)\int_0^{\hat{A}} f(\hat{A})dA][(1-p)\int_0^{\tilde{A}} \frac{p(2-p)}{(1-p)^2} f(\tilde{A})dA + (1-p)\int_{\tilde{A}}^{\infty} \frac{p}{1-p} f(\tilde{A})dA + p\int_0^{\bar{A}} \frac{p(2-p)}{(1-p)^2} f(\bar{A})dA + p\int_{\bar{A}}^{\infty} \frac{p}{1-p} f(\bar{A})dA]$$

1.4 Data

1.4.1 Albanian Demographic and Health Survey (ADHS)

We rely on data from the Demographic and Health Survey for Albania for the year 2008. This also coincides with the year in which the sex ratio at birth was the highest. The 2008-09 Albania Demographic and Health Survey (ADHS) is a nationally representative sample survey designed to provide information on population, demographic characteristics, household composition, and health issues in Albania for every household member.[7]

The survey is composed of questions answered by the female head of the family. It collects data on all household members, and it has very rich information on children's characteristics. The variables we are interested in measuring are *wife's bargaining power*, *relative preference for children* and *sex ratio*. The unit of observation is a cluster one, which are administrative unit divisions accurate to less than 15 meters.[8] ADHS 2008 includes 7584 female

[7]The ADHS was conducted by the National Institute of Statistics (INSTAT) and the Institute of Public Health (IPH), under the lead of the Ministry of Health of the Republic of Albania.

[8]The groupings of households that participated in the survey, known as clusters, are

respondents and 3013 men respondents of age 15-49. Survey fieldwork was conducted from October 2008 to April 2009. In the year 2008 Albania is an ethnically homogeneous country where all respondents are Albanian. The majority with approximately 55 percent, live in rural areas, with one-sixth of respondents in Urban Tirana, the capital. The average household size in Albania is 3.8 persons. Only 16 percent of households are headed by women.[9] We rely on this survey because it has very rich data on fertility, education, children information, gender and domestic violence, women's empowerment by accounting for gender attitudes, women's decision-making power, education and employment of men vs. women. The only downside of the survey is that it has data only on Albania in the Balkans region. As a result, we cannot compare it to other countries in the region.

1.4.2 Sex ratio of children

Sex ratio 1 is associated with the first-born child of the respondent, while sex ratio 2 is associated with the second-born child of the respondent. This ratio is about fertility in the past, children that respondents already had. We calculate them by dividing the total number of male children with the total number of female children for every cluster.[10] Their respective cluster average is 1.22 and 1.32. Since some clusters have very few children, we restrict our sample by considering only the upper 70th percentile of the number of children per cluster, which have a minimum of 16 children per cluster. We downsize to 214 clusters for our analysis. Table 1 shows a summary statistic of our variables of interest. The clusters have on average 23 children. With our new sample, the sex ratio for the first child is on average 1.18 with a standard deviation of 0.8. The sex ratio for the second child is on average 1.24 with a standard deviation

geo-referenced. Urban clusters contain a minimum of 0 and a maximum of 2 kilometers of error. Rural clusters contain a minimum of 0 and a maximum of 5 kilometers of positional error with a further 1% of the rural clusters displaced a minimum of 0 and a maximum of 10 kilometers.

[9] DHS Report on Albania 2008

[10] $sexratio1 = \frac{n_b^1}{n_g^1}$ and $sexratio2 = \frac{n_b^2}{n_g^2}$

of 0.95. The minimum for these sex ratios is 0.08 and it goes up to 4.5 and 7.5.

1.4.3 Relative preference for children

For the sex bias preference in children, we use a principal component analysis (PCA) of 3 measures at a cluster level: Time investment in children, education, and violence. PCA is a data reduction method to re-express multivariate data with fewer dimensions. This technique is used to uncover the common elements or to acquire fewer components for use in further analysis once surveys have been conducted. These measures are calculated as a ratio of female over male children for each cluster.

The first indicator, time investment in children, is composed of questions if parents read, name, count, draw, tell stories, sing songs, play, and take the child outside. These questions are answered for both female and male children by the mother, the father and also someone else in the family that might take care of the child.[11] The DHS children dataset has information on children up to the age of 5. There are 1616 children in this dataset. This variable is calculated by summing all activities that parents or someone else spends with the child. Then we sum all these activities for female versus male children, by accounting for the number of girls and boys for every cluster.

The second indicator is education. Considering that education is mandatory by law for children up to the age of 14, our variable of education is measured as the number of students enrolled in school for the age group 15-18 for every cluster, by taking into account the respective number of different sexes on each cluster.

The third indicator, violence, is measured by the number of female versus male children below 15 years who are abused physically by their parents.[12] By applying a principal component analysis on these 3 variables, we create the relative preference in children index represented by $\frac{x_g}{x_b}$. The scores are positive

[11]This is not specified and it can a grandparent, uncle/aunt, cousin or older sibling

[12]We get the same results even when we use children below 10. The reason we choose 15 years old is that on average it gets more difficult to abuse physically older children. The closer they get to 18, the more independent and powerful children become.

for time investment in children and education, and negative for violence. We transform this investment in girls index into taking values from 0.61 to 1, by assigning a value of 1 to the investment for boys ($x_b = 1$). Investment in children index is on average 0.8 at a cluster level with a standard deviation of 0.05. More about the construction of this variable and the PCA method can be found in Appendix A.2.

1.4.4 Women empowerment

We consider as women's empowerment the wife's bargaining power regarding household and health decision-making. It is expressed by θ in the model, which gives different weights to how utilities of parents and their respective cost of abortion, enter the utility of the family. The wife's bargaining power index is a principal component analysis on questions regarding final say for issues like own health care, making large household purchases, making purchases for daily needs, visits to family and relatives, and the decision on what to do with money husband earns. We build it by counting the total number of women per cluster that can take decisions on their own. We keep only respondents for whom we have data on all the variables we are interested. The index takes values from 0 to 0.53. A value of zero means that women have no power over decisions, while a value of 0.53 means that women are free to make decisions on their own or in cooperation with their partners. Women empowerment is on average 0.17 at the cluster level. Appendix A.2. shows a detailed construction of this variable.

Table 1.1. Summary statistics at the individual level of our sample

	Mean	Std. Dev.	Min	Max
Sex Ratio 1	1.18	0.80	0.09	4.5
Sex Ratio 2	1.24	0.95	0.08	7.5
Women empowerment θ	0.17	0.1	0	0.53
Investment in children x_g (when $x_b = 1$)	0.8	0.05	0.61	1
Number of children/cluster	23	4.6	16	40

1.5 Calibration

In order to do a counter-factual simulation analysis, we first do a calibration of the model, using the DHS data for the year 2008.

1.5.1 Identification

Some parameters are fixed ex-ante. We set γ, a scale parameter showing how parents enter utility of parents, equal to 1. When developing calculations, we see that γ does not affect the results. Moreover, we assign a value of 2 to r to show the extra cost burden for the mother. r shows that the cost of abortion for the father is only part of the cost of the abortion for the mother. Moreover, the value being greater than 1 is in line with the literature on abortion procedures. We assign the value 0.1 to the probability of stopping fertility q. This value is taken from the sample of DHS Data in 2008. It represents the share of couples that stop fertility after a first child. Then, we assign the value 0.52 to the probability of having a male child p. This is the biological probability to have a male child.

We proceed by finding m and b to match exactly the s^1 and s^2. m and b are the location and scale parameters of the logistic distribution of the cost of abortion. We solve the following system:

$$\begin{cases} s^1(m,b) = sexratio1 \\ s^2(m,b) = sexratio2 \end{cases}$$

The values we get from solving the system are $m = 0.05$ and $b = 0.00006$. A summary of the parameter values is provided in Table 2.

Over-identification:

We proceed with an over-identification strategy where m and b minimize the discrepancy across clusters between sex ratios in the data and the model. We minimize the difference between the model prediction and the observed sex ratios. We have 214 clusters in total, which leads to 428 targets, 2 parameters, and 426 over-identifying restrictions. To do so, we minimize the sum of the

Table 1.2. Parameters summary

	Meaning	Parameters	Target
γ	Parameter explaining how children enter utility of parents	1	normalization
r	Parameter showing the weight of the cost of abortion mother/father	2	a priori
x_b	Investment in the male child	1	normalization
q	Probability of stopping fertility	0.1	a priori
p	Probability of having a male child	0.52	a priori
m	Location parameter of logistic distribution	0.05	s^1
b	Scale parameter of logistic distribution	0.00006	s^2

squared errors for all clusters for both sex ratios, using several choice variables, to see whether the estimated parameter of interest remains the same or changes drastically. s_i^1 and s_i^2 depend from the location and scale parameters of the logistic distribution of the cost of abortion distribution m and b, investment in children post-partum for every cluster x_{g_i} (when $x_{m_i} = 1$), while the sexratio1 and sexratio2 represent the aggregate values of sex ratios for the first and the second child. To be considered that we set $r = 2$, $p = 0.52$, $q = 0.1$ and $\gamma = 1$. We first minimize the sum of the squared errors for all clusters with data on the investment in children, then introduce the women empowerment data.

$$\min_{m,b} \sum_i [s_i^1(m, b, x_{g_i}, \theta_i) - sexratio1) + (s_i^2(m, b, x_{g_i}, \theta_i) - sexratio2)]^2$$

We find the values of the location and scale parameter $m = 0.58$ and $b = 0.21$ by minimizing the sum of the squared errors for all clusters. In Appendix A.3 we show a plot of the values of m and $b$in the previous equation. There is a difference of 0.4 points between the observed and simulated values of sex ratios. The observed values of the sex ratios are 1.18 and 1.24, while the simulated are 1.22 and 1.20. 1.18 and 1.24 are different from the 1.12 value of the aggregate level, after limiting our data to clusters and individuals for which we can have information on all our variables of interest. In Figure 10 and 11 we plot the resulting fit for sex ratio 1 and 2.

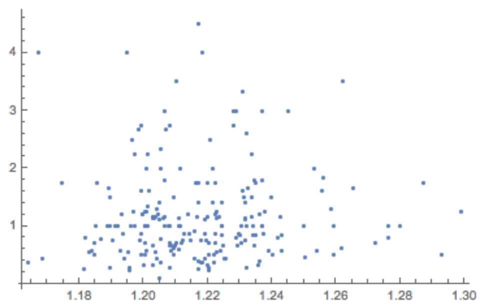


Figure 1.10. The fit of Sex Ratio 1

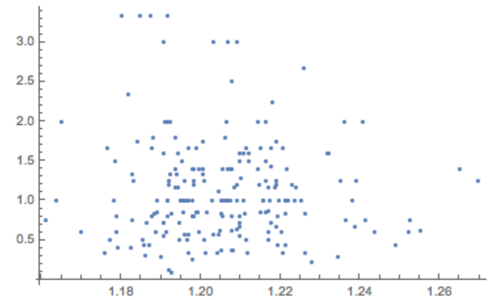


Figure 1.11. The fit of Sex Ratio 2

The next step consists in finding for what values of θ_i and x_{g_i} we can get normal values of Sex Ratio 1 and Sex Ratio 2. We find that θ_i and x_{g_i} are not powerful mechanisms in lowering sex ratios down to the biological values. There is still a residual of unobserved cultural traits that allows for this gap. We explain this simulation and residual in the next sessions.

Counterfactual simulation:
We continue with a counterfactual simulation of how sex ratio values change if the observed variables of investment in children and gender empowerment were gender equal and different from what we observe in the survey data prevailing in 2008. For instance, how do sex ratios change, when there is no gender gap in children and when parents invest the same in boys and girls.

We can see from the sex ratio equations that s_i^1 and s_i^2 depend on the values of the scale and location parameters, from the investment in children x_{g_i} and the gender empowerment θ_i. We start by assigning a gender equal value to x_{g_i}. Then we proceed by assigning a gender equal value to θ_i. Finally we assign gender equal values to both x_{g_i} and θ_i. Results are shown in Table 3. The first and the second sex ratio decrease for gender equal values of θ and x_g. The effect is stronger when we equalize preference in children x_g, than when we increase women empowerment θ. No children preference bias brings Sex Ratios 1 and 2 down to 1.16. No gender gap in bargaining power lowers Sex Ratio 1 and 2 down to 1.19 and 1.18, respectively. When we bring to gender equal values both θ and x_g, the sex ratios decrease at the same value as when we only adjust for x_g (1.16). This can be a result of cultural norms in the country where women themselves believe in the value of having a male child, which we express by the negative cost of abortion.

Table 1.3. Estimation of values for Sex Ratio for the first and the second child, average per cluster

	Sex Ratio 1	Sex Ratio 2	θ	x_g
observed	1.18	1.24		
simulated	1.22	1.20	0.16	0.8
no children preference bias	1.16	1.16	θ_i	1
no gender gap in bargaining power	1.19	1.18	0.5	x_{g_i}
gender equality for both children and adults	1.16	1.16	0.5	1
biological values	1.05	1.05		

1.5.2 Discussion of the Results

In our model we consider three reasons to abort: i) differentials in children's investment, x_g and x_b, according to parents' preferences, ii) differentials in bargaining power (θ) among partners, and iii) $G(0) > 0$, interpreted as resulting from unobservable cultural norms. We can discuss the following results regarding their effects on the sex ratios at birth:

i) from our data on ADHS having no post-natal children preference bias, can lower sex ratios at birth to 1.16. This is a biased value considering that the biological one is 1.05. As a result, although this mechanism can decrease the sex ratio values, it is not successful in decreasing them down to their biological value.

ii) when there is no gender gap in bargaining power among partners, sex ratios decrease to 1.19 and 1.18. These two channels can lower sex ratios at birth and decrease sex-selective abortion. Nevertheless, the values remain still above the biological value of 1.05, which means that there is still a presence of sex-selective abortion. When parents invest the same in boys and girls after birth, meaning when $x_g \to x_b$, sex ratio at birth is still biased (the limit of sex ratio 1 does not reach the biological value $s^1 \nrightarrow \frac{p}{1-p}$). This is because some parents have a negative cost of abortion which makes them abort also in cases when they have the same investment in children. The biased sex ratio can be written as $sr^1 = \frac{(1-p-\frac{1}{1+e^{\frac{m}{b}}})p}{(1-p)^2} = \frac{(1-p-G(0))p}{(1-p)^2}$.

iii) with equal bargaining power among parents ($\theta = 0.5$) and post-natal equal investment in children $x_g \rightarrow x_b$, sex ratios decrease at 1.16, which is again above biological value. This bias remains as people continue to abort girls. Pre-natal preferences of parents regarding the sex of their children make $G(0) > 0$. Our assumption regarding abortion of children is that women do not abort boys, making $G()$ the cost of aborting girls as well as capturing the residual in this model. It is important to distinguish between preferences in children before and after they are born. What x_b and x_g capture are investments and preferences in children post-partum while the cultural norms capture the preference pre-natal. This residual not explained from the model can capture the preference of children pre-birth.

In sum, cultural norms considering the ideology and conviction that male children are a blessing are represented in the model as having a negative cost of abortion. Parents desire a male child so much, that not only they do not mind the financial, emotional, and physical costs of abortion but also see it as an investment that has high returns in terms of cultural norms. This is translated into a higher utility for the family. The fact that investment in children is more powerful than women's empowerment in lowering sex ratios at birth can be in the fact that women themselves prefer males. Failure to give birth to a son may lead to consequences that include violence or rejection by the marital family. As a result, a deeply rooted preference for male children has been created leading to women choosing male children over female ones and reinforcing this norm furthermore. We use these results for Section 6, we discuss a possible interpretation of cultural norms considering historical data and the Albanian patriarchal context. Moreover, we discuss policies that account for gender equality while targeting biological sex ratios at birth.

1.6 Discussion

In this partial equilibrium theoretical framework, we analyze if equal investment in children and women empowerment can explain biased sex ratios of children. The sex ratio remains distorted, even when equalizing investment in boys and girls and when increasing women empowerment. There is still a residual that can be interpreted as social norms. It is important to discuss this residual and problematic due to the negative externalities it can have on the marriage market, the housing market, and women status in the society. Although a shortage of females might benefit them in the marriage market by marrying up, their intra-household bargaining power remains at a disadvantage. Other externalities include the increase in the likelihood of coerced marriages, bride abduction, and child marriage. These are strong reasons to focus on possible drivers and draft the correct policies and internalize these externalities.

Considering that this phenomenon comes from strong son preference, we focus on social norms. The fact that there is some sex selection in countries with an Albanian minority leads credence to the idea that this is cultural. Social norms can have roots in events that have happened in the past. This can be financial decisions, in times when having a son was an insurance for old age dependency, the continuation of family line, or just a preference for having a son in a patriarchal society where women do not have a say equal to men. Following these elements, we look at sex ratios in Albania historically. This allows for a better understanding of the influence that norms can have on the decision-making process and fertility decisions of parents.

1.6.1 Historical Background

Historically, families have chosen to take care, appreciate or neglect their children based on their sex. There is evidence that points to relatively high sex ratios at birth in European historical populations by looking at parish registers of vital statistics (Tapia (2019)). We look at Mosaic data that provides historically integrated census microdata for Albania in 1918. It only covers the northern and central part of present-day Albania, areas occupied by the

Austro-Hungarian Empire during World War I. In 1918, the first census has been carried out in the occupied territories. The census considers the household size, religion, literation, region, and age. It has 140611 observations and 27792 households. We are interested in analyzing the sex ratios at birth. For this, we try samples for different ages. We consider a sample for all children, one where children are below 20 years of age, another one where they are below 10 or below 5. In all these sample analyses, the sex ratios are skewed. Table 4 summarizes these values and it considers also sex ratios values that account for regional differences, urbanization, and religion. Values of the sex ratios in 1918 vary around 110 male children per 100 female children. These are high values considering that ultrasound technology was not available at the time and we cannot speak of sex-selective abortion. Problems of under-registration of female children may inflate sex ratios by under-enumerating girls. Postnatal care differentiation with lack of care due to gender discrimination that leads to a higher mortality can also be an explanation of the high SRB. Appendix A.7 explains more about sex ratios historically both in Europe and Albania.

Table 1.4. Sex Ratios in Albania 1918 *Source: Mosaic Census*

SR Measure	SR value <5	SR value <10	SR value <15
Aggregate	1.10	1.09	1.09
Urban	1.11	1.12	1.13
Rural	1.09	1.06	1.06
Regional (7 obs)	1.10	1.10	1.11
Provincial (274 obs)	1.24	1.18	1.16
Religion			
Muslim	1.11	1.10	1.34
Catholic	0.93	1.04	1.21
Roman Catholic	0.85	1.05	0.99
Orthodox	0.94	0.95	1.32

Considering that there are different reasons behind son preference and biased sex ratios at birth, we focus on religion. 76 % of the census population is registered as Muslim, 9.7 % as Orthodox and 9.8 % as Catholic, and 4.1 % as Roman Catholics. Among the others are Greek Catholic, Jewish, and Lutheran. The Muslim population has higher sex ratios than the orthodox and roman catholic ones, which are more in favor of female children. This might be explained by the financial wellbeing of the Roman Catholic population and the possibility to register their children. To be considered that during the Ottoman Empire, part of the population that did not convert to Islam had to pay a tax. Those that could afford to pay the tax could continue belonging to and practicing other religions. Another difference among religions is that the sex ratio increases with age for both Muslims and Catholics. This shows that the sex ratios are even more problematic when you consider more children from each household. This might also be that female children have left the household when they get married and migrated to areas that are not part of the sample. Other spatial differentials can be seen in terms of urbanization. The sex ratio in urban areas is higher than that in rural areas. The difference between the two grows even larger when age increases. Migration can be an explanation. There is no infant mortality data for this time period thus we cannot assume about neonatal discrimination and excess female mortality in childhood.

1.6.2 Policy Proposals

Negative externalities that sex ratios have on the society can be managed by developing the right policies. In order to lower sex ratios down to biological values we propose policies and directives which use as instruments the two channels that we explored in the theoretical framework: investment in children, women empowerment, cost of abortion and the cultural norm residual. The first policy we highlight is thus focused on eradicating the preference bias towards children of different sexes. Then we discuss why cultural norms are important and what can be the difficulties of a policy that accounts for them.

We continue with a policy that accounts equal bargaining power in decision making. A fourth one considers the increase of the cost of abortion. These policies are specific for Albania, since our results are based on the Demographic and Health Survey data of Albania. Nevertheless, they may also be beneficial for countries with large Albanian minority or even immigrant communities following evidence of transmission of norms when it comes to sex selection.[13]

In our first policy regarding investment in children we discuss how this post-birth preference of children affects people's decisions and why accounting for it is important. By relying on how we measure preference in children from the Demographic and Health Survey dataset, a proposal would be about equal treatment regarding time spent with each child and the same investment in education. Time spent with both genders is a key factor in determining the gender gap postpartum. Incentives in education are regarding achieving equality of genders in non-mandatory education. A possible strategy would be increasing awareness of gender-based violence like in South Korea, on legal and administrative protection and support services for victims of violence and abusers.[14] From the results that we get in the calibration part, an equal investment in children managed to lower sex ratios at birth but not down to the biological values. This is because we account for post-partum preferences. What could be important is to capture the preference of parents pre-partum and consider a policy that pushes for equal preferences in the expected child, irrespective of the future sex. This translates into a policy that fights cultural norms accounting for son preference. Creating a framework of how to target

[13] Closing the gender gap and advancing gender balance, promoting gender equality, ending gender based violence has been at the top of the priorities in the country by many institutions like the European Union, United Nations, World Bank.

[14] With regards to the prevention and reduction of gender-based violence and domestic violence, despite improvements especially in the legal framework in compliance with international and regional standards, Albania has to take a series of other actions to meet the requirements of ratified international conventions and to show no tolerance to violence. This is to be interpreted with caution considering the applicability of the law. Societal pressure keeps many women from making formal complaints. Moreover, in many cases, the prosecutor simply fails to prosecute a man for abusing his wife. Apart from the social, there are still legal barriers when reporting domestic abuse to the courts

and account for cultural norms is rather complex. This is because cultural transmission is strong. South Korea can be considered a successful country model in fighting the problematic of sex-selective abortion. It appears that a combination of factors contributed to this shift. Two decades of exceptional economic growth led to fundamental changes in Korean society with a shift away from a farm-based economy, increasing desire for small families, increasing urbanization, greater participation of women in the workforce with better employment opportunities, and parents having retirement savings for old age (Das-Gupta et al. (2003)). Promotion of women's rights and responsibilities within their birth family especially after marriage and recognition of women-headed households, were seen to be beneficial, as was a 'Love Your Daughter' media campaign (UNFPA).

We continue with the gender empowerment policy. During the 19th and the 20th centuries women in Albania lived in a patriarchal and traditional society.[15] Women today account for 24 percent of the seats in the Parliament and represent 42 percent of the Governmental cabinet. Although there is an increased representation of women in the public administration to 45 percent, the situation remains problematic at a micro household level regarding social norms. This affects furthermore people's behavior in terms of fertility decisions regarding the sex of their children. Considering this progress, for gender empowerment, the policy aims to empower furthermore females by increasing their participation in decision-making. Considering the heterogeneity of this measure in the data, this can be achieved by targeting the more problematic areas, ensuring the economic and social empowerment of women, addressing gender inequalities that lead to poverty and promoting social inclusion.[16]. Albania is a signatory to several important and binding international documents, which guarantee the equality of men and women and prohibit gender-based discrimination but looking at the data there is an obvious gender bias. [17]

[15]More about gender empowerment in Albania during this time period on Appendix A.6.

[16]The level of participation in the labor force for the male and female population of 15-64 years of age is respectively 72.2 % and 51.3 % (Males and Females in 2015, INSTAT)

[17]Gender equality has a prominent place in the agenda of the institutions, yet despite the

The above policies become less successful instruments in lowering sex ratios at birth when accounting for cultural norms. It is not easy to "fight" norms. They are defined as specific but tacit standards of what is socially and individually acceptable. One can argue that the strength of the rule of law is a mechanism for lowering sex ratios. Nevertheless people find a way. Many immigrants from countries where sex selection techniques are present, go back to their countries to undermine the procedure, considering the strength in the application of laws in developed countries. Sex selection procedure becomes a problematic in countries with strong cultural norms in son preference and not strong judicial systems. The goal of gender equality and the practice of gender mainstreaming focuses on how females and males experience problems in society differently, and how they relate to the societal forces that shape power relationships. It aims to identify the societal behaviors and structures that sustain gender inequality and make changes that are institutional and systemic. In a country where traditional patriarchal attitudes are still prominent, gender inequalities are present in all spheres of social and economic life, and violence against women is still widespread. To help fight this problematic, a stronger legal framework concerning the rights of women, for combating gender-based violence, as well as the adoption of legal enforcement for abortion after 12 weeks of pregnancy. An increase in the cost of abortion would affect also people that abort before the 12th week for other reasons or those that abort after 12 weeks for medical reasons.

1.7 Conclusion

We study the problematic of sex-selective abortion and consider a partial equilibrium theoretical framework where we try to lower sex ratios at birth by accounting for the two channels of son preference: women empowerment and investment in children. For this, we develop a model where the decision

progress that has been made to establish a legal and policy framework for the advancement of women's rights and gender equality, inequalities are pervasive. The country has restricted the use of modern technologies for sex-selection purposes. However, such prohibitions are often difficult to enforce, and they drive demand for these technologies underground.

to abort a female child depends on the sex bias in child preferences and on the gender gap in intra-household bargaining power among parents. We calibrated the model using the data of the Demographic and Health Survey in Albania for the year 2008 and we do a sensitivity analysis on the sex ratios at birth by changing the values of preference bias in children and women empowerment.

We find that equal preference in children lowers the sex ratio at birth from 1.22 to 1.16. Son preference is still deeply rooted in the decisions of the family's formation, which leads to the gradual buildup of a male surplus. Raising women's empowerment lowers the sex ratio, but less than equalizing preference in children. Sex Ratios decline from 1.22 to 1.19. An interpretation of this result can be that since women are often under immense family and societal pressure to produce sons, women themselves have a preference for male children. Failure to give birth to a son may lead to consequences that include violence or rejection by the marital family. As a result, a deeply rooted son preference has been created leading to women choosing male children over female ones and as a result further reinforcing this norm. When we consider both gender equality in decision making and no bias in children's preference, the sex ratio drops to 1.16. Our findings show that women's empowerment and investment in children help decrease the biased sex ratios but are not the main channels of this problematic. We propose policy recommendations on what our data and model tell us about preference bias in children and gender empowerment. Equal access to education, the time investment in the early years of life, and accounting for cultural norms are some of the policies that help lower skewed sex ratios at birth. This can be associated with other supportive measures like access to information, health care services, personal security, and the strengthening of the legislature. Future avenues for research can be dedicated to models with 2 abortion possibilities, increasing the number of children per family, increasing the utility cost of abortion, or the study of cultural norm transmission regarding fertility decisions within generations or between different Albanian communities in their country or abroad.

1.8 Appendix

1.8.1 Proofs

Proof 1: First child is a male, abortion of the second child The child will belong to the male sex with probability $1 - p$. In the model we assume that in this case, the pregnancy will always continue. There will be a second pregnancy with the same probabilities of belonging to the male and female sex. If in the second pregnancy, the fetus happens to be a female, then the family will evaluate the utility of continuing the pregnancy and the expected utility of abortion. In this last one it will be considered a third pregnancy with possible outcomes.
$U(female) = \ln(c) + \gamma \ln(x_b + x_g)$
$EU(abortion) = pU(male) + (1 - p)U(female)$
$EU(abortion) = p(\ln(c) + \gamma \ln(2x_b) - A) + (1 - p)(\ln(c) + \gamma \ln(x_b + x_g) - A)$
$EU(abortion) = \ln(c) + p\gamma \ln(2x_b) + (1 - p)\gamma \ln(x_b + x_g) - A$
The pregnancy will be interrupted when $EU(abortion) > U(G)$
This happens when $p\gamma \ln(2x_b) - p\gamma \ln(x_b + x_g) - A > 0$ and it is the second child that belongs to the female sex

$$A < \bar{A} = p\gamma \ln(\tfrac{2x_b}{x_b + x_g})$$

Proof 2: First child is a female. The decision is about aborting the first child that belongs to the female sex. The family has to evaluate the utility of keeping the child or terminating the pregnancy. The utility of having a female child is
$U(female) =$
$q[\ln(c) + \gamma \ln(x_g)] + (1 - q)[p(\ln(c) + \gamma \ln(x_b + x_g) + (1 - p)[\ln(c) + \gamma \ln(2x_g)]$
This is to be compared with the expected utility of aborting the child
$EU(abortion) = pU(B) + (1 - p)U(G)$

$$U(G) =$$
$$q[\ln(c) + \gamma \ln(x_g)] + (1-q)[p(\ln(c) + \gamma \ln(x_b + x_g) + (1-p)(\ln(c) + \gamma \ln(2x_g)]$$
$$U(B) = q[\ln(c)+\gamma \ln(x_b)]+(1-q)[p \ln(c)+\gamma \ln(2x_b)+(1-p)(\ln(c)+\gamma \ln(x_b+x_g)]$$

The abortion will be interrupted only when

$$EU(abortion) - U(female) > 0$$
$$EU(abortion) - [U(female) + A - A] > 0$$
$$pU(B) - pU(G) - A > 0$$
$$A < \hat{A} = p[U(B) - U(G)]$$

We can write

$$U(B) - U(G) = q\gamma \ln(x_b) + (1-q)p\gamma \ln(2x_b) - q\gamma \ln(x_g) - (1-q)p\gamma ln(x_b + x_g) + (1-q)(1-p)\gamma \ln(x_b + x_g) - (1-q)(1-p)\gamma \ln(2x_g)$$

Rearranging and writing $g = \frac{x_g}{x_b}$ we can rewrite

$$U(B) - U(G) = -q\gamma \ln(g) - (1-q)p\gamma \ln(\tfrac{1+g}{2}) - (1-q)(1-p)\gamma \ln(\tfrac{2}{1+\frac{1}{g}})$$

As a result $A < \hat{A} = p(-q\gamma \ln(g) - (1-q)p\gamma \ln(\frac{1+g}{2}) - (1-q)(1-p)\gamma \ln(\frac{2}{1+\frac{1}{g}}))$

Consider that the above can also be written as

$$A < \hat{A} = p(q\gamma \ln(\tfrac{1}{g}) + (1-q)p\gamma \ln(\tfrac{2}{1+g}) + (1-q)(1-p)\gamma \ln(\tfrac{1+\frac{1}{g}}{2}))$$

Proof 3: First child a female, abortion of the second child The child will be female with probability $1-p$. The pregnancy will continue only when $EU(abortion) < U(female)$. After the child is born there is a second pregnancy. A girl will be born with probability $1-p$ and a boy with probability p. We assume that if the child belongs to the male sex, pregnancy will continue. If the child belongs to the female sex, the pregnancy will continue only when $EU(abortion) < U(female)$. Consider that in case of abortion, another pregnancy will follow up with the probability p of having a boy and $1-p$ of having a girl.

$U(female) = \ln(c) + \gamma \ln(2x_g)$
$EU(abortion) = pU(male) + (1-p)U(female)$
$EU(abortion) = p(\ln(c) + \gamma \ln(x_g + x_b) - A) + (1-p)(\ln(c) + \gamma \ln(2x_g) - A)$
$EU(abortion) = \ln(c) + p\gamma \ln(x_g + x_b) + (1-p)\gamma \ln(2x_g) - A$
Comparing the $U(female)$ with $EU(abortion)$, the pregnancy will be interrupted when the later is greater than the first.
This happens only when:

$$A < \tilde{A} = p\gamma \ln(\tfrac{x_b + x_g}{2x_g})$$

Proposition 1: A second child who is female is more often aborted when first child is female, iff boys are preferred to girls: Since $\tilde{A} = p\gamma \ln(\frac{x_b+x_g}{2x_g})$ and $\bar{A} = p\gamma \ln(\frac{2x_b}{x_b+x_g})$ we can rewrite and simplify to $\frac{\tilde{A}}{\bar{A}} = \frac{\ln(2g)-\ln(1+g)}{\ln(\frac{1+g}{2})}$. Then $\tilde{A} > \bar{A}$ iff $x_b > x_g$ meaning $\frac{x_b}{x_g} < 1$, $x_b > 0$

Proposition 2: If families have two children, first girls are more often aborted than second girls. Since $\hat{A} = p(q\gamma \ln(\frac{1}{g}) + (1-q)p\gamma \ln(\frac{2}{1+g}) + (1-q)(1-p)\gamma \ln(\frac{1+\frac{1}{g}}{2}))$ and $\tilde{A} = p\gamma \ln(\frac{x_b+x_g}{2x_g})$ then $\frac{\hat{A}}{\tilde{A}} = \frac{\ln(g)}{2ln(2g)-2ln(1+g)}$. Ratio is always less than one, which means that $\hat{A} < \tilde{A}$

1.8.2 Principal Components Analysis

Principal Component Analysis is a data reduction methods to re-express multivariate data with fewer dimensions. We apply this method to acquire fewer elements when constructions the variable of gender empowerment and relative preference for children. For the construction of gender empowerment variable we use questions of the survey about the freedom of decision making of the female partner.

$\theta = dummy1 * 0.5375 + dummy2 * 0.1459 + dummy3 * 0.1227 + dummy4 * 0.1007 + dummy5 * 0.0932$

In the above equation dummy1 is the construction of a dummy variable from the survey question 'final say on own health care'. The variable takes value 0 when respondent does not have the final say on her own health and it takes value 1 when respondent has the final say alone or with in cooperation with her partner/husband. Following the same logic, dummy2 and stands for 'final say on making large household purchases', dummy3 for 'final say on making household purchases for daily needs', dummy4 for 'final say on visits to family or relatives' and dummy5 for 'final say on food to be cooked each day'. The highest proportion goes to the variable constructed from the question of taking decisions on her own health care and the rest is divided among the other dummy variables.

For the construction of relative preference in children, we do a PCA of three indicators:
i) Time investment in children below 5. This is composed by time spent reading, telling stories, singing, spending time outdoors, playing and organizing activities by both parents and someone else. We create dummy variables when one of the adults spend time on kids. We create a new variable composed by the sum of the time spend by each adult and divided by 6, which is the number of the total activities. We create this variable for both male and female children. We sum it up for each cluster respectively for boys and girls, by accounting for the different number of boys and girls in every cluster. Then the time variable is constructed by dividing the previous time cluster variables of girls over boys.
ii) Education of children aged 15-18. We calculate education differentials of male and female children. We calculate education years for both females and males for every cluster. The education variable is constructed by dividing the education of females over males for every cluster
iii) Violence towards children below 15. We create a violence differential variable by creating dummies for each female and male child whom have

undergone violence. We sum up the number of girls and boys for each cluster. The variable of violence is created by dividing the number of females of the number of males.
Considering the heterogeneity of this data we take the logarithms of each of them. Appying the PCA method on these 3 indicators leads to the following proportions that we use in the construction of $\frac{x_g}{x_b}$ variable.

$$\frac{x_g}{x_b} = 0.3621 * log_{time} + 0.3346 * log_{education} + 0.3033 * log_{violence}$$

The proportions are shared almost equally among time, education and violence variables. We normalize the new variable to an index that takes value zero to one.

1.8.3 Calibration and Identification

In a first step we find the values for the parameters m and b of the logistic distribution for both SR1 and SR2 by solving the system. In a second step we finding parameters of the logistic distribution - over identification - by minimizing sum of squared errors for all clusters. In Figure 12 to 14 we show fragments of the same 3D plot of m and b captured on different angles.

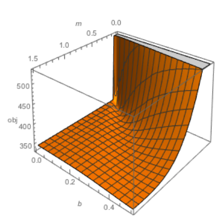


Figure 1.12

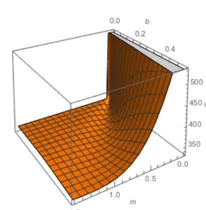


Figure 1.13

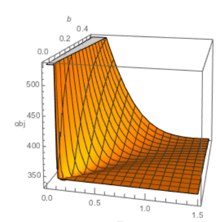


Figure 1.14

1.8.4 Gender bias worldwide

Countries with a strong son preference are mainly South Asian, East Asian, Central Asian, and Balkan countries (Guilmoto (2015)). Table 5 shows sex ratios at birth during 2008-2014 in a country aggregate level for both problematic countries and countries with biological normal values. A counter-example of high sex ratios at birth is sex ratios below the value of 1. An example of this is Minangkabau in Indonesia. This region has a preference for daughters. After marriage, the new couple moves to the house of the wife. A Minangkabau who gives birth to a boy will, on average, have a second child more quickly than one who gives birth to a girl.

1.8.5 Demographic indicators and gender bias in Albania

- Figure 15 and 16 show Sex Ratios at birth and TFR in Albania.[18] A low TFR has been observed in countries with skewed sex ratios. When fertility drops, sex selection is a technique to control for the preferred fertility.

- Figure 17 shows data on total births and declared abortion during 1994-2017. Figure 18 shows declared abortion per 1000 birth for the same years.

- Figure 19 shows infant female and male mortality. There is a convergence of the two throughout the years. Infant mortality is measured as the total number of male infants under 1 year old that have died divided by the total number of male children born that year. Since males outnumber females at birth, the decrease in male infant mortality leads to an increase in the sex ratios at a higher age, if the female infant mortality remains constant. In 2018 they are equal, leading to a higher number of males compare to female babies.

[18]Source: Albanian Institute of Statistics INSTAT

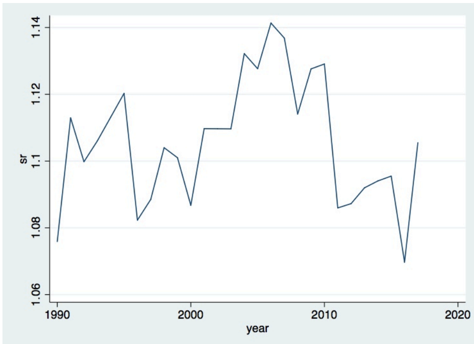


Figure 1.15. Sex ratio at birth

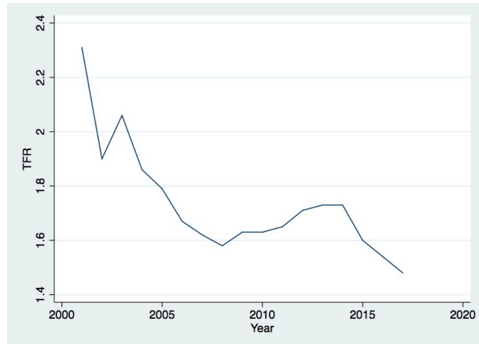


Figure 1.16. Total fertility rate

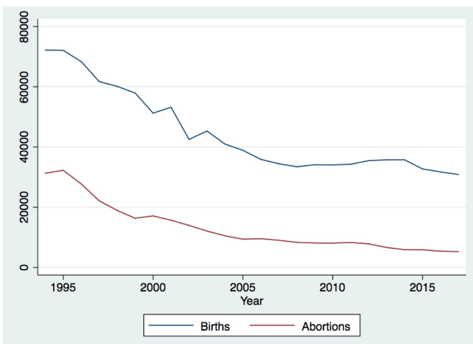


Figure 1.17. Births and Declared Abortions 1994-2017

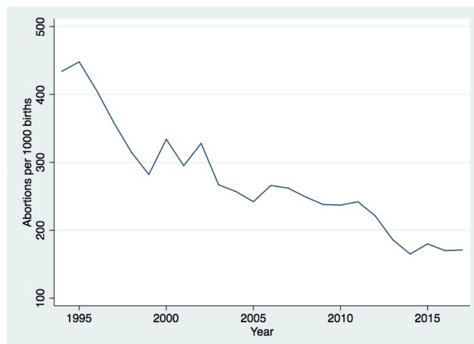


Figure 1.18. Declared Abortions per 1000 births

Figure 1.19. Mortality of male and female children under 1 year old

Table 1.5. Sex Ratio at birth 2008-2014. *Source: Guilmoto (2015)*

Country	SRB	Year	Source
China	115.9	2014	National Bureau of Statistics
South Korea	105.3	2013	Birth Registration
Hong Kong	109.3	2013	Birth Registration
India	110.0	2011-2013	Sample registration System
India	110.1	2012	Birth Registration
Singapore	107.0	2013	Birth Registration
Taiwan	107.4	2012	Birth Registration
Vietnam	112.2	2013-2014	Intercensal Survey
Azerbaijan	115.6	2013	Birth Registration
Armenia	114.0	2012-2013	Birth Registration
Georgia	108.0	2012-2015	Birth Registration
Albania	109.0	2012-2013	Birth Registration
Kosovo	110.4	2011-2013	Birth Registration
Macedonia (northwest)	110.4	2009-2013	Birth Registration
Germany	105.3	2013	Birth Registration
Brazil	104.8	2012	Birth Registration
France	104.8	2013	Birth Registration
USA	104.7	2012	Birth Registration
Japan	105.2	2012	Birth Registration
Russia	105.7	2013	Birth Registration
Turkey	105.8	2013	Birth Registration

1.8.6 Gender empowerment in Albania

There are still considerable differences between men and women in terms of their access to education, prenatal care, economic possibilities, and social standing despite the years of improvement. Albania's northern regions are seen to be more traditional, conservative, and patriarchal.[19] The Albanian

[19] de Haan *et al* (2006).*Biographical Dictionary of Women's Movements and Feminisms in Central, Eastern, and South Eastern Europe: 19th and 20th Centuries*

greeting "May it be a son" is customary due to the nation's predilection for sons. Women who follow these customs are under pressure to have boys. If a woman has daughters rather than males, she can possibly experience abuse or desertion. This may have its origins in the 500-year-old Kanun of Leke Dukagjini, a traditional code of conduct, and in which women's primary responsibilities are to care for their children and their homes. Females who were widowed or who desired equal rights with men had the option of swearing virginity in order to get full male privileges in return for their gender. By "taking an oath of virginity," these women may assume the position of the male head of the household, which would provide them the same privileges as males. This takes into account the freedom to behave like a man, to carry weapons, to own property, to move around freely, to dress like a man, to take on a man's name if they so choose, to assert their independence, to avoid being forced into marriage, and to associate with men while being treated as a man.

Under Albania's communist administration, gender equality advanced. They began encouraging equal roles and responsibilities for men and women, although this didn't end until the change of administrations in 1990. Women had less representation in the parliament than males in the first democratic election after the fall of communism. In the previous parliament of communist Albania, there were 75 female lawmakers. Today, there are just 9. The communist government carried out an effective literacy drive that reached even the most isolated areas. Although precise statistics are lacking, the majority of Albania's pre-war population lived in rural areas, where illiteracy was widespread. By the late 1980s, more than 90% of people were literate. Greater access to educational options helped women. Women were expected to maintain their function in the home while maintaining their loyalty to the party, just like men were. Due to the lack of birth control and the high fertility rate of 2.3 in the 1980s, women were expected to have children. Women were left to deal with the economic suffering brought on by the shift from a command economy to capitalism after communism.[20]

[20] Human Rights Watch 1996 *Human Rights in Post Communist Albania* and Suad Joseph; Afsana Nagmabadi (2003). *Encyclopedia of Women and Islamic Cultures: Family, Law and Politics*

1.8.7 Sex Ratios Historically

There is an increasing interest in the literature on historical sex ratios. Some of the reasons behind these historical imbalances in sex ratios are urbanization, conflict, and the economic value of children in an agrarian society. European society during the 16th to the 19th centuries has been characterized by the intensive disadvantaging of girls leading to the actual death of female infants and children. Industrialization and economic modernization certainly produced greater participation of women in the labor force but that did not change the disadvantages of females (Lynch (2011)).

Some interesting findings come from Minasyan and Mavisakalyan. They study if a societal preference for male births may be reinforced at the individual level due to gender-biased perceptions about group defenders such that fear of war and if this is associated with son preference. They rely on baptism data, where male baptisms, outnumber females. Another study is from Beltran-Tapia and Marco-Garcia. They show that baptism records exhibited exceptionally high sex ratios at birth, especially during the 19th century. On the other hand, having no previous male siblings increased both the probability of male baptisms and female mortality during the first day of life. Additionally, Beltran-Tapia and Raftakis argue that son preference resulted in gender discriminatory practices that unduly increased female mortality rates in infancy and childhood in Greece during the late 19th and early 20th centuries. In terms of spatial studies, Espana-Eljaiek and Fuentes-Vasquez show a structural deficit for the girl-boy ratios in the rural areas, especially during the first half of the 20th century, by considering a higher economic value of boys in agrarian economic activities. This raise in the interest in historical sex ratios comes from better historical data access. The importance and contribution of this studies are on disentangling the cultural, financial, and religion drivers behind the biased sex ratios. [21]

[21] References for this section are working papers from the Sex Ratios Seminar Series:
i) Minasyan and Mavisakalyan, *"The Role of Conflict in Sex Discrimination: The Case of Missing Girls"*
ii) Beltran-Tapia and Marco-Garcia Death, *"Sex and fertility: Female infanticide in rural Spain, 1750-1950"*

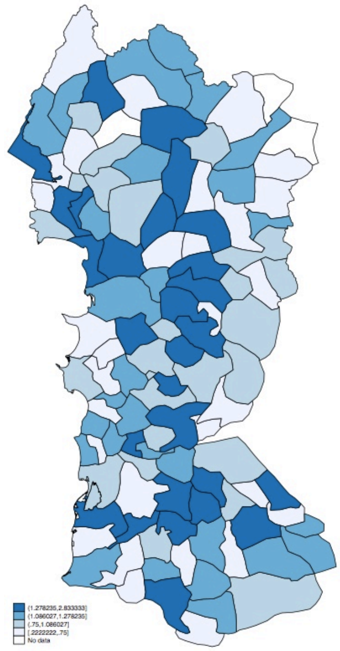


Figure 1.20. Sex Ratio in Albania in 1918, Source of data: Mosaic Cencus

In Figure 20, the map of Albania under the Austro Hungarian Regime, with the biological (normal values, below 1.05) sex ratio in light blue and the distorted sex ratio (above 1.05) in dark blue. The dark blue colored provinces are in line with the mountainous map of Albania, which is also known for being more patriarchal (they used their own set of laws called Kanun). Sex selection techniques were not present in the day, but neglecting to register a child was common. Under registration of children was the main cause of biased sex ratios at birth. There was no echography and sex detection in 1918 so sex-selective abortion is excluded as a possible reason but this bias. Difficult infrastructure in terms of mobility, patriarchal norms, children preference can be some of the reasons behind the under-registration of female children in these areas.

---

iii) Beltran-Tapia and Raftakis, *"All little girls, the bad luck! Sex ratios and gender discrimination in 19th and early 20th century Greece"*

iv) Barbiera, Castiglioni, and Dalla Zuanna, *"Missing Women in the Italian Middle Ages? Data and Interpretation"*

v) Espana-Eljaiek and Fuentes-Vasquez, *"The Economic Value of Girls: Urban and rural gaps in sex child ratio during 20th century"*

2

Internal Migration as a Response to Soil Degradation: Evidence from Malawi [1]

Co-authored with Stefanija Veljanoska

[1] Authors would like to thank Frédérick Docquier, Jean-Francois Maystadt, Giuseppe Maggio, European Historical Demography Society (Madrid), NCCR on the move (Université de Neuchatel), University of Namur, UCLouvain seminar for great comments and discussions.

Abstract
We study how the slow deterioration of soil, caused by climate change, affects internal migration and household resettlement. Rural households are expected to move when they face worsening soil conditions, as soil degradation is detrimental to agricultural productivity. The other possibility is that they can get stuck in a poverty trap. We use the Integrated Household Survey in Malawi for the years 2010-2016. Soil depletion is not a random process and to account for its endogeneity, we instrument soil degradation by using distant climate shocks and controlling for recent weather conditions. We find that severe soil nutrient constraints push households to send their members away. The underlying mechanism is that soil degradation is harmful to agricultural productivity, and therefore food security, which incentivizes households to seek better opportunities by pushing their members to migrate.

"*Men argue. Nature acts.*"
- Voltaire

2.1 Introduction

According to the UNHCR, climate refugees account for more than half of all migrants.[2] A category of these migrants are the environmentally motivated one.[3] These are people who leave a constantly deteriorating environment to pre-empt the worst effects due to environmental stressors. Myers (1997)

[2]UNHCR's Refugee Population Statistics Database

[3]According to "Internal Displacement Monitoring Center", weather-related disasters already force an average of 21.8 million people to flee their homes every year, without including displacements associated with slow-onset disasters such as drought and environmental degradation. Storms caused 12.9 million displacements worldwide 55 per cent of all weather-related disasters - by triggering mass displacement of populations living in vulnerable areas.

describes environmental refugees as 'individuals who no longer gain a secure livelihood in their traditional homelands because of what are primarily environmental factors of unusual scope'. There are at least 20 million environmental refugees worldwide, more than those displaced by war and political repression combined. Land degradation is one of the aspects of environmental degradation. Between 2010 and 2015, 20 percent of the worldwide land was degraded with Sub-Saharan Africa experiencing land degradation above the Globe's average.[4] In 2000, 1.33 billion people worldwide were located on degrading agricultural lands, of which 1.26 billion were in developing countries and this impacts significantly future poverty in developing countries (Barbier and Hochard, 2016).

The profound relationship between migration and the environment is not a new phenomenon, but the acceleration of climate change introduces more complexity to this relationship. Gradual environmental change, slow-onset, and sudden-onset natural disasters influence population migration decisions and patterns. Even though sudden-onset natural disasters are more likely to induce mass displacement, still a large share of people is estimated to migrate because of gradual degradation of the environment and a depletion of resources. Slow-onset disasters and gradual environmental deterioration such as reduction of soil fertility, desertification, coastal erosion, and sea-level rise, which may be associated with climate change, can have serious implications for existing livelihood patterns and production systems and may in turn induce different types of migration.

The complexity of this environment-migration nexus is subject to a fast-growing research (Millock (2015) does a summary of the literature review of this nexus). This literature focuses mainly on the climate change impact on migration decisions and finds mixed evidence (for an exhaustive review, see Section 2.2). The study of the incidence of slow degradation of land on migration received little attention from researchers from different fields (an exception are (Gray, 2009; Call and Gray, 2020), see Section 2.2). The scarcity of research on this relationship is probably related to the i) lack of granular

[4] UNSTATS

data on the soil quality and ii) that its degradation is not a random process iii) since land degradation can happen as a slow process, the available datasets provide a downsized sample of migrants. This means that what is available in the dataset, is a sample of people that are left behind after a considerable number of people have already migrated.

In this paper, we study whether soil degradation, exacerbated through climate change, acts as a driver of internal migration. We use three waves of the Malawi Integrated Household Panel Survey (IHPS) to which data on soil quality from Harmonized World Soil Database (HWSD) is matched. We use an Instrumental variable approach (IV) to deal with the endogeneity of land degradation. Our IV consists in considering the average number of climatic shocks 15 years ago as an instrument of current land degradation while controlling for actual climate conditions. Our results, whether considering the endogeneity of soil quality or not, show that poor soil quality (soil with limited nutrient availability) initiates people to migrate. In particular, households with lands in states of severe constraints regarding soil nutrients have more family members that move between the two survey interviews.

To unpack the underlying mechanism, we further explore whether the channel through which land degradation influences migration decisions is agricultural productivity. We find that a lack of soil nutrients decreases agricultural yield, which could in turn encourage farmers to migrate. These findings are in line with a number of studies on the climate-driven migration that highlight climate change consequences on agricultural yields. We run a series of robustness checks: i) we change the type of indicator we use for soil quality; ii) we use population pressure as a proxy for degradation of soils; and iii) we run a sensitivity analysis in terms of the time window and shock thresholds in the construction of the IV. Our results are consistent across the different specifications.

The remainder of the paper is organized as follows. Section 2.2 reviews the related literature and describes the background of Malawi. Section 2.3 explains data sources, variable construction and provides summary statistics. Our econometric framework and estimation results are presented respectively in Section 2.4 and 2.5. We conclude our findings in the final section.

2.2 Literature Review and Background

In this section, we review the literature that studies the relationship between environment and migration. The framework of this literature is represented in Figure 2.1. The change of climate can affect land degradation in a fast or a slow process. Events such as hurricanes, floods, and droughts are considered fast shocks while other events can take a slower onset, like soil erosion, sea level rises, or land degradation. It is important to distinguish between the two because their effect on migration goes through different channels. Slow onset affects migration through productivity, which is different from the fast one where people resettle immediately after the shock.

Our study focuses on a slow onset process, which affects productivity and the decision to migrate internally. This fills a gap in the existing literature that considers mainly fast shocks and international migration, due to the difficulty in measuring the slow onset of land degradation. We develop the existing literature in the following sections here below.

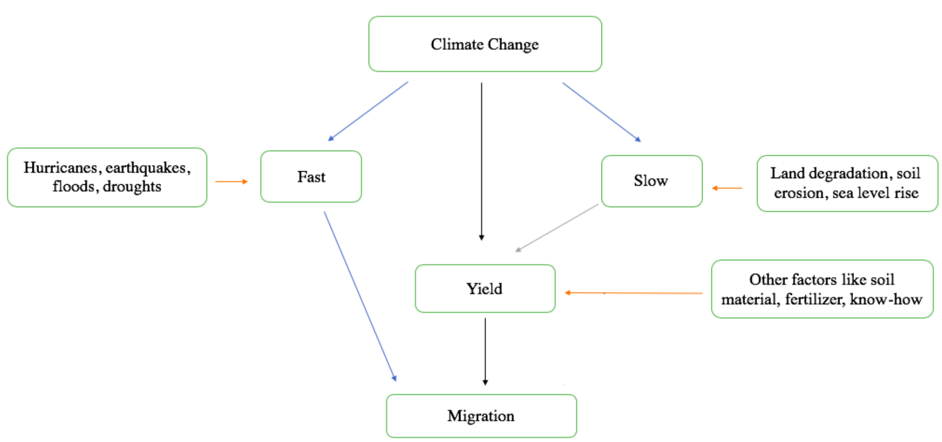


Figure 2.1. Framework of the link between climate change, land degradation and migration

2.2.1 Literature Review

Climate change and migration

Economists, ecologists, and human-environment researchers have drawn attention to the rapid rate of environmental change in many rural areas, including soil degradation, deforestation, climate change and the related displacement of potentially large numbers of environmental refugees. While most existing evidence in the literature focuses on international migration, internal migration remains understudied. Marchiori and Schumacher (2011) find that climate change will increase migration and that small impacts of climate change have a significant impact on the number of migrants. Dallmann and Millock (2017) study the impact of climate variability on internal migration by considering meteorological indicators of climate variability. They show that drought frequency in the origin state increases inter-state migration. This effect is stronger in agricultural states, and in such states, the magnitude of drought also increases inter-state migration significantly. Drought frequency in the origin state acts as a push factor on inter-state migration and it has the strongest effect on rural-rural inter-state migration. Gray (2009) finds that mean annual rainfall did not affect local migration but had a negative impact on internal and international migration. Harvest fluctuations, in contrast, increased local and regional migration, but not international migration. He concludes that international migration is the least influenced by environmental conditions, which confirms the literature's ex ante hypothesis that migration costs increase with the distance migrated. Sub-Saharan African agriculture is especially dependent on rainfall compared to most other developing countries, which triggers the potential impact of rainfall variations on economic activity. These impacts are intensified in rural areas where agriculture concentrates, and thus potentially affect rural-urban migration patterns. Drought-related changes in migration patterns, for instance, are more likely to emerge in areas where land degradation is severe (Hermans-Neumann and Hilderink, 2015). Human migration has been identified as a potentially important response to climate change. If habitation is difficult in some areas, people respond by moving into other areas. Gray and Mueller (2012a) investigate the effects of drought in rural Ethiopian highlands on population mobility over a 10-year period. They find that floods had no significant impact on migration but that weather-related crop failures increased migration. The results indicate that men's labor migration increases with drought and that land-poor households are the most vulnerable. However, marriage-related moves by women also decrease with drought. Coniglio and Pesce (2015) find that a large fall in average temperature during the rainy season increases out-migration, in particular in less developed economies. The magnitude of these effects is substantially smaller if these shocks occur during the dry season. [5] Policies regarding extreme weather events are especially valuable in resource-poor settings where climate-related vulnerability is high, such as for smallholder farmers in the developing world. Brausmann et al. (2019) show that there can be potentially large welfare gains from incorporating improved meteorological services into their decision-making process.

Environmental determinants of migration

Environmental factors that act over the long term, [6] such as soil erosion and decline of water availability and quality can affect migration decisions. LeBlanc and Perez (2008) find a positive correlation between rainfall and population density in the African continent below an annual rainfall of 900 mm. Density increases across the continent should lead to a significant increase in

[5] Also Bohra-Mishra et al. (2014) find that in 7000 households in Indonesia, an increase in temperature (e.g., due to natural variations or global warming) and, to a lesser extent, variations in rainfall are likely to have a greater effect on permanent out-migration of households than natural disasters. They find that rainfall has a quadratic effect such that in conditions that are initially dry, a decline in rainfall tends to increase migration, whereas, in wetter conditions, an increase in rainfall increases migration.

[6] To be distinguished from sudden shocks like natural disasters (Gray and Mueller, 2012c) and Beine and Parsons (2015) find higher internal migration after natural disasters and other find ambiguous or even negative effects like Gray and Mueller (2012b) and Tse (2011). Tse (2011) finds that earthquakes reduce the household size, earnings, and non-business assets, each of which tends to reduce migration rates. Volcanic eruptions on the other hand raise the value of farmland, which, in turn, reduces migration

the extent of water-stressed zones. Changes in rainfall, the pattern of which remains inherently uncertain today mitigate those effects. Wisner et al. (2004) writes that not only nature but also economic, social, and political factors are also affected by the consequences of natural hazards. There are a fast-growing number of people who can no longer gain a secure livelihood in their homelands because of drought, soil erosion, desertification, deforestation, and other environmental problems like in (Myers, 1997). Adverse environmental conditions have a tendency to increase migration but this is not always the case. Gray (2011) finds that soil quality significantly reduces migration in Kenya, particularly for temporary labor migration, but marginally increases migration in Uganda.

Climate change and productivity

Environmental changes affect agriculture through production and income effects. Climate change has changed urbanization in sub-Saharan Africa and it has been an important determinant of rural-urban migration. Weather, proxied by rainfall and temperature can be considered an input in the production function. It can also affect income by adjusting wages and general equilibrium effects like in (Barrios et al., 2006; Marchiori et al., 2012). Brausmann et al. (2019) study the high climate vulnerabilities of smallholders in the Peruvian Altiplano and they show that access to existing meteorological services is empirically associated with avoided losses in agricultural production that amount to 18 dollars on average per household and per year.

A mechanism by which the environmental variables influence migration is temperature, which is known to have a dominant influence on economic conditions (Schlenker and Roberts, 2009; Schlenker and Lobell, 2010). Dell et al. (2009) show that for the year 2000 countries are, on average, 8.5 percent poorer per capita per 1 degree Celsius warmer. [7] Hsiang (2010) finds that

[7] In another world sample from 1950 to 2003, they find that being 1 degree C (Celsius) warmer in a given year reduces per capita income by 1.4 %, but only in poor countries. Moreover, they estimate the model with lags in temperature and they find that this large effect is not reversed once the temperature shock is over, meaning that temperature is affecting not only income levels but also growth levels, which have economic consequences in

national output falls by 2.5 % for a 1 degree Celcius increase in temperature in Caribbean countries during 1970-2006. Schlenker and Roberts (2009) find that for US, yields increase with temperature up to 29 degrees C for corn, 30 degrees C for soybeans, and 32 degrees C for cotton but that temperatures above these thresholds are very harmful. Malpede and Percoco (2021) find that areas that experienced large soil aridification reduced the GDPs of African and Asian countries by 12 % and 2.7 %, respectively.

Productivity, wealth and migration

There is evidence that sudden-onset environmental events affect migration. Dust storms have affected the movement of people by causing a decrease in output in agriculture and the value of land (Gutmann et al., 2005). Gröger and Zylberberg (2016) find that internal labor migration facilitates shock coping in rural economies. Following a massive drop in income, due to a disastrous typhoon shock, households cope mainly through labor migration to urban areas. Non-migrant households react by sending new members away who then remit, and households with settled migrants ex-ante receive more remittances. Schlenker and Lobell (2010) studies the negative impact of climate change on African agriculture.[8] Hornbeck (2012) finds that adaptation through changes in crop choice was relatively minor and that adaptation occurred mainly through migration.

Many regions, especially developing countries, are expected to experience significant declines in agricultural yields as a result of projected warming. Feng et al. (2010) study the relationship between crop yields and migration. They find that a 10 percent reduction in crop yields in Mexico would lead an additional 2 percent of the population to emigrate to the United States. Dell et al. (2012) find that higher temperatures have substantial negative impacts in poor countries. They substantially reduce economic growth in poor countries,

the longer run.

[8]Crops affected are maize, sorghum, millet, groundnut, and cassava. They also find that countries with the highest average yields have the largest projected yield losses, suggesting that well-fertilized modern seed varieties are more susceptible to heat-related losses.

as well as agricultural output, industrial output, and political stability. [9]

2.2.2 Background

We focus our study on Malawi, a country in Sub-Saharan Africa, which relies highly on agriculture and that has been affected strongly by land degradation. Malawi is both one of the least urbanized countries in the world and has one of the highest rates of urban population growth.[10] [11]

Agriculture and Climate in Malawi

According to the World Bank, the agricultural sector contributed to the Malawian GDP by 30 percent in 2017. It generates over 80 percent of national export earnings and it concerns 64.1 percent of the labor workforce comprising mostly the smallholders subsistence farmers. Approximately 83 percent of the Malawian population leaves in rural areas and agriculture remains the primary economic activity. The majority of the population engages in smallholders, rain-fed agriculture production and only about 4 percent of the cultivatable land is under irrigation. The growing season is between November and May, whereas the rest of the year is characterized by a dry season. During the dry season, only a small number of farmers can exploit the residual moisture in valley floors (dambos).[12] Malawi has experienced

[9] Lack of resources can also affect conflicts which lead to migration. Environmental distress, changes in landscape, and soil material associated with scarcity of resources can also cause conflicts. Poor societies will be particularly affected since they are less able to buffer themselves from environmental scarcities and the social crises they cause. These societies are, in fact, already suffering acute hardship from shortages of water, forests, and especially fertile land (Homer-Dixon, 1991, 1994)

[10] It is estimated that in Africa, 51 percent of the poorest people occupy marginal lands; in Asia, 60 percent; and in Latin America, 80 percent (Myers, 1997).

[11] International Organization for Migration IOM: Migration in Malawi, A Country Profile 2014.

[12] Agriculture is also the main contributor to the national and household food security and nutrition. The main crops are maize, rice, sorghum, bananas, cassava, sweet potatoes, Irish potatoes, and legumes and the main industries deal with the agricultural processing of tobacco, tea, sugar, and timber products.

extreme weather episodes in the last years, both droughts and floods, which made the country highly vulnerable to climate variability, which is only expected to worsen under a warmer climate future. It is one of Southern Africa's most densely populated countries, with a population of about 17 million people spread over an area of 118,484 square kilometers. Its land is under extreme pressure with 200 people per square kilometer. Malawi is a landlocked, low-income, and agricultural-based economy (World-Bank, 2017). According to the Agriculture Public Expenditure Review report (2013), Malawi is getting around 21 percent of the national budget expenditure exceeding the Maputo declaration which recommended a 10 percent support to agriculture. Despite this, the land for agriculture in Malawi is becoming limited and affected by erosion linked to the increasing deforestation.

Migration in Malawi

Our hypothesis is that households opt for other off-farm opportunities if land degrades or they consider migration to avoid poverty.[13] Agricultural environments become more attractive and can pull households from seeking better opportunities elsewhere and thus migrate. Among other factors motivating migration in Malawi are economic reasons, marriage, or divorce. The country has witnessed a massive increase in internal migration, both rural-to-rural and rural-to-urban, and the bulk of migration experiences for Malawians today remain internal (Englund, 2002). Following natural disasters since the late 1970s related to drought, fires, floods, and landslides, the country has experienced numerous incidents of internal displacement. [14] International migration in Malawi has been decreasing from 1990 to 2013 for both sexes.[15] Rural residence is predominant: the vast majority of the

[13] Addressing the Land Degradation, Migration Nexus: The Role of the United Nations Convention to Combat Desertification 2019, Gray and Mueller (2012a)

[14] It is estimated that between 1979 and 2008, natural disasters in Malawi affected an estimated 21.7 million people and killed 2,596 people (International Disaster Response Law (IDRL) in Malawi 2015)

[15] International Organization for Migration IOM: Migration in Malawi, a country profile 2014

Malawian population resides in rural areas, and the majority of rural migrants move to other rural areas instead of urban centers (Anglewicz et al., 2017)

2.3 Data

We use two sources of data to analyze the interplay between the slow degradation of land and migration decisions. The household and plot data is from the Malawi Integrated Household Panel Survey (IHPS) which is a three-year survey conducted by the National Statistical Office of Malawi, with assistance from the World Bank. The panel survey develops from the Integrated Household Survey (IHS), which collects information on individuals, households, plots, and community-level characteristics of 12,288 households interviewed during the rainy season 2010/2011 - Third Integrated Household Survey (IHS3).[16] The panel component of this sample includes 1,619 that were further re-tracked and interviewed in 2013 - Fourth Integrated Household Survey (IHS4). Our sub-sample includes 1334 households in each wave where original nucleus households present in the three waves are taken into consideration. In Figure 2 can be found a timeline of the data on temperature, land quality, and migration, the main variables that we use in this study.

Our dependent variable is constructed using the IHPS information on individual migration outflows and household resettlement. The former indicates whether a family member migrated in the 6 months preceding the interview. The latter indicates whether the household has resettled at least 10 km away from the origin place in the first survey wave. We choose 10 km as

[16]The longitudinal panel dataset IHPS consists of five core questionnaire instruments; the Household Questionnaire, the Agriculture Questionnaire, the Fishery Questionnaire, the Community Questionnaire, and the Individual Questionnaire. We use the agriculture and the household questionnaire. We use this survey data because it is the only one that provides a panel (longitudinal dimension), the use of technology: GPS for households location and plots, descriptively exploits geo-spatial variables, household characteristics, the decision to migrate, and because it is open data. Malawi data has more than 12000 respondents different from the other LSMS countries that have around 3000-4000

threshold. [17] With this specific distance we do not consider the very short moves. Yet, taking into account that the distance to a main road is on average 7.5 km, 10 km can be considered a resettlement distance for Malawi standards.[18] According to Table 6, 6% of households have a migrant that has left in the past 6 months and 0.09 is the average number of migrants that left households in the last 6 months. 5% of households have resettled more than 10 km away since the previous wave.

We use two measures to evaluate soil quality, our variable of interest. The first measure is a subjective one, that captures the perceived land quality from the perspective of the farmer. The IPHS includes descriptive information on the farmer's perceived soil quality of each of the households operated plots.[19] The objective measure is incorporated as a part of the Geo-variables provided by the World Bank together with IPHS and originally comes from the Harmonized World Soil Database (HWSD v 1.21, used by WB) established by international institute for applied systems analysis and FAO.

We follow the development economics literature in using weather variables as instruments to identify shocks. By 'weather' we refer to temperature and precipitation at a specific time and place. By 'climate change' we refer to a permanent change in the average long-term levels of temperature and precipitation, and by 'climate variability' we refer to a permanent change in the long-term variance around the average levels (Deschênes and Greenstone, 2007). We are interested in climatic changes, rather than permanent cross-country climatic differences in levels.

To study migration in response to slow-onset changes, periods longer than

[17] Chort et al. (2018) uses 5km distance while Kubik and Maurel (2016) uses 10km. Other papers use like Hirvonen (2016) 'nearby village' as distance measure

[18] In the Robustness Section we account for different distance thresholds to study not only the link between land degradation and migration, but also the one between land degradation and distance of move.

[19] The question asked in the survey is the following: "What is the soil quality of this plot"? The proposed answer options are 1 - Good; 2 - Fair and 3 - Poor. To have an overall evaluation of the soil quality evaluated by a given farmer on all of its land holdings, we construct a weighted subjective measure where we weight each type of land quality, good, fair, and poor, by the size share of the corresponding plot in the total land holdings.

a few years must be taken into account, given the long response time of these changes. We use data from the National Oceanic and Atmospheric Administration (NOAA) and from the European Centre for Medium-Range Weather Forecasts (ECMWF). According to the World Bank, the mean annual temperature in Malawi increased by 0.9 degrees Celcius during the period 1960 to 2006 (World Bank, 2012). The average increase in temperature is concentrated in the rainy summer season (December–February).

We compute a Standardized Precipitation Index SPI based on the information included in the previous dataset, which is a widely used index to characterize meteorological drought on a range of timescales. On short timescales, it uses soil moisture as a short term and groundwater and reservoir storage as a long time indicator.[20] The SPI specifically addresses the intensity of meteorological drought or precipitation deficit. The shortage of precipitation is a fundamental, intuitive metric for drought. The IHPS survey provides the GPS location (latitude and longitude) of the 197 enumerator areas addressed by the survey, so we can merge the socio-economic data with the rainfall and temperature records from two sources. Our SPI score accounts for rainfall anomalies starting from 15 years before the start of the survey until 25 years backward. For example for the survey of 2010, which is answered in 2008-2009, SPI measures are for the years 1983-1993.[21]

We also account for observed intrinsic household characteristics like age, sex, education, household size, and a dummy on rural or urban areas. Degradation of environmental conditions can contribute to out-migration, while in-migration can cause environmental changes at the immigrants' destination. High population density places additional stress on local natural resources (Hermans-Neumann et al., 2017). Thus, we also control for district population density. We consider these characteristics to further explore different behavior in the households and community. Our sample has a

[20] John Keyantash, National Center for Atmospheric Research Staff (Eds.), The Climate Data Guide: Standardized Precipitation Index (SPI) (2018)

[21] For the survey in 2013, whose data was collected during 2011-2012, SPI measures are collected for the years 1986-1996. For the survey in 2016, whose data was collected during 2014-2015, SPI measures are based on the years 1989-1999.

household size of 5 members on average and 2.45 of them belong to the age group 14-60. 25% of households live in urban areas. 23% of households are of a female head. The average years of education in the family is 5.56. 56% of households are credit constrained. The distance of a household from the main road is 7.75 km on average. More information on descriptive statistics can be found in the summary statistics of Appendix D.

Maps of Malawi regarding land quality, migration patterns, and climate change are shown in Appendix E. Land quality is perceived as worst in the southern part of the country. The south and the center are the most exploited areas in the country in terms of agriculture. They were also the areas in which the estate and state sector was established, which had a negative effect on the production capacity of the smallholders' sector (Green, 2007). The subjective land quality is correlated to a high Topographic Wetness Index (TWI), shown in Figure 4 of the same appendix.[22] The objective land quality shows a decrease in nutrient availability in 2 regions, respectively in the center and the south of the country. Internal migration patterns are heterogeneous but they seem to increase in the central part of the country from 2010 to 2013 and in the southern part of Malawi from 2010/2013 to 2016. Internal migration in the north decreases during these years. Resettlement gets stronger in the central area of the country and one region of the north. An explanation is a search for unused, arable land, considering both a primarily agricultural base and a high birth rate (IOM, 2014).

[22]Topographic Wetness Index (TWI) expresses the terrain-driven balance of the catchment water supply and local drainage.

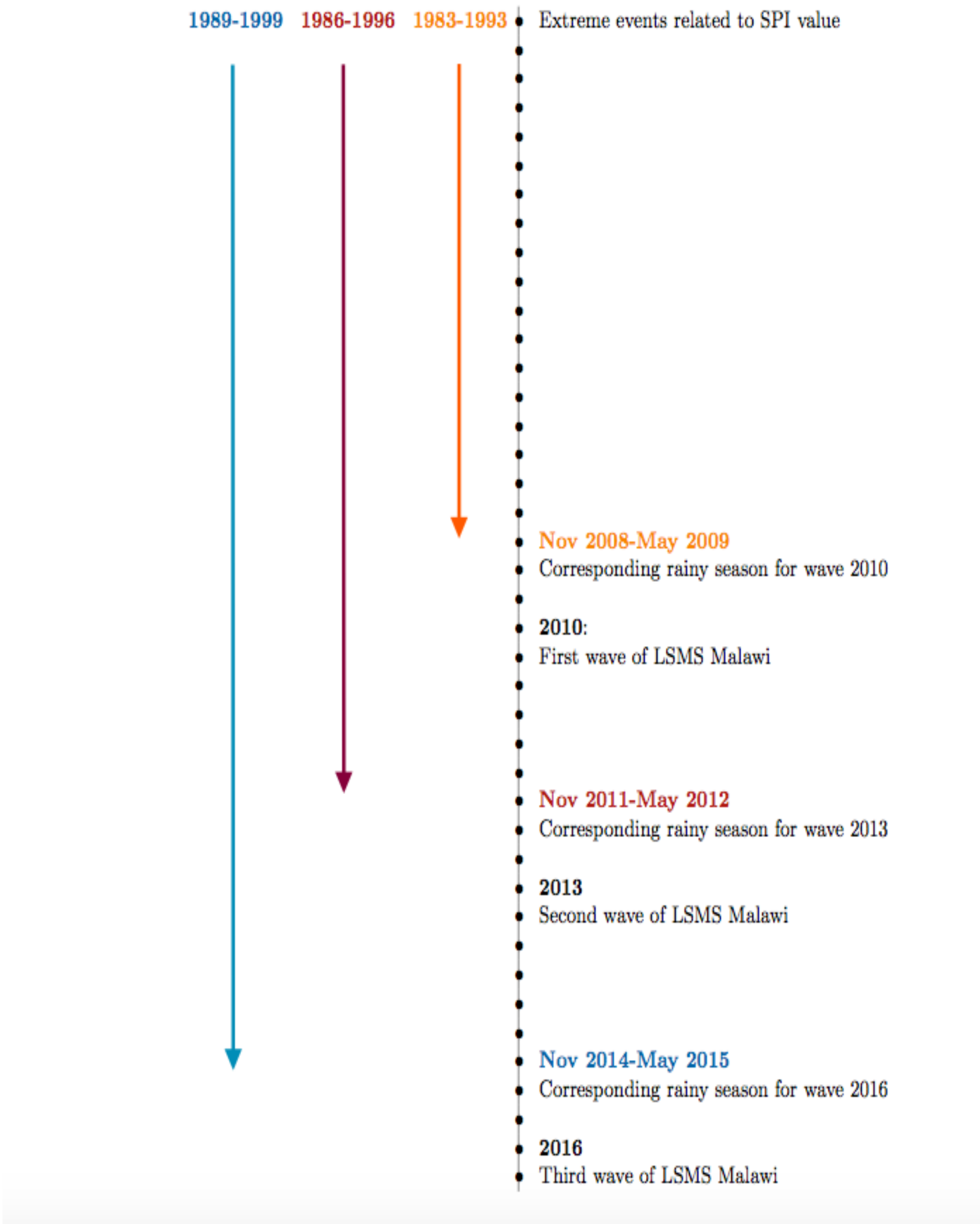


Figure 2.2. Timeline of the data on SPI, land quality and migration

2.4 Identification Strategy

In this section, we discuss the empirical approach that we adopt to study the impact of land degradation as a possible push migration factor. We describe the different identification issues and propose an IV framework to reduce the bias in the estimates. The main estimation equation is as follows:

$$M_{ht} = \alpha_0 + \alpha_1 LD_{ht} + X'_{ht}\alpha_2 + \alpha_3 R_{ct} + \eta_t + \varepsilon_{ht} \tag{2.1}$$

where M_{ht} is the migration of a given household h or member of a household h in year t. We use three different variables: a dummy variable if a member of a household (HH) migrated in the 12 months preceding the interview; a variable of the total number of HH migrants; and whether the whole HH moved more than 10 km away. We aim to consider different types of movements as migration responses to soil degradation might be heterogeneous. In particular, we would like to evaluate to which extent soil degradation affects individual migration, both on the intensive and extensive margin, as well as the collective, household migration. LD_{ht} stands for land degradation with capturing the soil nutrients availability. It is a categorical variable based on the qualitative evaluation of the Harmonized World Soil Database.[23] It takes value one if a HH faces severe nutrient availability constraints, zero otherwise. Household characteristics are represented in the vector X_{ht}, which includes the household head's gender, age, average years of education, whether the HH lives in an urban area, a dummy variable indicating if a HH is credit constrained and wealth quantiles determined by principal component analysis. R_{ct} is a dummy variable indicating if there was a rainfall anomaly based on the SPI score (SPI<-1.28) in the community c where household h belongs in year t. η_t gathers time dummies.

If Equation (3.1) is estimated by using Ordinary Least Squares (OLS), $\hat{\alpha_1}$ - $\hat{\alpha_3}$ will be biased. Land degradation is not a random process - it depends on different environmental factors that can influence migration, and other individual or household decisions while modifying/controlling for its degree.

[23] Harmonized World Soil Database 2012

In this sense, endogeneity sources can be of the three traditional kinds: selection bias, omitted variables, and reverse causality. First, households that tend to improve land quality or operate with the land of better quality might be systematically different from those who have access to the land of poor quality. Such systematic differences could affect different household decisions including migration ones. Second, there might be some observable or unobservable individual and household characteristics (e.g. entrepreneurial spirit of farmers) that could simultaneously affect migration decisions and the degree of land quality of the operated land. Introducing household fixed effects in Equation (3.1) would purge the time-invariant household characteristics. However, the time-variant unobservable characteristics would remain. Land degradation being a gradual process that varies mildly over time, including household fixed effects in Equation (3.1) would limit the possibility to estimate any effect of land degradation on migration because of low within variability. Therefore, we estimate our main specification by using a pooled OLS model, we cluster the error term, ε_{ht}, at the household level and we tackle omitted variable bias by adopting an IV approach. Finally, reverse causality between migration and land degradation can arise and bias the estimates. In particular, households who have migrants have a lower endowment of farm labor which can have an incidence on the adoption of other farming inputs such as fertilizers. We address the different types of bias by instrumenting the degree of land degradation using climatic events that occurred far away in time while controlling for current events through R_{ct}.

The first stage estimation equation is defined as follows:

$$LD_{ht} = \beta_0 + \beta_1 R_{ct-15} + \beta_2 R_{ct} + X'_{ht}\beta_3 + \eta_t + \nu_{ht} \tag{2.2}$$

where LD_{ht} indicates whether a household faces severe soil nutrient availability constraints in year t. R_{ct-15} is the instrumental variable and it represents the average number of negative precipitation anomalies (SPI<-1.28) since 1983, related to the first date of the LSMS data availability, till 15 years before the year of survey in the community c where the household currently belongs.[24] The drivers of land degradation can be of environmental

[24]In the Robustness checks section we do the analysis modifying this threshold and the

and human nature (Olsson et al., 2019). Among the environmental drivers, the climatologist literature has identified gradual changes in precipitation, temperature, and wind, including distribution and intensity of extreme events among the main climate-change drivers of land degradation (Lin et al., 2017; Olsson et al., 2019). Distant past rainfall changes have long-run consequences through their direct implications for land degradation, but they are less likely to directly influence current migration decisions which are more likely to be subject to recent weather events. To account for such temporal correlation weather conditions, we control for both past rainfall changes R_{ct-15}, and current R_{ct} precipitation anomalies. In this sense, we assume that past rainfall changes affect migration decisions through their impact on land degradation.[25] Further, we account for the same household characteristics that are represented in the vector X_{ht} and time dummies included in η_t.

2.5 Results

In this section, we discuss the different results of the empirical specification described in the above section. We first focus on commenting on the estimation results from the specification that does not account for the endogeneity issues (LPM hereafter). We then turn to discuss the estimates from the IV specification. In particular, we first focus on the stage regression results and comment on the relevance of the instrument, and then we discuss and compare the second stage results with respect to the LPM specification.

The results are shown in Table 2.1. The table is organized into 3 panels. It shows the impact of soil nutrient constraints on 3 types of migration: 1) when the household has a migrant, 2) the number of household migrants and 3) when the household resettles more than 10 km away. We find that severe

results in the first and second stage of our analysis are consistent.

[25]It should be noted that past extreme weather events could have other political (e.g. conflicts) and socio-economic (e.g. prices and food security variations) events consequences which affects human movements. Therefore, it is important to acknowledge that exclusion restrictions are not fully satisfied in this framework. Yet, we show the test scores of standard IV tests to examine the relevance of the instrument we select.

constraints of soil nutrients have a statistically significant and positive effect on the probability of having a migrant and on the number of migrants within a household. Similarly, households facing severe soil constraints have more family members that move away than households that do not face such constraints. This is not the case when we examine the impact of severe soil constraints on resettlement. The absence of such an effect can be related to a 'poverty trap' argument, as household resettlement might imply higher costs of migration than individual migration. When land degrades, its productivity lowers, and cannot provide sufficient income that covers costs to resettle. In terms of magnitude, severe soil constraints increase the probability of having a migrant by 5.4 % compared to households that do not face such constraints. Regarding the intensive margin, the impact is relatively low: the number of migrants of households dealing with severe soil constraints is 0.05 higher than households with better soil conditions.

Regarding wealth, we find that richer people resettle more. Households belonging to the third wealth tercile are associated with a higher probability of household resettlement than the households with lower wealth levels. However, this impact is only significant at a ten percent level of statistical significance and is not observed for the other types of migration. Being in the third tercile of wealth increases the probability of resettlement for more than 10 km by 1.8 % compared to households belonging to the other wealth terciles. Previous studies like Dustmann and Okatenko (2014) and McKenzie and Rapoport (2007) also show that wealth is an important determinant of migration.

Education of the household head is positively associated with household having a migrant and the number of migrants per household. It is statistically significant with the household resettlement. One more year of education increases the probability of resettlement by 0.2 %. Our results are in line with the ones of Docquier et al. (2014); Mountford and Rapoport (2011, 2012) as more educated people not only resettle more but also have more migrants that left home. Age does not affect the first and second type of migration. Age is statistically significant at a one percent level for household resettlement. An increase in the age of the household head by one year decreases the probability of resettlement by 0.1 %. In the literature, younger people migrate more like in Chort et al. (2018) and Stark and Bloom (1985). Here we look at the age of the household head, where experience might be useful when making decisions to resettle the whole household. Female household heads increase the probability that the household has a migrant by 3.3 % compared to male household heads. It is in line with Chort et al. (2018) but also with literature that gender matters in migration (Bertoli and Marchetta, 2012). This is significant at the one percent level. It also increases the number of migrants by 0.07 at a one percent significance level. It is not statistically significant for the resettlement measure and it is associated negatively to it.

Household size is significant at the 1 percent level for both measure 1) and 2) of migration but not for measures 3). An increase by one member in household size increases the probability that the household has a migrant by 1 % and the number of migrants by 0.023. It is associated negatively with resettlement. The credit constraint of the household is associated negatively with measures 1) and 3) of migration but positively with measure 2). We can say the same about the location of the household, specifically urban areas households. The biggest driver to explain migration and the number of migrants is household size. 1 standard deviation increase in household size boosts migration by 9 % and the number of migrants by 11 %. Severe constraints is the second biggest driver explaining migration. 1 standard deviation increase in the severe constraints, boosts the number of migrants by 8 %. Age is the determinant in resettlement. 1 standard deviation increase in the age of household head, diminished the number of migrants by 7 %.

Negative deviations from the rainfall level is not significant. It is positively associated with the 3 measures of migration. We do not find such an impact because there were not too many shocks during the period.

Instrumental Variable: We consider the average number of climate events 15 years ago as an instrument of climate change and its effects on soil nutrients. Since soil quality is difficult to measure we consider a proxy. We use this instrument as it is strongly correlated with soil quality and it is independent of migration.

The first stage results are presented in Table 2.2. We find that the average number of extreme climate events that occurred 15 years prior to the year of the survey is a significant and strong determinant of severe constraint in terms of soil nutrients. An increase of the average number of climate events by 1, increases the probability of having a highly severe soil constraint by 59.3 %. Negative rainfall shock is not significant but it is associated positively with the severe constraints of the soil nutrients. This is the biggest driver of soil quality. 1 standard deviation increase in negative rainfall shock, boosts severe constraint by 59 %. Looking at the household characteristics, we see that the education of household head and the urban location of the household are also statistically significant. Being in an urban area affects negatively the severity of soil in terms of nutrients. Location in an urban area is the second biggest driver of soil quality. Having a high education of the household head affects negatively the severity of soil in terms of nutrients. An increase of 1 year of education for the household head of the family, lowers the constraint severity of the soil nutrients by 0.3 %. When the household is in an urban area, the constraint severity of soil nutrients lowers by 17.2 %. Household size is significant at the ten percent level. An increase of the household by 1 more member raises the constraint severity of the soil nutrients by 0.3 %. Having a female household head is associated with fewer soil nutrient constraints. Age does not affect the first and second type of migration. There is no difference between different ages of the household head. Location is statistically significant at the one percent level. Households that have land in urban areas have 17.2 % more severe constraints regarding the soil nutrients of their lands. Wealth and credit constraints are associated positively with the dependent variable but they are not statistically significant.

The second stage results are presented in Table 2.3. The table is organized into 3 panels. Severe constraints on soil nutrients is statistically significant in households that have a migrant that has left home. An increase in severity of 1 degree of the soil nutrients raises the probability that households have a migrant by 22.9 %. It is associated positively with the 2) and the 3) measures of migration but it is not significant. Severe constraints is the biggest driver of migration and the number of migrants. An increase of 1 standard deviation of severe constraints boosts migration by 31 % and the number of migrants by 18.8 %. When rainfall shocks are negative, it increases the probability that the household has a migrant by 2.9 %. This is significant at the five percent level. It also raises the number of migrants by 4 % and it is significant at the ten percent level. Furthermore, it is associated positively with the 3) measure of migration. Focusing on the household characteristics, the education of the household head is positively associated to the 1) and the 2) measure of migration. It is significant at one percent for the third measure of migration. An increase of 1 year of education for the household head, raises the probability of household resettlement by 0.3 %. Age does not affect the first and second type of migration. Age is significant at the one percent level for the 3) measure of migration. It is the biggest driver of resettlement, although its effect is negative. One year older heads of household resettle less by 0.1 %. Having a female as head of household is significant for the first and the second measure of migration. It is associated negatively for the third measure. Having a female household head increases the chances that the household has a migrant by 3.2 % and increases the number of migrants by 7.1 %. Household size is the biggest driver of migration out of all household characteristics. It is significant at the one percent level for the first and the second measure of migration. It is associated negatively with the third measure of migration. An increase in the household size by 1 member, increases the chances of households having a migrant by 1 % and increases the number of migrants by 2.3 %. Being in an urban area is negatively associated with migration. It is significant at the ten percent level for the first measure of migration. Being in an urban area lowers by 3.5 % the chances that the households have migrants. Credit constraint is negatively associated with migration. Being in the second tercile of wealth is positively associated with the first and the second measure of migration but negatively with the third. Being in the third tercile of wealth is positively associated to the first measure of migration, negatively to the third and it is significant at the 10 % level for the third. Households that are in the third tercile of wealth resettle by 1.7 % more. Overall our results that livelihoods are increasingly challenged by the dual threats of land degradation and climate change are consistent with Call and Gray (2020).

Table 2.1. The impact of soil nutrient constraint of different forms of migration

	(1) HH has a migrant	(2) Number of HH migrants	(3) Resettlement > 10km
Severe constraint: soil nutrients	0.054***	0.050*	0.014
	(0.02)	(0.03)	(0.03)
Negative rainfall shock	0.024	0.036	0.002
	(0.02)	(0.03)	(0.01)
Education of HH head	0.002	0.002	0.003**
	(0.00)	(0.00)	(0.00)
Age of HH head	0.000	0.000	-0.001***
	(0.00)	(0.00)	(0.00)
Female HH head	0.033***	0.070***	-0.003
	(0.01)	(0.03)	(0.01)
HH size	0.010***	0.023***	-0.001
	(0.00)	(0.01)	(0.00)
HH is credit constrained	-0.013	-0.001	-0.002
	(0.01)	(0.01)	(0.01)
Urban HH	-0.004	0.006	-0.015
	(0.01)	(0.02)	(0.01)
2nd tercile wealth	0.006	0.015	-0.002
	(0.01)	(0.02)	(0.01)
3rd tercile wealth	0.006	0.001	0.018*
	(0.01)	(0.02)	(0.01)
Obs.	3,694	3,694	3,694
R^2	0.021	0.021	0.039

Note: Standard errors in parenthesis are clustered at household level. The negative rainfall anomaly is a binary variable indicating if SPI is lower than -1.28. Time dummies are included in the specification. Significant coefficients are denoted with stars as follows: *** p<0.01, ** p<0.05, and * p<0.1.

Table 2.2. First Stage Results: Past Climate as IV

	Severe constraint: soil nutrients
Average number of climate events 15 years ago	0.593***
	(0.06)
Negative rainfall shock	0.008
	(0.01)
Education of HH head	0.003***
	(0.00)
Age of HH head	-0.000
	(0.00)
Female HH head	-0.008
	(0.01)
HH size	0.003*
	(0.00)
HH is credit constrained	0.007
	(0.01)
Urban HH	0.172***
	(0.01)
2nd tercile wealth	0.014
	(0.01)
3rd tercile wealth	0.010
	(0.01)
Nr. observations	3,694

Note: Standard errors in paranthesis are clustered at household level. The negative raifall annomaly is a binary variable indicating if SPI is lower than -1.28. Time dummies are inclded in the specification. Significant coefficients are denoted with stars as follows: *** p<0.01, ** p<0.05, and * p<0.1.

Table 2.3. Second Stage: the impact of soil nutrient constraint of different forms of migration

	(1) HH has a migrant	(2) Number of HH migrants	(3) Resettlement > 10km
Severe constraint: soil nutrients	0.229**	0.188	0.031
	(0.10)	(0.18)	(0.08)
Negative rainfall shock	0.029**	0.040*	0.002
	(0.01)	(0.02)	(0.01)
Education of HH head	0.001	0.002	0.003***
	(0.00)	(0.00)	(0.00)
Age of HH head	0.000	0.000	-0.001***
	(0.00)	(0.00)	(0.00)
Female HH head	0.032***	0.071***	-0.004
	(0.01)	(0.02)	(0.01)
HH size	0.010***	0.023***	-0.001
	(0.00)	(0.00)	(0.00)
HH is credit constrained	-0.014*	-0.002	-0.002
	(0.01)	(0.01)	(0.01)
Urban HH	-0.035*	-0.018	-0.015
	(0.02)	(0.04)	(0.02)
2nd tercile wealth	0.003	0.013	-0.003
	(0.01)	(0.02)	(0.01)
3rd tercile wealth	0.004	0.000	0.017*
	(0.01)	(0.02)	(0.01)
Nr. observations	3,694	3,694	3,694
R^2	-0.010	0.015	0.038
Kleibergen-Paap LM P-val	0.0	0.0	0.0
Cragg-Donald Wald F	97.4	97.4	97.4

Note: Standard errors in parenthesis are clustered at household level. The negative rainfall anomaly is a binary variable indicating if SPI is lower than -1.28. Time dummies are included in the specification. Significant coefficients are denoted with stars as follows: *** p<0.01, ** p<0.05, and * p<0.1.

2.5.1 Discussion of Results

The difference between OLS and IV results in our study can be interpreted by the attenuation bias. The coefficient of the severe constraints in the OLS is 0.054 while in the second stage of IV is 0.229. Attenuation bias, also called regression dilution, is a bias in model coefficients caused by measurement error or noise in the independent variables. These errors in the independent variable bias the linear regression slope towards zero.

Moreover, our results should be interpreted with prudence, especially regarding the exclusion restriction. This caution is due to the fact that we cannot formally test the exclusion restriction but by controlling for present climate we account for temporal correlation of weather conditions. Past climate events can influence income-generating decisions, and risk aversion regarding weather, which later on can influence migration decisions. Another cautious aspect is regarding our sample, which might already be downsized. Considering we have only 6 years of data in our study, people might have already migrated and we are left with a sample of people that are not willing to move.

Furthermore, our study comes with limits that future and better datasets might eliminate. Future research can aim at more precise data on soil quality like precise chemical components that shed light on how they exactly affect productivity in agriculture. There is also a need for a longer time span of the data. We have only 3 years, but land degradation happens more slowly over time, which leads to a need for a longer timeframe. Another limit is the definition of migration, where 6 months is arbitrary. With the data we have, the results we show are actually an underestimation of the effects. Land degradation is a slow effect, where many people have already migrated throughout the years that the land degraded, leaving a downsized sample in the dataset. Furthermore, it would be valuable checking in what activities people are involved in after they migrate, if they continue working in agricultural activities in a better land quality plot, or if they shift to another activity. With the data that we have available, we are limited and cannot analyze it not knowing the destination area. Future studies, if data allows, should consider checking it.

2.6 Robustness checks

This section is organized as follows: We first run a set of robustness checks concerning how soil health measures are defined. In particular, we use another measure for soil constraints faced by farmers. We focus on the soil workability features of the land that is measured similarly to our baseline definition, the availability of nutrients in the soil.

We then discuss the underlying mechanisms behind our results. In a reduced-form estimation framework, we show that poor soil quality (soil with difficulty of workability) contributes to lower agricultural yields. In this sense, poor health of soil drives within households migration, as a way to adapt to soil degradation, which leads to income losses.

2.6.1 Soil quality measured as ease of tillage

One should consider soil workability as ease of tillage that depends on the different soil characteristics. The ease of working with a particular parcel with a given soil quality depends on its characteristics, such as texture, organic matter, and stones on the profile (for an in-depth discussion, see Appendix C). Such characteristics influence the ease of using labor and machines in the field which have an incidence on yields. Similarly, we construct a dummy variable indicating whether a soil plot has severe workability constraints and re-estimate our main equation. We account for the endogeneity of soil health as farmers' past decisions and activities could contribute or control to the present state of soil quality. We employ the same 2SLS strategy, whose results are presented in Tables 2.4 and 2.5.

Our previous findings are consistent across the different specifications. On the other, hand we do not consider indicators like slope, elevation, and tropical wetness index as they are fixed features that do not unfold details about the quality and composition of the land, but they are implicitly accounted for in the soil indicators used in the present paper. Moreover, we do not consider population density which is also a popular measure in the literature. The drawback of the way population density is measured is that it is invariable over time as it is based on the 2009 Malawi Census data.

Table 2.4. First Stage Results: Past Climate as IV

	sq71_d Severe constraint: soil workability
Average number of climate events 15 years ago	0.986***
	(0.10)
Negative rainfall shock	-0.026
	(0.02)
Education of HH head	0.002
	(0.00)
Age of HH head	-0.001***
	(0.00)
Female HH head	-0.016
	(0.02)
HH size	-0.014***
	(0.00)
HH is credit constrained	-0.010
	(0.01)
Urban HH	0.172***
	(0.01)
2nd tercile wealth	0.014
	(0.02)
3rd tercile wealth	0.088***
	(0.02)
	(0.02)
Nr. observations	3,694

Note: Standard errors in parenthesis are clustered at household level. The negative rainfall anomaly is a binary variable indicating if SPI is lower than -1.28. Time dummies are included in the specification. Significant coefficients are denoted with stars as follows: *** p<0.01, ** p<0.05, and * p<0.1.

Table 2.5. Second Stage: the impact of workability constraint of different forms of migration

	(1) HH has a migrant	(2) Number of HH migrants	(3) Resettlement > 10km
Severe workability constraints	0.139**	0.118	0.015
	(0.06)	(0.11)	(0.05)
Negative rainfall shock	0.034**	0.044*	0.003
	(0.01)	(0.03)	(0.01)
Education of HH head	0.002	0.003	0.003***
	(0.00)	(0.00)	(0.00)
Age of HH head	0.000	0.000	-0.001***
	(0.00)	(0.00)	(0.00)
Female HH head	0.032***	0.071***	-0.003
	(0.01)	(0.02)	(0.01)
HH size	0.012***	0.025***	-0.000
	(0.00)	(0.00)	(0.00)
HH is credit constrained	-0.011	0.000	-0.001
	(0.01)	(0.01)	(0.01)
Urban HH	-0.021	-0.006	-0.013
	(0.02)	(0.03)	(0.01)
2nd tercile wealth	0.004	0.015	-0.003
	(0.01)	(0.02)	(0.01)
3rd tercile wealth	-0.005	-0.004	0.013
	(0.01)	(0.02)	(0.01)
Nr. observations	3,694	3,694	3,694
R^2	-0.031	0.011	0.037
Kleibergen-Paap LM P-val	0.0	0.0	0.0
Cragg-Donald Wald F	101.1	101.1	101.1

Note: Standard errors in parenthesis are clustered at household level. The negative rainfall anomaly is a binary variable indicating if SPI is lower than -1.28. Time dummies are included in the specification. Significant coefficients are denoted with stars as follows: *** p<0.01, ** p<0.05, and * p<0.1.

2.6.2 Agricultural Productivity

The underlying mechanism of why one would expect soil degradation to act as a push factor for individual and household migration is via its negative incidence on farm income. In this section, we attempt at verifying whether the intuition holds for the farmers included in our sample. The first challenge is to quantify production for farmers that are engaged in subsistence agriculture. The literature has identified problems of measurement errors and recall bias in survey data when farmers are asked to provide subjective evaluation of the quantity they produced, used for self-consumption, or sold at the local markets (Eitzinger et al., 2018; Methorstab et al., 2017). In addition to production measurements, land size is often subject to similar biases (Dillon et al., 2016; Dillon and Nagraj Rao, 2018).

We have two types of production information depending on the timing of the interview with the respect to the agricultural calendar. If the harvesting period did not occur or finish prior to the interview, farmers provide information on the expected yields. In the other case, if the agricultural season is over, then data on effective production is provided.[26] We make use of both types of data. As farmers produce a crop portfolio, one needs prices to evaluate the overall agricultural yields. Part of the farmer's agricultural production is used for consumption, and part of it is sold. The dataset we use contains information on the value of the quantities sold, which allows us to compute prices for different crops. To reduce recall bias, we use average crop prices at the community level.[27] Agricultural yields include both self-consumed and sold quantities per operated acre of land evaluated at community price levels (Malawian Kwacha (MWK)) per acre.

Further, we model yields as a function of land quality, operated land size,

[26]In some cases, depending on the timing of the agricultural season, farmers could provide both types of information.

[27]If the data regarding price information are missing at the village level, then we use the national level as a reference. We can also have different types of the same category of product (for example: different types of maize). We proceed by aggregating them under one category, in this case, 'maize'. In particular, we aggregate the production for the following product categories: maize, tobacco, groundnut, and rice.

and other agricultural inputs fertilizer use, both organic and inorganic, pesticide use, livestock use in tropical livestock units (TLU) and labor in the total number of days worked on the farm by all the adults belonging to a given household, maximal temperature and total precipitation level in mm within the agricultural season of the current year. Rainfall intensity and temperature count as agricultural inputs in economies where irrigation systems are absent (Torres et al., 2019). The results of the estimation of expected and effective yields are presented in Table 7.

Intuitively, operating in an environment with severe soil nutrient constraints reduces on average the effective agricultural yields by 200760 MWK per hectare compared to farmers that do not face such constraints.[28] This impact is statistically significant at the five percent level and economically important as it represents about 50 percent of the average sample yield. It is therefore natural to expect that farmers who are dealing with soil constraints might opt for moving away as searching of better economic opportunities. The difference in terms of expected yield in between farmers facing severe soil constraints and those who do not is statistically insignificant. Concerning the standard agricultural inputs, as expected using fertilizer increases effective agricultural yields by 280084 MWK/hectare whereas an increase of one degree Celsius in the maximal temperature, decreases effective productivity by 55818 MWK/hectare.

Regarding expected yields, the maximal temperature has a consistent detrimental effect, however current livestock, and precipitation intensity slightly reduce expected yields. One additional day of agricultural work increases expected yields by 26.6 MWK/hectare, whereas the effect is insignificant for effective yields. The differences between effective and expected yields as well as the slight counter-intuitive effects of precipitation and livestock could result from the following potential explanations: First, yields are highly volatile therefore it might be difficult for farmers to evaluate well-expected outcomes. Second, the timing of the season when farmers are interviewed might have an influence on the estimation of future outcomes.

[28] This value corresponds to around 195 euros/hectare

Third, the sample we use contains missing values both for RHS and LHS variables. Therefore, interpretations of the different findings should be considered with caution.

Table 2.6. The impact of soil nutrient constraint on yield

	(1) Effective yield	(2) Expected yield
Fertilizer	280084.066**	10242.001
	(131507.95)	(31128.24)
Pesticides	317012.555	111249.794*
	(268117.38)	(60813.42)
Ag. labor in days	9.005	26.674*
	(18.33)	(13.75)
TLU	10946.539	-1203.024
	(12452.19)	(1217.29)
Season precipitations	-391.966	-195.348
	(376.93)	(119.06)
Season max temperature	-55818.757	-16674.843
	(35680.87)	(12224.75)
Severe constraint: soil nutrients	-200760.799**	16087.941
	(90934.04)	(51482.47)
Obs.	2,817	2,817
R^2	0.022	0.025

Note: Significant coefficients are denoted with stars as follows: *** p<0.01, ** p<0.05, and * p<0.1.

2.7 Concluding Remarks

In this paper, we examine the intensive and extensive margin that land degradation has on internal migration and household resettlement. Our study shows how trends in rainfall affect soil quality indicators, which later on have an effect on migration. The results have important implications for internal migration dynamics given that climate change is expected to worsen, thereby threatening agricultural livelihoods, especially in risk-prone areas. We find that slow onset land degradation affects internal migration. We provide these results by first showing how climate change affects soil quality, and then secondly how soil quality affects migration.

In this study, we also focused on the climate acting through its effect on agricultural productivity yields. These results have important implications for future studies of migration streams and the relationships between migration, usage of the land, and the environment. This analysis of environmental effects on migration supports the overall importance of environmental factors by considering agricultural productivity as a push or pull factor for internal migration. It also suggests that environmental factors influence migration through multiple pathways. These results complement the existing findings that suggest an influential role of climate change on migration.

Migration and environmental policies are interlinked and cannot be optimally designed without their effects on one another being taken into account. Therefore our study aims in adding new information and results on environmental policies. It is important to note that future climatic change can be understood as an ongoing process rather than a permanent shock.

Previous literature proposes mitigation and adaptation as possible ways of dealing with climate change. We consider adaptation as an uncertain step for the following reason. There are regions that are already extremely poor and vulnerable and when climate change affects them, they are unable to adapt. As a result, unless in the presence of a poverty trap, the only solution is to move away from those areas and find areas with better living and working conditions. Given the uncertainty over adaptation, and the spread of income risk or remittances, the estimation of climate change effects on the economy and migrations is not precise. The results do, however, provide clear guidance on the link between global climate change, land productivity, and migration.

Soil degradation plays an increasingly important role in the present and future of agriculture, given the prospect of global warming. Although the underlying soil quality/migration relationship is complex and the evidence is sometimes indirect, there is a need not only in forecasts to prepare for future immigration flows in response to land degradation but also on tackling the right determinants.

2.8 Appendix

2.8.1 The inter-link between migration and environmental changes

In the 1970s, Lester Brown of the Worldwatch Institute popularized the idea of environmental refugees. It was originally used in a document titled "Environmental Refugees" published by the United Nations Environment Programme in 1985. A forced environmental migrant is someone who has no choice but to leave their home due to an environmental stressor, as opposed to an environmentally motivated migrant, who may choose to migrate as a result of an environmental stressor.

There are different types of migration of interest for studying the environment-migration nexus. They include internal versus international migration, rural-rural versus rural-urban migration, and temporal versus permanent migration. We focus on the internal movement of people only. Internal migration occurs within a given country by crossing sub-national boundaries (for example, by moving between districts or provinces), different from international migration which involves crossing national borders. There is a gap in the studies that link internal migration flows with slow environmental changes. For this reason, we try to link them by analyzing possible responses of SSA crops to climate change. Roughly 17 percent of GDP was derived from agriculture in Sub-Saharan Africa, with this fraction in excess of 50 percent in some countries (World-Bank, 2009). Climatic factors affect economically relevant outcomes, agricultural output, economic growth, health, and conflict, thus a careful understanding of such effects may be essential to the effective design of contemporary economic policies and institutions (Dell et al., 2014). Changes in crop yields, and water scarcity that result from climate changes occur over broad geographical areas and are likely to lead to long-term population shifts. This is different from climate processes such as sea-level rise, salinization of agricultural land, desertification and growing water scarcity, and climate events such as flooding, storms, and glacial lake outburst floods, which happen in specific areas. Crop yields in agriculture are especially relevant to developing countries, that typically have large rural populations that derive a living directly from agriculture. The ability to migrate is a function of mobility and resources (both financial and social). In other words, the people most vulnerable to climate change are not necessarily the ones most likely to migrate. The decision to migrate varies widely and deciding causality between economic 'pull' and environmental 'push' is often highly subjective. Moreover, dis-aggregating the role of climate change from other environmental, economic and social factors requires an ambitious analytical step into the dark. By interacting with and aggravating already-existing issues with food security, water shortages, and the productivity of lands, climate change puts a strain on the ability of many diverse groups to adapt. When the land can no longer support livelihoods, people are compelled to move to places with better prospects.

Sudden shocks like floods, hurricanes, rainfall, and temperature anomalies are environmental factors that are exogenous to the individual or the household making the migration decision. They can have major destructive effects as well and migration is seen as a potential adaptation strategy to environmental changes. Degradation of forests and loss of biodiversity, soil erosion, and water quality degradation are endogenous environmental changes and can have economic and social origins. They can be caused by misuse of resources, poor planning, poor infrastructure, and poor governance and monitoring. This mismanagement of resources, over-exploitation from mass industry, and industrial pollution affect climate change (change in rainfall patterns, sea-level rise, increased frequency of heat waves, and so forth, depending on location) it can be foreseen that more of the global population will be facing environmental stresses in the future, whose direct drivers are climate change, nutrient pollution, land conversion leading to habitat change, over-exploitation, and invasive species and diseases, indirect drivers to be demographic, economic, socio-political, scientific, technological, cultural and religious factors. (Millennium Ecosystem Assessment (MA), Millock (2015).

2.8.2 Climate Change and environment

Climate change

Climate is usually described in terms of the mean and variability of temperature, precipitation, and wind over a period of time, which ranges from months to millions of years. Gutmann et al. (2005) find evidence of temperature and rainfall deviations acting as push factors for migration in the US Great Plains counties during the 1930s and the combined 1950s and 1960s. For this reason, we consider rainfall data. Drought is a natural phenomenon associated with a deficit of water availability resulting from low precipitation compared to long-term average. Drought has become more frequent and severe in arid and semi-arid lands (ASALS) than in humid areas. Drought is a disaster that affects large areas and for a longer period compared to other natural disasters such as floods. [29]

We extract rainfall data from the Africa Rainfall Climatology version 2 (ARC2), a project of the National Oceanic and Atmospheric Administration's Climate Prediction Centre (NOAA-CPP). For the period January 1983-July 2013, ARC2 delivers public re-elaborated data on the amount of rain decreasing over 10-day intervals with a spatial resolution of 0.1 degrees. Data on the maximum, minimum, and average temperature are delivered with a spatial resolution of 0.25 degrees from the European Centre for Medium-Range Weather Forecast (ECMWF). They provide two databases both recording temperature with a time resolution of 10 days. SPI data shows observed precipitation as a standardized departure from a selected probability distribution function that models the raw precipitation data. The raw precipitation data are typically fitted to a gamma or a Pearson Type III distribution, and then transformed to a normal distribution. The SPI values

[29] Rainfall variability at a time scale from years to days is as much a characteristic of climate as the total amounts recorded. Low values, however, do not necessarily lead to drought, nor is drought necessarily associated with low rainfall Gommes and Petrassi (1996). The meteorological drought is associated with long time intervals of significantly low or no precipitation and increased air temperature. The deficiency in rainfall leads to low infiltration, decreased runoff, and groundwater recharge.

can be interpreted as the number of standard deviations by which the observed anomaly deviates from the long-term mean. The SPI can be created for differing periods of 1-to-36 months, using monthly input data.[30] In Figure 9 we show deviation in precipitation (SPI Index).

The link between climate change and environment

Agricultural systems face new challenges from climate change, through increased frequencies and severity of extreme events such as droughts and intense precipitation, as well as shifts in the temporal distribution of rainfalls. Vulnerable populations in developing countries are particularly exposed to financial and technical dangers due to insufficient capacities to manage climate risks (IPCC, 2014). Lacking resources to adapt, developing countries are also the most vulnerable to future warming (World-Bank, 2009). According to Barrios et al. (2006), agricultural drought occurs when the water supply is insufficient to cover crop or livestock water. In addition to reduced rainfall, several factors may lead to agricultural drought, some of them not always obvious. This "invisible" agricultural drought prevents farmers at the subsistence level from achieving regular and high yields. "Invisible" drought is brought about by environmental degradation as much as by climate requirements. The African continent has a long history of rainfall fluctuations of varying lengths and intensities.

When determining whether a change in the local climate will cause people to move, it is important to take into account several variables, including the type of climatic shock, the characteristics of the population affected, the various socioeconomic implications, and the availability of alternative coping mechanisms. Increasing greenhouse gas emissions brought on by human activities are the primary reason for the Earth's climate is changing fast. The "poorest nations" will be particularly heavily struck by climate change, and millions may be driven even farther into poverty (Stern, 2006). Focusing on the driving force behind environmental migration can greatly benefit policymaking. Climate change presents many critical challenges to

[30] Hayes et al, An evaluation of the standardized precipitation index, 2010

humankind, one of them being the displacement of large numbers of people. The prospect of global mass migration induced by climate change in Stern (2006) is predicted 200 million 'environmental migrants' or 'climate refugees'. Growth and development may be seriously impacted by climate change. Especially in Africa, the development difficulties of providing food security and alleviating poverty will be considerably exacerbated by the effects of climate change on agriculture. Particularly in comparison to the industrialized world, Sub-Saharan Africa will be far more sensitive to slow-onset climate change and increasing climatic variability (Thornton et al., 2009).

Environmental policies

Sudden climate migration could exacerbate a range of problems, including deterioration of ecosystems, slowing of regional economic development, and increased internal or international conflicts (Gleditsch et al., 2007). The current estimates from IPCC (2014) suggest that a temperature increase of 2-3 degrees Celsius will potentially raise the number of people at risk of hunger by 30-200 million. Varieties with greater drought and heat tolerance, improved and expanded irrigation systems, rainwater harvesting technologies, disaster relief efforts, and insurance programs will likely all be needed to foster agricultural development and adaptation to warming. The existing theoretical and empirical research makes clear that migration is only one of several potential strategies used by households to cope with environmental change. The use of migration as an adaptation strategy depends on the frequency of exposure to droughts and the existence or not of public policy aimed at reducing the risk of a disaster or assistance after the event. Many authors have discussed the potential for environmental degradation to displace 'environmental refugees', with some estimating the number of those displaced in millions (Myers (2002), Hugo (1996), Westing (1992)). Climate change influences soil moisture levels by direct climatic effects (precipitation, temperature effects on evaporation), climate-induced changes in vegetation, plant growth rates, rates of soil water extraction by plants, and the effect of enhanced CO_2 levels on plant transpiration. Changes in soil water fluxes may also feed back to the climate itself and even may contribute to drought conditions by decreasing available moisture, altering circulation patterns, and increasing air temperatures. The drivers of climate change such as moisture, temperature, and CO2 are expected to have variable effects on various soil processes and properties having relevance to soil fertility and productivity. Policies that take into account and put a stop to systems and processes that degrade the climate and environment are the adequate ones to be proposed.

2.8.3 Plot Variables Description

Subjective Land Quality

Land quality is the ability of the land to perform its function of sustainable agriculture production and enable it to respond to sustainable land management. Some of the indicators in the LSMS Malawi dataset, that help understand the quality of land are slope, elevation, total wetness level and index and fallowness.

We calculate subjective land quality in a survey where respondents have different size parcels. Thus we create weights for their size. This weight is the size of the plot out of total sum of parcels. We have both subjective answer from farmers and objective data from satellites. In cases when farmers do not know the size of their parcel, we consider the gps satellite size. We multiply the weights that we constructed, with the perceived quality of lands, in order to have a weighted quality of land measure. We take the mean value of the land quality of all parcels of the household. We do this for the wet and the dry season, years 2010, 2013 and 2016. For the permanent crops there is data only for 2010 and 2013 and there is no subjective quality of land indicator.
We keep only the rainy season since it is the main one and also because there are not all the indicators in the dry season (support it with literature). Their irrigation system lies only on the rain and not water reserves. These subjective data often aim at capturing the overall soil quality level (in a categorical manner) as well as key soil quality indicators, such as soil color and texture. One might assume that respondents who spend ample time working on the land would be able to assess the health of the soil with reasonable accuracy. Soil health, however, is a highly complex subject, and this assumption can be

misguided.[31] The objective land quality uses geo plot variables provided from the world bank but collected from other satellite data sources for the year 2013 and 2016. We consider elevation, slope and total wetness index as indicators of land quality. We construct variables of these indicators according to the weight of the parcel index. We then sum up all the indexed parcels.

Objective Land Quality

The absence of regular, high-quality data on soil health and how it is evolving under previous and present management has a significant impact on the study of agricultural production that is currently being done. We utilize the Harmonized World Soil Database (HWSD), a 30 arc-second raster database that contains over 15000 distinct soil mapping units and incorporates global regional and national updates of soil information from the European Soil Database, SOTER, Soil Map of China, and WISE. Utilizing a standardized structure, they may link attribute data with GIS to display or query the composition in terms of soil units and the characterization of certain soil parameters (organic Carbon, pH, water storage capacity, soil depth, cation exchange capacity of the soil and the clay fraction, total exchangeable nutrients, lime and gypsum contents, sodium exchange percentage, salinity, textural class and granulometry).

Traditional wet chemistry techniques for soil nutrient extraction are part of conventional soil analysis, which is regarded as the gold standard in soil testing. Basic physical analysis, such as the determination of water-holding capacity, is also included. The Mehlich 3 extractant technique is among the most used

[31] Several factors that might affect the credibility of subjective results are: a) subjective assessments of soil health should be provided at the plot level rather than farm level b) appropriate respondent(s) may not be available to answer questions during the time that the survey team or the enumerator will be visiting the associated enumeration area (EA), the use of proxy respondents will be one of the factors mediating the reliability of the information sought. c) in rural areas where mobility is limited, a respondent's reference is only the soil in and around his or her farm. d) It may not be clear to respondents that the intention is to isolate the quality of the soil itself and not other plot characteristics or production outcomes. Source: LSMS guidebook, Spectral Soil Analysis & Household Surveys from S.Gourlay, E. Aynekulu, C. Carletto, and K.Shepherd, World Bank

methods for extracting nutrients in conventional wet chemistry. Micronutrient and macronutrient levels are estimated using the Mehlich 3 technique.[32]

We use as indicators of soil quality the following: [33] and uses parameter scale from 0 to 100% divided in 'no or slight constraint', 'moderate constraint', 'severe constraint', 'very severe constraint', 'not suitable'. The rating system is adapted from Sys et al[34] The indicators are the following:[35]

1. Nutrient availability: Soil texture, soil organic carbon, soil pH, total exchangeable bases
2. Nutrient retention capacity: Soil Organic carbon, Soil texture, base saturation, cation exchange capacity of soil and of clay fraction
3. Rooting conditions: Soil textures, bulk density, coarse fragments, vertic soil properties and soil phases affecting root penetration and soil depth and soil volume
4. Oxygen availability to roots: Soil drainage and soil phases affecting soil drainage
5. Excess salts: Soil salinity, soil sodicity and soil phases influencing salt conditions
6. Toxicity: Calcium carbonate and gypsum
7. Workability (constraining field management): Soil texture, effective soil depth/volume, and soil phases constraining soil management (soil depth, rock outcrop, stoniness, gravel/concretions and hardpans).[36] [37]

---

[32] Mehlich, 1984

[33] The classes used in the Soil Quality evaluation are: 1. No or slight limitations, 2. Moderate limitations, 3. Sever limitations, 4. Very severe limitations, 5. Mainly non-soil, 6. Permafrost area, 7. Water bodies.
Classes 1 to 4 are corresponding to an assessment of soil limitations for plant growth. Class 1 is generally rated between 80 and 100% of the growth potential, class 2 between 60 and 80%, class 3 between 40 and 60%, and class 4 less than 40%.

[34] Sys, C., Van Ranst, E., Debaveye, Ir.J Beernaert, F. 1993. Land evaluation, part III. Crop requirements. Agriculture publication, no.7, General Administration for Development Cooperation. Brussels, Belgium, 166p.

[35] Source: Food and Agriculture Organization and Harmonized World Soil Database

[36] More details to follow up in Appendix D

[37] Source: FAO, Food and Agriculture Organization

For analyzing soil qualities, maize was selected as reference crop because of its global importance and wide geographical distribution. In the table 1 below, all the indicators of land quality we use in our study.

Table 2.7. Land quality indicators

Subjective Land Quality	Data declared from respondents of LSMS Malawi where they describe the quality of their land as good, fair and poor by using the numbers 1, 2 and 3. The higher the index, the worst the land quality.
Slope	The topography of a field greatly determines how sensitive the soil is to water erosion. The steeper and longer the slopes are in a field, the greater the soil erosion potential. In comparison to a mild slope, a steep slope will have a higher effect on the soil's composition. Organic particles rich in nutrients from the topsoil will flow down the slope as rain falls. When vegetation is limited and water levels rise, this occurs more frequently. The slope's angle toward the sun affects the soil as well but we do not have data on that. Slope <8 percent is good, 8-15 medium, 15-25 poor [38]

[38] *Schiefer et al 2015* Indicators for the definition of land quality as a basis for the sustainable intensification of agricultural production

Elevation	Higher altitude is associated to lower temperatures which further decreases the mineral rate. This can reduce litter decay, soil organic matter (SOM) decomposition rates, soil water balance and geologic deposition processes. Moreover, variation in soil properties due to topography contributes to soil physical, biological and chemical quality variations at different elevation categories. [39]
Topographic wetness index	It is constructed to predict the spatial distribution of soil moisture and surface saturation as well as to assess the impact of local topography on hydrological processes. This index is formulated as TWI = ln(a/tanb), where 'a' is the upslope contributing area per unit contour length (or Specific Catchment Area, SCA) and 'tanb' is the local slope gradient for estimating a hydraulic gradient. The computation of both 'a' and 'tanb' need to reflect impacts of local terrain on local drainage. TWI is an indicator that measures the potential on where water tend to accumulate. A high index value denotes high potential of water accumulated due to low slope.[40]

[39] *Fantaw et al., 2006b; Jing et al., 2011; Asmamaw and Mohammed, 2012*

[40] *Qin et al 2009*, An approach to computing topographic wetness index based on maximum downslope gradient

Nutrient availability (SQ1)	This soil quality is decisive for successful low level input farming and also for intermediate input levels. Diagnostics related to nutrient availability are manifold. Important soil characteristics of the topsoil (0-30 cm) are: Texture/Structure, Organic Carbon (OC), pH and Total Exchangeable Bases (TEB). For the subsoil (30-100 cm), the most important characteristics considered are: Texture/Structure, pH and TEB. The soil characteristics relevant to soil nutrient availability are to some extent correlated. For this reason, the most limiting soil characteristic is combined in the evaluation with the average of the remaining less limiting soil characteristics to represent soil quality SQ1. [41] [42]

[41] Soil Qualities for Crop Production from Food and Agriculture Organization of the United Nations

[42] Fischer, G., F. Nachtergaele, S. Prieler, H.T. van Velthuizen, L. Verelst, D. Wiberg, 2008. Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008). IIASA, Laxenburg, Austria and FAO, Rome, Italy.

Nutrient retention capacity (SQ2)	Nutrient retention capacity is of particular importance for the effectiveness of fertilizer applications and is of special relevance for intermediate and high input level cropping conditions. The ability of the soil to hold onto more nutrients against leaching-related nutrient losses is referred to as "nutrient retention capacity." Plant nutrients are stored in the soil by the clay fraction, organic matter, and clay-humus complex, which offer exchange sites. The amount of leaching varies depending on the pace at which soil moisture drains through the soil profile. Through its impacts on the accessible exchange sites on the clay minerals and through soil permeability, soil texture influences nutrient retention capacity in two ways. The soil characteristics used for topsoil are respectively: Organic Carbon (OC), Soil Texture (Text), Base Saturation (BS), Cation Exchange Capacity of soil (CECsoil), pH, and Cation Exchange Capacity of clay fraction (CECclay). Soil pH serves as indicator for aluminum toxicity and for micro-nutrient deficiencies. The most limiting of these soil characteristic is combined with the average of the remaining less limiting soil characteristics to estimate nutrient retention capacity SQ2. Source of information in footnote 40 and 41.

Rooting conditions (SQ3)	Rooting conditions include effective soil depth (cm) and effective soil volume (vol. %) related to presence of gravel and stoniness. The presence of a soil phase may alter rooting conditions by reducing the effective volume available for root penetration or by restricting the effective rooting depth. Rooting conditions discuss numerous connections between crop development and soil conditions in the rooting zone. The assessment takes into account the following variables: Adequate foothold, or enough soil depth for the crop to anchor; available soil volume and penetrability for roots to extract nutrients; space for root and tuber crops to expand and produce an economically viable amount of crop in the soil; and absence of shrinking and swelling properties (vertic) affecting root and tuber crops. Root penetration is impacted by soil depth/volume restrictions, which might also limit yield production (roots and tubers). The following pertinent soil characteristics are taken into account: soil depth, soil texture/structure, vertic properties, gelic properties, petric properties, and the presence of coarse particles. This soil quality is calculated by multiplying the soil depth restriction by the most restricting attribute of the soil or soil phase. The soil phases that are important for rooting conditions vary slightly depending on the source of the soil map and the kind of soil being used. In the HWSD (FAO 74) soil phases are: stony, lithic, petric, petrocalcic, petrogypsic, petroferric, fragipan and duripan. Other soil phases are (FAO 90): rudic, lithic, pertroferric, placic, skeletic, fragipan and duripan. ESB soil phases and other soil depth/volume related characteristics: stony, lithic, petrocalcic, petroferric, fragipan and duripan, and presence of gravel or concretions, obstacles to roots (6 classes), and impermeable layers (4 classes). Source of information in footnote 40 and 41.

Oxygen availability (SQ4)	The drainage properties of soils have a key role in determining oxygen availability in soils. Based on methods established at the FAO (FAO 1995), which consider soil type, soil texture, soil phases, and terrain slope, soil drainage classes are determined. In addition to drainage parameters, soil and terrain variables may have an impact on the soil quality and oxygen availability. Source of information in footnote 40 and 41.
Excess salts (SQ5)	Salinity might be brought on by salt accumulation. Soil salinity, which is the excess of free salts, may be determined by measuring the electric conductivity (EC in dS/m) or the saturation of the exchange complex with sodium ions (sodicity or sodium alkalinity), which is determined by the exchangeable sodium percentage (ESP). Crops are impacted by salinity because it prevents the intake of water. Low salt levels may cause crop death, whereas moderate salinity diminishes yields and inhibits growth. Sodicity influences soil structure, causing vast or coarse columnar structures with limited permeability due to salt toxicity. Saline (salic) and sodic soil phase characteristics may have an impact on crop growth and yields. When both processes occur, the most restricting combination of soil salinity and/or sodicity conditions and the occurrence of saline (salic) and/or sodic soil phase is chosen. Source of information in footnote 40 and 41.

Toxicities (SQ.6)	Low pH causes different shortages, such as phosphorus and molybdenum deficiency, as well as acidity-related toxicities, such as aluminum, iron, and manganese toxicities. Calcareous soils typically have micronutrient shortages, such as those in iron, manganese, and zinc, and occasionally molybdenum toxicity. Gypsum severely restricts the moisture content of the soil. Crops' susceptibility to gypsum and calcium carbonate varies greatly.[43] High calcium carbonate and gypsum concentrations and low pH are incompatible. SQ1, nutritional availability, and SQ2, nutrient retention capacity, respectively, account for acidity-related toxicities such as aluminum toxicities and micronutrient shortages. Therefore, only toxicities associated with calcium carbonate and gypsum are included in this soil quality SQ6. For the purpose of quantifying SQ6, the excess calcium carbonate and gypsum present in the soil, as well as the presence of petrocalcic and petrogypsic soil phases, are considered to be the most restricting factors. Source of information in footnote 40 and 41.

[43] FAO, 1990; Sys et al, 1993

Workability (SQ7)	Diagnostic characteristics to indicate soil workability vary by type of management applied. Workability or ease of tillage depends on a variety of interrelated soil properties, including texture, structure, organic matter content, bulk density, and the presence of gravel or stones in the soil profile or on the soil surface, as well as continuous hard rock at shallow depths and rock outcrops. Heavy agricultural equipment might not be able to be used due to irregular soil depth, gravel and stones in the profile, and rock outcrops. While soil restrictions linked to uneven soil depth and stony and rocky soil conditions are primarily effecting mechanized land preparation and harvesting operations of high-level input mechanized farming LUTs, they are also especially harming low and intermediate input farming LUTs. Therefore, workability limitations for low/intermediate and high inputs are addressed differently. The physical impediment to agriculture and the cultivation restrictions brought on the texture/clay mineralogy are both included in the workability soil quality SQ7. The most restricting soil/soil phase characteristic is combined with the average of the other attribute conditions to create the soil quality SQ7. Soil phases are stony, lithic, petric, petrocalcic, petroferric, fragipan and duripan (FAO '74), and lithic, petroferric, rudic, skeletic, duripan and fragipan (FAO '90). Source of information in footnote 40 and 41.

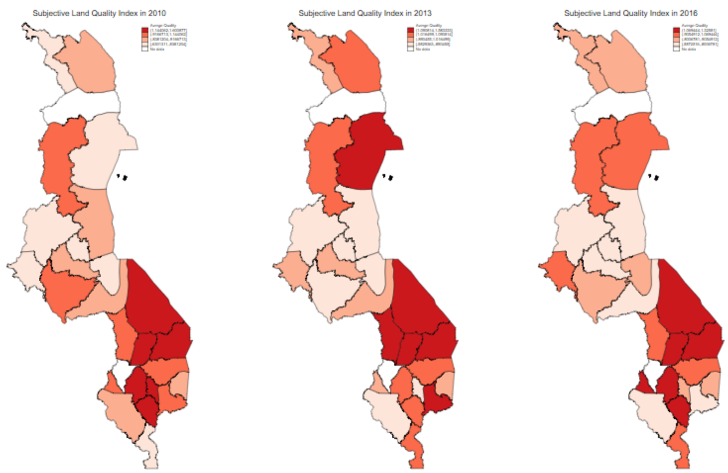
2.8.4 Summary Statistics

Table 2.8. Summary Statistic of all variables

Label	Description	Min	Max	Mean	St.dev
migration					
migrants	Dummy where hh has a migrant that has left in the past 6 months	0	1	0.06	0.24
hhmigrant	Number of migrants per hh that left in the past 6 months	0	8	0.09	0.46
resettle	Dummy if the household has resettled since the previous wave	0	1	0.05	0.21
land specification					
landown	Land owned in hectares	0	3.67	0.56	0.5
landless	Household does not hold any land	0	1	0.08	0.27
slope	Slope of the territory where household owns land	0	47	4.5	4.7
elevation	Elevation of the territory where household owns land	0	1610	842	343
twi	Total wetness Index	0	28	13.4	3
sq1	Nutrient availability	1	7	1.64	1.45
Household characteristics					
hhsize	Household Size	1	20	5.16	2.31
urban	Dummy where household lives in an urban area	0	1	0.25	0.43
femhead	Dummy where household head is a female	0	1	0.23	0.42
agehead	Age of household head	16	113	45	15
educave	Average years of education in the family	0	18.5	5.56	2.86
hhlabor	Household members older than 14 and younger 60 years old	0	9	2.45	1.38
wealth	Non-agricultural wealth index	-0.76	12.6	0.21	1.33
agwealth	Wealth scores	-1.61	7.38	-0.01	1.02
dist	Household distance from a main road	0	46	7.75	9.34
credit	Household is credit constrained	0	1	0.56	0.5

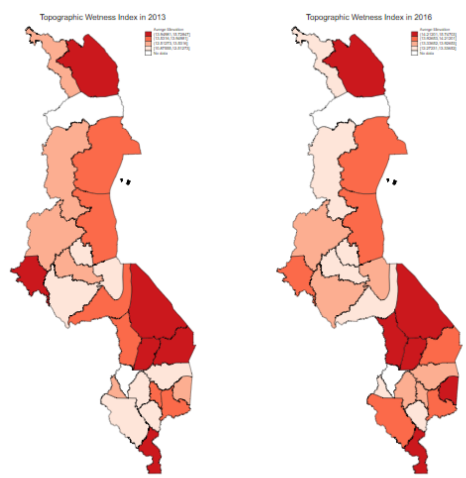
2.8.5 Maps

Subjective Land Quality and TWI



(a) 2010 **(b)** 2013 **(c)** 2016

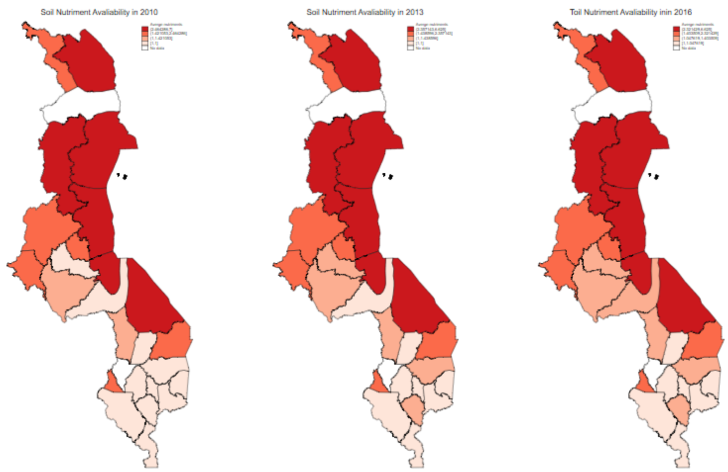
Figure 2.3. Land quality



(a) 2010 **(b)** 2013

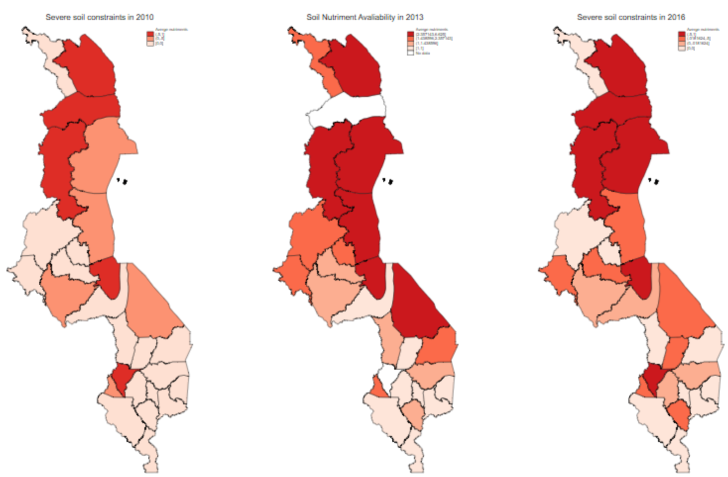
Figure 2.4. Total wetness index

Objective land quality



(a) 2010 **(b)** 2013 **(c)** 2016

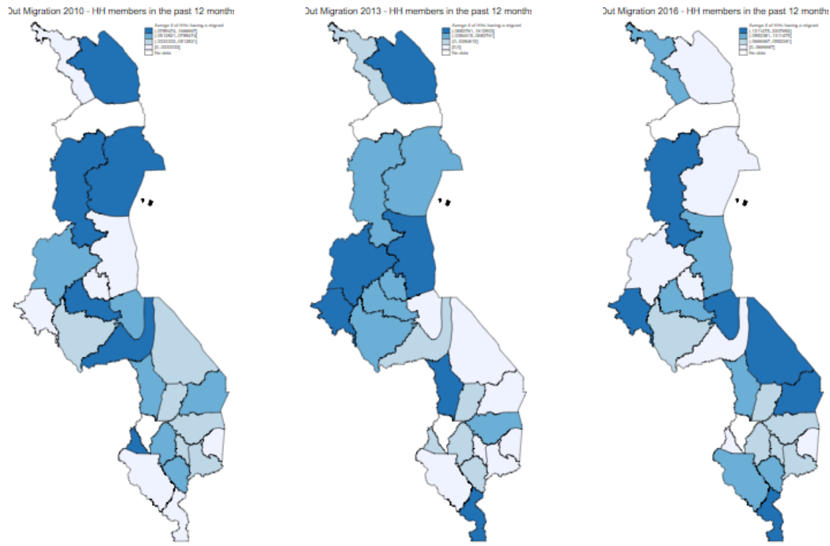
Figure 2.5. Objective land quality (nutrient availability)



(a) 2010 **(b)** 2013 **(c)** 2013

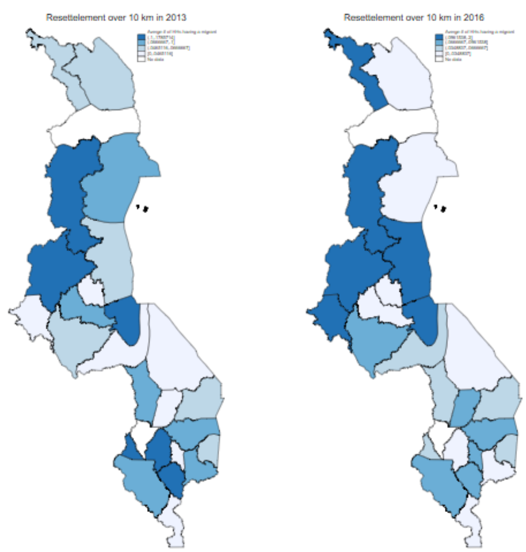
Figure 2.6. Objective land quality (nutrient availability) dummy with 'severe' or 'not severe constraints'

Migration



(a) 2010 **(b)** 2013 **(c)** 2016

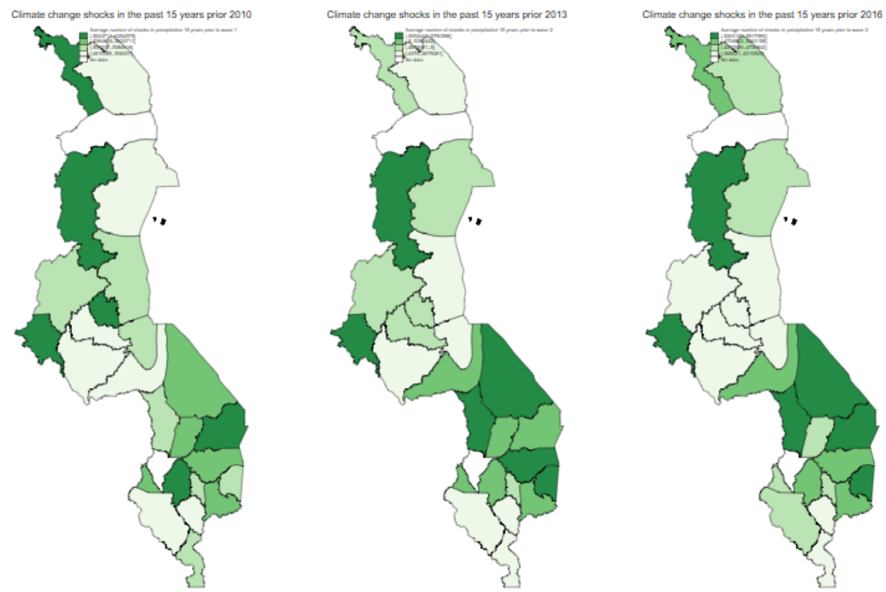
Figure 2.7. Who left in the past 12 months



(a) 2013 **(b)** 2016

Figure 2.8. Resettle: households that resettled in more than 10 km

Climate change



(a) 2010 **(b)** 2013 **(c)** 2016

Figure 2.9. Deviation in precipitation SPI

3

Integration Vs Cultural Persistence: Fertility and Working Time among Second-Generation Migrants in France [1]

Co-authored with Thomas Baudin and Maeva Planard

[1] We would like to thank Simone Morriconi, David de la Croix, Jean-Francois Maystadt, Luca Pensieroso, Stefanija Veljanoska, Bastien Chabé-Ferret and the participants of the Doctoral Workshop in UCLouvain for suggestions and great discussions.

Abstract
We study whether cultural norms in the origin country, measured at different times, affect fertility and labor force participation of second-generation migrant women in France. We investigate empirically and follow an epidemiological approach to test that the culture of origin affects people's behavior and decisions. We use the dataset TeO (Trajectoires et Origines) on population diversity in France in 2008. We find that: 1) cultural norms affect people's fertility and labor working time decisions, confirming the results of Fernandez and Fogli (2009) also for the French context; 2) the timing when the norm is measured is crucial. The later the norm is measured in time, the most powerful its effect, suggesting that the effect of the norms transmitted from peers is stronger than that of norms transmitted from parents. The explanatory power of norms holds also when controlling for socio-economic characteristics such as age, siblings, education of the respondent, spouse, and parents; 3) the feeling of being French moderates the persistence of cultural norms differently for fertility and labor force participation, while the perceived feeling of being discriminated does not alter the persistence of the cultural norms.

3.1 Introduction

People incur considerable economic costs to comply with cultural norms. According to Tylor (1871), culture is defined as the social behavior and norms found in human societies, as well as the knowledge, beliefs, arts, laws, customs, capabilities, and habits acquired by man as a member of society. Culture serves as a guide for the behavior and expectations of certain social groups. It gets transmitted through social learning and interactions and it can affect human decisions and economic outcomes. Considering this and relying on the epidemiological approach, in this study we show that the results of Fernandez and Fogli (2009), where cultural norms affect fertility and labor force participation of second-generation female migrants in the USA, stand also for the French context. We nuance this result by showing how the timing of the norm can be crucial in the transmission of culture.

Additionally, we document how perceived discrimination and feelings of integration alter the effect of the norms on fertility and labor time decisions. In order to investigate cultural transmission, we follow an empirical investigation and apply it to France using the exceptional dataset Trajectoires et Origines (TeO) where on top of labor force participation and the number of children, there is information on the migration history of the respondent, the feeling of discrimination and integration. We measure cultural norms in the origin country at different times, to capture the preferences and beliefs regarding fertility and labor working time in the origin country throughout the life span of the respondent. Second-generation woman migrants are born and raised in France, with the same rights and legal obligations as the other French woman but with different cultural backgrounds. Their parent's country of origin is different from France leading to the presence of different cultural norms in the household. Second-generation French women are influenced by both their parent's culture and French culture.

We focus on the dimension of culture that is transmitted from strong ties with the origin country. This includes transmission through first-generation migrants towards second-generation ones. It can also include the effect of peers, which is the impact of the present society in the origin country. In particular, we use as cultural proxies total fertility rate (TFR) and labor force participation (LFP) in the origin country of the parent during different times: i) the year of birth of the second-generation French woman; ii) the time when they finish their studies; iii) in 2008, the year of TeO survey. We choose these cultural proxies because they indicate heterogeneity of culture in different countries. A given women's decision depends on her preferences and beliefs which are a composition of French values and parents' origin country values. The TeO survey allows to study the impact of origins on the living conditions and social trajectories of second-generation immigrants, while taking into account other sociodemographic characteristics, namely social environment, neighborhood, age, cohort, sex, and educational level. TeO asks questions on feelings of discrimination and integration perceptions such as the feeling of being French. We analyze how these perceived measures moderate or magnify the persistence of cultural norms from the origin country. We complement this dataset with other data on fertility and labor market participation of the origin country provided by the World Bank. The access to such rich data allows understanding how the culture of different origin countries, that differ in economic and demographic situation, affect second-generation immigrants decision on the number of children and labor time.

We find that norms measured in more recently explain better fertility and labor working time, compared to norms of origin country measured at the time of birth. The norm that explains better fertility behavior is the one measured at the end of the education, and for labor time, it is the one measured at the time of the interview.[2] This implies that the effect of peers in transmitting culture can be more effective than that of parents, suggesting that horizontal transmission of culture is stronger than vertical one. Moreover, discrimination does not mediate the effect of norms on the behavior of second-generation immigrants, while integration indicators like feeling French moderates the persistence of the norms. Feeling French, serving as an integration measure, makes second-generation women work 0.12 hours more a week and have 0.2 more children. This perceived integration moderates the persistence of the norm by decreasing the power of the norms from the origin country. On the contrary, perceived discrimination does not alter the effect and power of the cultural norm. Our main results are robust to alternative specifications like a Hurdle Model for labor working time and a Poisson Model for fertility. Results are maintained when we limit our analysis to women who have finished their fertility cycle (above 40), samples for both men and women, or considering both parents' country of origin. These findings contribute to the literature on cultural persistence of origin country cultural norms on second-generation migrants by filling the gap about which norms can explain better behavior and how discrimination versus integration alter the effect of these norms.

Our paper is organized as follows. Section I contains a brief review of the

[2]Correlation of number of children with TFR at the end of education in origin country it is 30%, and hours worked with LFP in 2008 in origin country is 17%. These are also the highest values compared to correlations with other measures of fertility and labor force participation norms

empirical literature. Section II describes the data, and Section III contains an institutional context of culture. Section IV develops the empirical strategy. In section V we summarize our results and we conclude in Section VI following also a discussion.

3.2 Literature Review

In this section, we review the literature that focuses on the impact of culture on migrants' fertility and labor market participation. Sociologists and philosophers have studied and written about the cultural interlink with economic outcomes. Although the differences in their opinions, what their views had in common is that culture matters in decision making. We write more about the interlink between culture and economic decisions in Appendix K. A thorough literature review of a historical perspective of culture can be found in Guiso et al. (2006).

3.2.1 Cultural impacts on fertility

Culture is an important aspect of the decision making and it affects economic outcomes through beliefs and preferences. One of these decisions is also fertility and how many children to have. Traditionally, the economic literature sheds light on the role of prices and incomes as determinants of fertility, while the demographic literature discusses whether assimilation to local norms or disruption effects, influences fertility. Both disciplines agree that there is a strong interlink between culture and economic outcomes. For the U.S. context, a study by Fernandez and Fogli (2009) studies the cultural transmission of origin country on fertility and labor force participation of second-generation immigrants. They find that cultural norms have a positive significant explanatory power even after controlling for education and spousal characteristics and that results are not explained by unobserved human capital. Assimilation or transmission of these cultural norms depends on family social capital, structure, and values. Considering this, in another study Fernandez and Fogli (2006) show that culture measured by the total fertility rate in the woman's country of ancestry, and family experience, measured by the woman's number of siblings, are both significant determinants of fertility, even after controlling for several individual and family-level characteristics. Our results of cultural norms' effect on fertility are in the same line and magnitude.

Gender roles of the origin country can be powerful in affecting people's behavior, through a vertical transmission of cultural norms. Nevertheless, immigrants can integrate in the origin country and their fertility, education, and labor market outcomes converge to the local values (Blau (2015)). In general, as immigrants adapt, they are less and less influenced by their origin social environments ((Kahn (1988))). This difference in vertical and horizontal transmission of norms depends on many factors. In the U.S. for example, the high fertility of immigrants, compared to natives can be explained by compositional differences with respect to age, education, income and ethnicity (Kahn (1994)). Although immigrants are 'expected' to have higher fertility than similar natives, they tend to adapt their fertility 'goals' over time, both within and across generations. We consider and explore this vertical and horizontal transmission of culture in our study by focusing on the heterogeneity of the second-generation immigrants and their country of origin. The difference in cultural norms of the origin country measured at different points in time affects immigrants' decisions making differently. Depending on the strength of adherence to home country norms, immigrants coming from high-fertility origin countries, can either converge to destination country fertility levels or continue to have higher fertility than native-born women.[3] Deviating from the economic optimum comes at a cost, so not always cultural influence prevails.[4] Other observed differences can arise from the sociocultural distance between parents' country of origin and the host country, as well as from structural determinants. For the French context, Pailhé (2017) finds that women of Southeast Asian descent deviate from the fertility pattern of their parents,

[3] Behrman and Weitzman (2019), Afulani and Asunka (2015)

[4] Originating from a high fertility origin country and being an immigrant in France, correlates to having larger families but it does not predict earlier entry into motherhood (Chabé-Ferret (2019)).

while those of Turkish descent preserve their parents' cultural heritage. The similarity of this study to ours is in the fact that the author uses the same dataset and studies the fertility behavior of second-generation immigrants but the study does not account for the timing of the cultural norm.[5]

3.2.2 Cultural impacts on labor market outcomes

Furthermore, we exploit how culture stands as a determinant in explaining labor working time. This is relevant because culture through labor market outcomes can have important effects on the state of the economy and growth (Weil (2004)).[6] Other relevant characteristics for labor market outcomes are age, number of siblings, education, partner and parents' education, and gender. Culture can also serve as a determinant in the gender equality of the labor force. Considering this, we distinguish between second-generation male and female immigrants. Their behavior in the workplace can differ due to their cultural background. Labor market gender roles among migrants can be transmitted from the country of origin (Alesina et al. (2013)).[7] Culture of the country of origin can explain entirely female labor market outcomes (Read (2004)). On the contrary, source country female LFP rates can extend beyond gender role attitudes. These attitudes can contribute to labor market skills which help immigrant women navigate the host country's labor market (Frank and Hou (2016)). Labor market outcomes are country specific. In the U.S., second-generation immigrants can exhibit integration in the labor market. Fogli and Veldkamp (2011) and Fernandez (2013) show that female

[5]The author analyses the transition to first, second, and third births among different groups of immigrants daughters (from the Maghreb, sub-Saharan Africa, Turkey, and Southeast Asia) and compares them to native-born women using discrete-time logistic regressions, also using the TeO dataset.

[6]More on the literature about the interlink between culture, institutions, and economic outcomes can be found in Alesina and Giuliano (2015)

[7]In this study, second-generation immigrants, with a heritage of traditional plough use, exhibit less equal beliefs about gender roles today. This was measured using reported gender-role attitudes and female participation in the workplace, politics, and entrepreneurial activities.

labor participation in the U.S. has increased throughout time due to the assimilation effects of intergenerational social learning of immigrants. In France, other cultural differences can be seen in family ties. Children of immigrants stay significantly longer in the parental home, partly because their parents come from societies characterized by strong family ties, and partly because they have greater difficulties in becoming economically self-sufficient (Ferrari and Pailhé (2016)).[8] Cultural influence depends on the strength of the norms of the origin country, as well as the host country (Kok et al. (2011)).[9] Depending on the length of stay of immigrants, the effect of culture might differ. The difference between women from high and low participation countries diminishes with the length of residence due to the persistent effect of the host country (Neuman (2018)).

3.2.3 Discrimination and Integration

When we speak about multiculturalism and immigration, racism and discrimination are present and well documented.[10] Although social mobility does exist from one generation to another, social status among second-generation immigrants changes more slowly, despite a general improvement in education and occupational positions (Simon (2003)). There is evidence that second-generation immigrants in France are not treated equally in between them, both in the private and public sectors (Berson (2009)). However, the wage gap is determined by characteristics like the number and gender of immigrant parents and not only by the country of origin. The study of integration and discrimination is relevant because of

[8]For children of immigrants from North Africa, especially women, the entry into adulthood is slower and is less marked by union formation, whether cohabitation or marriage. Moreover, Ferrari and Pailhé (2016) finds that children of immigrants from Southern Europe behave more like native French.

[9]Kok et al. (2011) find that both differences in home-country female participation and the trend in native female participation in the Netherlands affect the participation of migrant women.

[10]Especially after 1980 when many parties in local and national elections were making immigration part of their main campaign issues.

evidence that persistent exposure to discrimination can lead individuals to internalize the prejudice or stigma that is directed against them. [11] Non-European origin in Europe is associated with greater disadvantage in finding employment not only among first-generation immigrants but also among second-generation (Gorodzeisky and Semyonov (2017)). The likelihood of attaining a high-status job can be influenced mostly by the immigrant status, regardless of region of origin and gender. In sum, discrimination matters for labor force participation, which is also what Baudin et al. (2021) find.

Our study contributes to the field of literature that studies how culture affects behavior, and how culture can persist in doing so intergenerationally. We not only exploit the hypothesis that culture in the origin country affects behavior related to fertility and labor working time but we specify which cultural norm is more important. Our paper focuses on the intersection of literature on cultural transmission in second-generation migrants' behavior and the literature on discrimination against migrants. We exploit how perceived discrimination or integration affects the cultural transmission of norms from the country of origin to the destination country.

3.3 French Institutional context

France has a long history of migration. The big inflow of people after war world I and II to compensate for casualties and the lack of labor shortages, led to a high number of immigrants. Nevertheless, in the late 20th century, European Union countries with France included became more resistant to the admission of asylum seekers. During economic downturns waves of racial discrimination and social opposition were present in the political environment. The French history of immigration is rather complex. In Appendix L we provide a detailed historical context of French Immigration and Baudin et al. (2021) do a detailed

[11] This, later on, can affect their behavior. According to United Nations in "Prejudice and discrimination: Barriers to social inclusion", discrimination affects people's opportunities, their well-being, and their sense of agency. Discrimination is prejudice based on concepts of identity, and the need to identify oneself as being part of a certain group.

summary of it.

Considering this big flow of immigrants, there are also a high number of their descendants in France. It is important to study the behavior of second-generation immigrants considering their share of the population. According to the 2006 French census, 25% of the population have at least one parent or grandparent who immigrated to France (INSEE 2008). In 2008, an INSEE census estimated that 5.3 million foreign-born immigrants and 6.5 million direct descendants of immigrants (born in France with at least one immigrant parent) lived in France, representing a total of 11.8 million people, or 19% of the total population in metropolitan France (62.1 million in 2008). 40% of immigrants live in the Parisian urban area (Greater Paris) making it the area with the largest proportion of immigrants. According again to INSEE, in 2008, there were 12 million immigrants and their direct descendants (2nd generation) making up about 20% of the population. There were also 6.7 million second-generation immigrants (with one parent born outside of France) living in France in 2008, corresponding to 11% of the total population. Immigrants aged 18-50 count for 2.7 million (10% of the population aged 18-50) and 5.3 million for all ages (8% of the population). Second-generation immigrants aged 18-50 make up 3.1 million (12% of 18-50) and 6.5 million of all ages (11% of population). TFR in France in 2007 was 1.98. Social and family policies seemed to have played a stabilizing role to keep TFR at this threshold. Fertility remained at a high level between 2008 and 2014. France has been leading the European countries since 2012 (INSEE 2021). Many migrants come from high fertility countries, where the age at first birth is low, which is especially the case in sub-Saharan Africa. Accounting for the different ethnic compositions, French language and cultural immersion courses have been long considered integration measures.[12] Considering that discrimination and integration are part of migrants' trajectories, we focus furthermore on these perceived feelings. We interpret the feeling of discrimination as not being accepted while the feeling of being French as belonging. We investigate how discrimination and

[12]The number of assigned hours of training significantly increases the labor force participation. The language classes appear to have a larger effect on labor migrants and refugees relative to family migrants (Lochmann et al. (2018)).

the sense of belonging interfere with cultural norms when studying their effect on fertility and labor time in second-generation migrants. Integration in France is perceived as a process in which immigrant characteristics uniformly converge toward the average characteristics of French society. Nevertheless, integration and feeling French are two different levels not to be mixed. One does not automatically mean the other one holds. This difference between the two concepts can serve as a mechanism to detect if people behave differently in terms of their fertility choice and labor market participation. This is important for second-generation migrants where one direction of integration does not exclude the feeling of being of national of a certain country.[13]

3.4 Data

We study fertility and labor working time of second-generation immigrants in France by exploiting the Trajectories and Origins (TeO) survey on population diversity in France.[14] This survey provides information on living conditions and social trajectories while taking account of other sociodemographic characteristics, like social environment, neighborhood, age, cohort, sex, and educational level. It also provides information on parents' origin and level of education. Moreover, this survey provides data on discrimination linked to origin and integration. There are 24000 respondents and the survey was carried out during October 2008-February 2009, out of which 9600 native descendants of immigrants.[15] TeO covers all populations living in

[13]For the population of European ancestry, the feeling of belonging, while weaker at the beginning, gets stronger with time and it is correlated with the feeling of being accepted. On the contrary, for people of African and Southeastern Asian ancestry, the feeling of being French is always stronger than the feeling of being accepted. This gap persists with the acquisition of nationality, not only among the first generation but also among the second-generation immigrants (Jayet (2016)).

[14]The TeO survey was conducted jointly by INED and INSEE. It was closely supervised by the official bodies which oversee the collection and use of public statistics (CNIS; CNIL)

[15]The respondents of the TeO dataset are as follows: 9600 immigrants, 9600 native descendants of immigrants, 800 people from the DOM-TOM/COM, 800 native descendants of DOM origin, 3200 native descendants of native-born

metropolitan France (mainland and Corsica). We chose this survey because it provides immigrant trajectories, immigrants' descendants, demographic characteristics, and measures of perceived discrimination and integration. The TeO survey has all our variables of interest which include: fertility, labor force participation, parent's origin, personal and partner characteristics, perceived discrimination, identity, and identification in terms of national belonging perceptions. In France, national surveys of integration and discrimination are sparse and recent and including all the other indicators that we need for this study, we can narrow them down to only TeO. We use the first and only wave available in this dataset.[16]

We restrict our sample to women aged above 18, married, and who have finished their studies. Our variables of interest are fertility, which we measure by all births ever, and female labor force participation, which we measure by hours worked. We define hours of work by considering 3 measures: i) the number of hours worked listed as family-support (aide familiale) and independent workers (personnes non-salariees); ii) number of hours worked in regular employment for active and employed people, in case the first measure is missing; iii) the number of hours worked of people who have multiple jobs and thus more hours worked in total than in their regular job. We limit hours worked to 80 by discarding this way 0.05% of our sample.

Other variables that we consider from this survey are age, education, spouse's age and education, parents education, number of siblings, and regional dummies. For age, we create age dummies for different age groups and we base our measure on age group '40-50'. We consider a polynomial on age for both respondent and partner (age and squared age). In doing so we take into account the difference between the year when the norm was measured in the origin country and the year when the respondent answered the survey. The estimation difference, depends on the age difference when the norm was measured, for example, two respondents of Italian and Algerian origin that were interviewed in the year 2000 compared to 1999, or 2 Senegalese women whose norms are measured at different times.

[16]The second wave is available in Autumn 2022

TeO dataset also includes data on discrimination, resource availabilities, and obstacles. We use these variables to see if individuals' origins is a parameter that can explain observed heterogeneity in behavior and decision making. The survey questions focus on the discriminatory processes that affect the careers and living conditions of immigrants and their descendants. It also hampers their social and economic integration. The key situations in which discrimination may occur are examined regarding job-seeking, dismissal, promotion, relations with colleagues, the search of accommodation as well as health degree, medical insurance, political and social participation. Discrimination is addressed in two ways: indirectly via the respondent's labor, health, and fertility situation as compared to people of different origin, and directly by asking them about their experience of discrimination. This is perceived discrimination seen from a subjective point of view. Apart from discrimination, we also consider the integration of second-generation migrants by relying on the question if the respondent feels French.

For studying how culture of the origin country influences behavior, we merge the TeO data with data from the origin country made available from the World Bank. These data give information on the total fertility rate (TFR) and female labor force participation (LFP) in the origin country of the parents of the respondent during the time that the respondent was born, finished their studies and in 2008, the year of the TeO survey. In our sample, women have on average 2.2 children, while the TFR from the origin country depending on when it is measured varies between 2.42 to 5.11. Our respondents work an average of 20 hours a week, are on average 40 years old, their partner is on average 45 years old and they have on average 4 siblings. LFP of origin country varies between 26 to 40. The variance in cultural norms tends to decrease through time for fertility but it increases for LFP. A summary of statistics of our variables of interest in the dataset is shown in Table 1. In Appendix H we also provide summary statistics by country of origin.

Table 3.1. Summary statistics at the individual level of our sample

	Obs	Mean	Std. Dev.	Min	Max
Number of children	4045	2.20	1.41	0	10
TFR at birth	2749	5.11	1.94	1.35	7.94
TFR at the end of education	3491	3.61	1.91	1.13	7.94
TFR in 2008	3503	2.42	1.20	1.34	6.80
Worked hours	4045	19.98	18.69	0	80
LFP at birth	2175	26.85	18.47	2.44	83.89
LFP at the end of education	3211	32.79	18.65	2.16	94.4
LFP in 2008	3479	40.02	19.44	13.13	81.17
Age dummies					
18-20	19	19.57	0.69	18	20
21-25	225	23.48	1.28	21	25
26-30	498	28.11	1.39	26	30
31-35	697	33.07	1.44	31	35
36-40	767	38.02	1.39	36	40
41-50	1182	45.39	2.89	41	50
More than 50	657	54.91	2.77	51	60
Age	4045	39.95	9.63	18	60
Squared Age	4045	1689.07	788.54	324	3600
Partner Age	3930	44.45	10.80	17	91
Squared Age of Partner	3930	2093.30	1009.86	289	8281
Siblings	4302	4.58	3.50	0	60

Table 3.2. Education of respondents (TeO Dataset)

	Less than Primary	Primary	Lower -Secondary	Upper -Secondary	College
Diploma of the respondent	23.09%	6.55%	9.43%	31.97 %	28.98%
Diploma of the partner	23.49 %	4.39%	6.23%	36.19%	29.70%
Diploma of the mother	68.17%	8.26%	5.07%	10.36%	8.14%
Diploma of the father	58.91%	9.74%	4.50%	13.04%	13.81%

3.5 Empirical strategy

The empirical approach that we first adopt to study the influence of cultural norms on second-generation migrants in France is shown by estimating the following model:

$$D_{ijr} = \alpha_0 + \alpha_1 \bar{y}_j + \alpha_2' X_i' + \alpha_3 g_{ijr} + \alpha_4 r_{ijr} + \alpha_5 i_{ijr} + \varepsilon_i \tag{3.1}$$

We run an Ordinary Least Square (OLS) regression where D_{ijr} is the outcome variable, either the number of children or the number of hours worked, of the respondent i, which is the second-generation migrant in France, originally from country j residing in the region r in France. $\bar{y}_j$ is our proxy for culture, which is fertility measured by TFR in the origin country when the outcome variable is fertility. It becomes LFP in the origin country when the outcome variable is hours worked.

We consider different proxies for culture: TFR and LFP in the origin country at i) the time of birth of the respondent, ii) when the respondent finished the studies, and ii) in 2008, the year when the respondent filled in the TeO survey. X_i is a set of individual characteristics that include age, partner's age, education, partner and parents' education level, number of siblings, and region. g_{ijr} shows other cultural indicators like LFP in the origin country of the respondent in the year 2008 when the outcome variable is fertility, and it takes the measure of TFR at the end of the education when the outcome variable is hours worked. r_{ijr} is discrimination indicators of the respondent in France regarding age, health, origin, place of living, religion, gender, outfit, and skin. i_{ijr} is the perceived feeling of integration of the respondent, which relates to the feeling of being French. We apply this model to the sample of married second-generation immigrants, above 18 years of age. We expand our study to other samples of all women and all population above 18 in Appendix A.

Regarding fertility decisions, apart from the OLS regression model, we also consider a Poisson model since fertility is an integer variable. Our dependent variable Y has a Poisson distribution and assumes the logarithm of its expected value can be modeled by a linear combination of unknown parameters. This is an appropriate model when the dependent variable is a count and it allows to study the impact of the TFR in the origin country, on the number of children of the respondent. Regarding labor working time, hours worked exhibit an excess of zero observations. To take into account the high number of zero in the dataset, we consider a Hurdle model. This model is used when a random variable is modeled using two parts, the first, which is the probability of attaining value 0, and the second part models the probability of the non-zero values. This method allows explaining both the number of hours worked (with a Poisson model) and the decision not to work at all (with a Logit model).

3.6 Results

3.6.1 Fertility Behavior

The effects of the TFR of the origin country, on second-generation immigrants' fertility decisions in France for our sample of married women can be found in Table 2. The table is organized into 6 models. We consider the mother's country of origin and we measure as cultural norms three variables: the TFR at the birth year of the respondent, the TFR in the origin country at the year when the respondent finishes education, and the TFR in 2008, which is the year of the TeO survey. We find that the closest the cultural norm measure is to the year of the survey, the stronger the effect on the fertility of second-generation immigrants. What seems to matter the most is the cultural norm in the origin country at the actual time of the survey, compared to the norm measured in the previous years. An increase of the TFR in the origin country by one child is associated with an increase in the respondent's fertility by 0.08 children when the norm is measured during her year of birth, it increases to 0.18 when the norm is measured at the end of her education cycle and it reaches 0.21 more children when the norm is measured in 2008. Nevertheless, in terms of magnitude of coefficients, we see that it is TFR at the end of the education that explains most of the variability of the number of children. 1 standard deviation increase in TFR at the end of education in the origin country raises number of children of second-generation migrants by 0,32 compared to 0,25 from 1 standard deviation increase of the TFR in 2008 or 0,14 from 1 standard deviation increase in TFR at birth. The results are in line with Fernandez and Fogli (2009) where TFR in the origin country measured in 1950 has a positive and significant effect on the number of children ever had. They show that women that come from countries with a higher TFR have more children. In their study, an increase in the TFR by 1, results in 0.2 more children among the second-generation migrant women.

We proceed by studying how labor working time in the origin country in 2008 alters the effect of the norm and how it affects the fertility of the respondent. We consider both models with (Models 1,2 and 3) and without this variable (Models 1.b, 2.b and 3.b). We find that LFP does not alter the effect of the fertility norm and that it affects negatively the number of children of the respondent.[17] In terms of personal characteristics, we find that age is significant in explaining fertility decisions when considering each type of cultural norm for different age groups. Respondents aged below 40 have fewer children compared to those above 40, which have finished their fertility cycle. Age of the partner is significant and it affects positively the number of children. The older the partner, the more children the partnership has. Moreover, we find a positive sibling peer-effect. The number of siblings of the respondent is significant in explaining the number of children of the respondent when we consider as the cultural norm the TFR at birth, as it is correlated with this cultural measure through the parent's culture of origin. It loses significance for the second model, and it gains back significance for model 3.b where we consider as cultural norm TFR in 2008. Having a primary education affects positively the number of children of the respondent, and the relationship changes sign when we move to the lower, upper-secondary, and college education. Partner's education on the other hand affects negatively the number of children of the respondent and is significant in the three models we consider. We also control for parent's education and regional dummies.

Considering that cultural norms have a positive and strongly significant effect on women's fertility, we check how this effect stands when it comes to

[17]We also run a Variance Inflation Factor test for different measures of culture regarding TFR and LFP and values are well below 10, excluding multicollinearity.

men. In Table 4 we show the results of the effect of cultural norms in the origin country, measured in three points on time, on the number of children of second-generation immigrant men. The results are in the same line as the results that we show for the second-generation immigrant women. Men whose mother is originally from a country with a high TFR, have more children. As for the women's analysis, the norm that shows the strongest effect is the one measured later in time.

Table 3.3. OLS Model: How fertility in origin country affects number of children of second generation woman immigrant in France

	(1)	(1.b)	(2)	(2.b)	(3)	(3.b)
TFR at birth	0.0728***	0.0883***				
	(0.0161)	(0.0127)				
TFR at the end of education			0.171***	0.184***		
			(0.0155)	(0.0120)		
TFR in 2008					0.210***	0.190***
					(0.0426)	(0.0312)
LFP in 2008	-0.00348**		-0.00635***		-0.0105***	
	(0.00123)		(0.000881)		(0.000818)	
Age						
18-20	-1.411***	-1.364***	-1.734***	-1.755***	-1.980***	-1.989***
	(0.380)	(0.391)	(0.310)	(0.308)	(0.329)	(0.318)
21-25	-1.005***	-0.987***	-1.230***	-1.271***	-1.445***	-1.489***
	(0.227)	(0.228)	(0.186)	(0.191)	(0.198)	(0.201)
26-30	-0.754***	-0.738***	-0.860***	-0.880***	-1.037***	-1.057***
	(0.212)	(0.212)	(0.149)	(0.148)	(0.160)	(0.160)
31-35	-0.283*	-0.271*	-0.307**	-0.323**	-0.431***	-0.451***
	(0.152)	(0.151)	(0.110)	(0.113)	(0.121)	(0.126)
36-40	-0.184*	-0.178	-0.155	-0.162*	-0.220**	-0.230**
	(0.106)	(0.104)	(0.0929)	(0.0896)	(0.0972)	(0.0957)
More than 50			0.0959	0.0954	0.130	0.119
			(0.147)	(0.162)	(0.141)	(0.164)
Squared age	7.32e-05	7.42e-05	-0.000105	-0.000155	-3.83e-05	-0.000119
	(0.000207)	(0.000211)	(0.000127)	(0.000121)	(0.000118)	(0.000118)
Age of partner	0.130***	0.125***	0.0573***	0.0439**	0.0559***	0.0399**
	(0.0187)	(0.0192)	(0.0173)	(0.0185)	(0.0176)	(0.0183)
Squared age of partner	-0.00134***	-0.00127***	-0.000487**	-0.000337*	-0.000457**	-0.000265
	(0.000256)	(0.000266)	(0.000194)	(0.000188)	(0.000192)	(0.000181)
Siblings	0.0192**	0.0192**	0.0145	0.0156	0.0205	0.0278*
	(0.00830)	(0.00747)	(0.00990)	(0.0101)	(0.0140)	(0.0142)
Education						
Primary	0.256**	0.260**	0.145*	0.139*	0.0923	0.0799
	(0.105)	(0.108)	(0.0752)	(0.0732)	(0.0757)	(0.0806)
Lower Secondary	-0.187*	-0.167*	-0.166*	-0.155*	-0.222**	-0.210**
	(0.0922)	(0.0905)	(0.0943)	(0.0893)	(0.0858)	(0.0848)
Upper-Secondary	-0.414***	-0.411***	-0.317***	-0.322***	-0.399***	-0.423***
	(0.0637)	(0.0677)	(0.0486)	(0.0475)	(0.0527)	(0.0540)
College	-0.669***	-0.675***	-0.473***	-0.474***	-0.609***	-0.633***
	(0.0979)	(0.0975)	(0.0829)	(0.0800)	(0.0826)	(0.0845)
Partner's Education						
Primary	-0.549***	-0.534***	-0.387***	-0.393***	-0.411***	-0.428***
	(0.108)	(0.0948)	(0.0960)	(0.0809)	(0.101)	(0.0929)
Lower Secondary	-0.232**	-0.220**	-0.310***	-0.293***	-0.316***	-0.293***
	(0.0824)	(0.0838)	(0.0844)	(0.0907)	(0.0926)	(0.103)
Upper-Secondary	-0.273***	-0.263***	-0.278***	-0.274***	-0.296***	-0.299***
	(0.0728)	(0.0726)	(0.0727)	(0.0783)	(0.0756)	(0.0836)
College	-0.341**	-0.333***	-0.338***	-0.342***	-0.374***	-0.388***
	(0.122)	(0.115)	(0.0784)	(0.0810)	(0.0769)	(0.0811)
Mother's Education	yes	yes	yes	yes	yes	yes
Father's Education	yes	yes	yes	yes	yes	yes
Region	yes	yes	yes	yes	yes	yes
Constant	-0.418	-0.537	1.155**	1.220*	1.471**	1.548**
	(0.466)	(0.475)	(0.553)	(0.595)	(0.588)	(0.608)
Observations	2,290	2,305	2,938	2,957	2,948	2,967
R-squared	0.299	0.297	0.311	0.305	0.299	0.284

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.4. OLS Model: How fertility in origin country affects number of children of second generation immigrant men in France

	(1)	(1.b)	(2)	(2.b)	(3)	(3.b)
TFR at birth	0.0448***	0.0539***				
	(0.0107)	(0.00987)				
TFR at the end of education			0.132***	0.140***		
			(0.0139)	(0.0134)		
TFR in 2008					0.173***	0.159***
					(0.0270)	(0.0172)
LFP in 2008	-0.00151		-0.00361***		-0.00670***	
	(0.000880)		(0.000839)		(0.000869)	
Age						
18-20	0.427	0.429	0.00426	-0.000879	0.00660	-0.00385
	(0.372)	(0.380)	(0.243)	(0.257)	(0.216)	(0.230)
21-25	0.0641	0.0605	-0.176	-0.180	-0.198	-0.193
	(0.276)	(0.286)	(0.183)	(0.198)	(0.163)	(0.180)
26-30	-0.0442	-0.0393	-0.154	-0.146	-0.190	-0.171
	(0.177)	(0.183)	(0.125)	(0.136)	(0.118)	(0.129)
31-35	0.0131	0.0268	0.0121	0.0332	-0.0351	-0.00129
	(0.154)	(0.157)	(0.114)	(0.116)	(0.110)	(0.111)
36-40	0.0919	0.0987	0.144**	0.155**	0.125**	0.140**
	(0.0919)	(0.0971)	(0.0560)	(0.0621)	(0.0542)	(0.0597)
More than 50			0.155	0.135	0.165	0.142
			(0.143)	(0.140)	(0.139)	(0.136)
Squared age	0.000931***	0.000938***	0.000630***	0.000631***	0.000751***	0.000759***
	(0.000168)	(0.000168)	(9.88e-05)	(0.000102)	(0.000101)	(9.95e-05)
Age of Partner	0.336***	0.332***	0.230***	0.228***	0.240***	0.239***
	(0.0279)	(0.0274)	(0.0234)	(0.0234)	(0.0218)	(0.0217)
Squared Age of Partner	-0.00459***	-0.00453***	-0.00294***	-0.00292***	-0.00307***	-0.00308***
	(0.000373)	(0.000369)	(0.000286)	(0.000289)	(0.000259)	(0.000262)
Siblings	0.0202***	0.0171**	0.0169**	0.0145*	0.0176***	0.0191***
	(0.00681)	(0.00701)	(0.00716)	(0.00738)	(0.00585)	(0.00630)
Education						
Primary	-0.0736	-0.0852	-0.206**	-0.202**	-0.227**	-0.222**
	(0.109)	(0.110)	(0.0856)	(0.0865)	(0.0906)	(0.0937)
Lower Secondary	-0.0861	-0.0844	-0.129*	-0.120	-0.152**	-0.133
	(0.0598)	(0.0617)	(0.0701)	(0.0790)	(0.0663)	(0.0789)
Upper-Secondary	-0.125**	-0.129**	-0.151**	-0.150**	-0.207***	-0.205***
	(0.0494)	(0.0507)	(0.0615)	(0.0613)	(0.0573)	(0.0593)
College	-0.238***	-0.248***	-0.158**	-0.158**	-0.264***	-0.259***
	(0.0504)	(0.0486)	(0.0756)	(0.0707)	(0.0622)	(0.0619)
Partner's Education						
Primary	-0.0220	-0.0377	-0.213**	-0.217***	-0.269***	-0.283***
	(0.144)	(0.146)	(0.0759)	(0.0753)	(0.0708)	(0.0722)
Lower Secondary	-0.201**	-0.228***	-0.351***	-0.365***	-0.408***	-0.429***
	(0.0817)	(0.0707)	(0.0956)	(0.0859)	(0.0991)	(0.0883)
Upper-Secondary	-0.443***	-0.459***	-0.507***	-0.510***	-0.541***	-0.555***
	(0.0682)	(0.0661)	(0.0671)	(0.0684)	(0.0763)	(0.0801)
College	-0.737***	-0.755***	-0.758***	-0.767***	-0.805***	-0.830***
	(0.0746)	(0.0748)	(0.0577)	(0.0609)	(0.0669)	(0.0739)
Mother's Education	yes	yes	yes	yes	yes	yes
Father's Education	yes	yes	yes	yes	yes	yes
Region	yes	yes	yes	yes	yes	yes
Constant	-5.258***	-5.269***	-3.292***	-3.427***	-3.367***	-3.584***
	(0.751)	(0.735)	(0.539)	(0.545)	(0.531)	(0.521)
Observations	2,826	2,856	3,685	3,722	3,701	3,738
R-squared	0.425	0.424	0.425	0.423	0.424	0.418

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

3.6.2 Labor Force Participation

The OLS results of cultural impact on labor working time when we restrict our sample to the married women are shown in Table 5. The Table is organized into 6 panels studying the effect of 3 types of labor force participation cultural norms, while considering also how TFR at the end of education in the origin country affects these results. Like for the TFR, here we find that the norm is stronger the closer we come to the survey year. We start with LFP at birth which is not significant and it affects negatively the hours worked of the respondent. In Model 2, we find that LFP at the end of education is significant at the 1% level and it affects the hours worked by the respondent positively. We find a significant effect of the LFP in 2008 in the origin country on the labor working time of the respondent at the 1% level. This shows that the cultural norms transmitted through peers are stronger than the ones transmitted through parents. LFP in 2008 is the norm that explains more of the variability of hours worked. 1 standard deviation increase in LFP in 2008 in origin country raises hours worked of second-generation migrant in France by 2.41 hours. The effect of the norm is in the same direction as in Fernandez and Fogli (2009) where an increase in LFP in the origin country, measured in 1950, raises hours worked by 0.04 to 0.07.

In Models 1,2 and 3 we check if TFR at the end of education alters the LFP cultural norm and how it affects the hours worked by the respondent. Even though the inclusion of TFR reduces systematically the size of the effect of the norm, suggesting that a part of the effect of the norm is passing through the negative association between fertility norms and labor force participation norms. We control for the age of the respondent and their partner, as well as the number of siblings. Age of partner is not significant in explaining hours worked. Having a sibling makes respondents work fewer hours, and the coefficients are significant for Models 2.b, 3, and 3b. Having only a diploma from primary school affects negatively hours worked by the respondent while having a diploma from secondary school and college is significant and it affects positively how many hours the respondent works. Having a college diploma raises the time worked per week by approximately 10 hours. On the partner's side, having a primary school diploma is significant and it affects positively the number of hours worked by the respondent. The effect is not significant when the partner holds a lower or upper secondary school diploma, but when the partner has a college degree, his wife works fewer hours. In all these results we control for parents' education and regional dummies.

Table 3.5. OLS Model: How LFP in origin country affects hours worked of second generation immigrant in France

	(1)	(1.b)	(2)	(2.b)	(3)	(3.b)
LFP at birth	-0.0261	-0.0219				
	(0.0269)	(0.0246)				
LFP at the end of education			0.0566***	0.0666***		
			(0.0176)	(0.0169)		
LFP in 2008					0.124***	0.130***
					(0.0319)	(0.0295)
TFR at the end of education	-0.995***		-0.681***		-0.320	
	(0.341)		(0.196)		(0.327)	
Age						
18-20	-8.178	-7.642	-16.35***	-15.71**	-15.62***	-15.30***
	(6.955)	(7.050)	(5.459)	(5.618)	(4.739)	(4.728)
21-25	-3.918	-3.517	-10.45**	-10.03**	-10.16**	-9.935**
	(5.661)	(5.668)	(4.076)	(4.093)	(3.777)	(3.578)
26-30	-1.494	-1.230	-6.910**	-6.572**	-6.283**	-6.086**
	(4.028)	(4.050)	(2.759)	(2.783)	(2.306)	(2.231)
31-35	-1.826	-1.661	-5.789**	-5.544**	-5.348***	-5.219***
	(2.882)	(2.891)	(2.131)	(2.179)	(1.619)	(1.659)
36-40	-0.592	-0.653	-3.521**	-3.372*	-2.987**	-2.903**
	(1.933)	(1.954)	(1.628)	(1.665)	(1.328)	(1.349)
More than 50			-2.121	-2.056	-2.605	-2.644
			(2.489)	(2.494)	(2.010)	(1.988)
Squared age	0.00496	0.00542*	0.000189	0.000396	-8.59e-05	1.08e-05
	(0.00304)	(0.00307)	(0.00196)	(0.00202)	(0.00163)	(0.00160)
Age of partner	-0.280	-0.321	0.157	0.137	0.0862	0.0845
	(0.472)	(0.460)	(0.333)	(0.335)	(0.319)	(0.322)
Age of partner Squared	0.00361	0.00362	-0.00265	-0.00277	-0.00209	-0.00227
	(0.00528)	(0.00517)	(0.00322)	(0.00318)	(0.00317)	(0.00310)
Siblings	-0.301	-0.459	-0.173	-0.288**	-0.245*	-0.301**
	(0.254)	(0.281)	(0.123)	(0.119)	(0.135)	(0.115)
Education						
Primary	-1.684	-1.899	-0.250	-0.126	0.916	0.853
	(1.539)	(1.520)	(1.281)	(1.289)	(1.304)	(1.345)
Lower Secondary	4.076***	4.251***	4.979***	5.062***	5.059***	5.083***
	(1.404)	(1.394)	(0.956)	(0.985)	(0.848)	(0.854)
Upper-Secondary	9.591***	10.19***	8.701***	9.075***	8.802***	8.968***
	(0.804)	(0.802)	(0.722)	(0.715)	(0.542)	(0.513)
College	10.96***	11.87***	10.84***	11.45***	11.33***	11.62***
	(0.646)	(0.714)	(0.665)	(0.672)	(0.729)	(0.700)
Partner's Education						
Primary	7.795***	8.150***	4.576***	4.765***	4.087***	4.180***
	(2.125)	(2.132)	(1.005)	(1.061)	(1.107)	(1.111)
Lower Secondary	-0.520	-0.464	-1.647	-1.572	-2.275	-2.290
	(2.137)	(2.174)	(2.000)	(2.030)	(1.884)	(1.906)
Upper-Secondary	0.616	0.837	-0.120	0.0535	-0.432	-0.429
	(1.055)	(1.067)	(1.057)	(1.049)	(0.854)	(0.844)
College	-0.597	-0.553	-1.836	-1.736	-2.204**	-2.212**
	(1.169)	(1.174)	(1.304)	(1.296)	(1.029)	(1.029)
Mother's Education	yes	yes	yes	yes	yes	yes
Father's Education	yes	yes	yes	yes	yes	yes
Region	yes	yes	yes	yes	yes	yes
Constant	21.16	18.82	20.99*	18.85	19.06*	17.95*
	(13.03)	(13.17)	(11.09)	(11.07)	(10.86)	(10.29)
Observations	1,855	1,855	2,740	2,740	2,921	2,931
R-squared	0.171	0.167	0.133	0.130	0.142	0.142

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

We moreover consider running regressions of horse-racing explanatory variables to find out how their significance and magnitude evolve. We place in the same estimation equation, the three measures of cultural norms, or we consider different combinations of them. Regarding fertility, only TFR at the end of education remains significant and at the same magnitude, irrespectively of which other norm we consider in the same regression. TFR in 2008 loses significance but it remains positive. TFR at birth remains significant but it changes signs when we consider it together with TFR at the end of education, it loses significance but it remains positive when we consider it with TFR in 2008. When we consider the three cultural norm measures in the same regression, TFR at birth changes sign and remains significant, TFR at the end of education remains significant and at the same magnitude and TFR in 2008 loses significance. When we redo the same analysis for Labor force participation cultural norm variables, LFP in 2008 is the cultural norm proxy that remains significant, irrespectively of the other norm we consider in the regression. The results are shown in Appendix D.

3.6.3 Discrimination

We confirm the finding of Fernandez and Fogli (2009) on France and show that the time at which the norm is measured matters, suggesting that socialization of peers is an important aspect in explaining decisions about fertility and labor working time. Moreover, considering that discrimination and integration are part of immigration trajectories, we study whether perceived discrimination-related variables alter the effect that these norms have on labor working time and fertility decisions. For this, we run an OLS regression where we estimate the effect of LFP in 2008 in the origin country, a set of discrimination measures, the interaction between discrimination measures, personal, partner, and parents characteristics on the labor working time. The aim is to see how the discrimination measures, together with their interaction will affect hours worked and will interfere with the effect of the norm on hours worked per week. Results of how different types of discrimination affect hours worked and how they interfere with the cultural norms are shown in Table 6. Respondents answer questions where they declare if they have felt discriminated about their age, health, origin, place of living, religion, gender, outfit, and skin color. We control for age of respondent and partner, education of respondent, that of partner and parents, number of siblings, and region.

Hours Worked: The results show that whatever the measure specification of discrimination, the cultural variable (LFP in 2008) remains significant in explaining hours worked. Being discriminated based on age has a negative effect on hours worked and it is not significant. This type of discrimination does not alter the magnitude or significance of the cultural norm. Being discriminated based on health affects the number of hours worked by the respondent negatively and it does not change the significance of the cultural norm. Even when LFP in the origin country raises by 1, respondents that were discriminated based on their health, work 0.17 hours less in a week. Being discriminated based on origin country affects negatively hours worked by the respondent and it does not alter the persistence of the norm. We can say the same about discrimination based on place of living. The interaction variable is significant and tells us that even when the LFP in the origin country raises by 1 for those that felt discriminated, second-generation migrants work 0.3 hours less in France. Considering these variables of discrimination we can say that it affects negatively hours worked but it does not alter the persistence of the origin country norm on the effect that it has on the labor working time of second-generation migrants. The feeling of being French affects positively the labor working time of the respondent and it is significant at the 1% level. It also raises the power of the norm by increasing its coefficient from 1.18 to 1.65. People that feel French participate more in the labor market. The interaction term between the feeling of being French and the LFP is negative and significant. Coming from a country with one unit higher LFP and being an immigrant that feels French, makes second-generation immigrants work less. This combination term moderates the effect of the cultural norm on hours worked. The feeling of being discriminated tends to reduce or not alter the persistence of cultural norm. On the other hand, the feeling of being French increased the hours worked while not altering the persistence of norms. Moreover, we do not find

any significance for other types of discrimination like clothes, gender, or skin. They do not alter the magnitude of the norm. Coefficients are in the same range as in the first model without discrimination.
Then we consider the same measures of discrimination for analyzing its effect on hours worked for both men and women. The effect of the norm is even higher when we consider the whole population while the discrimination measures coefficients do not change their magnitude. The results are shown in Appendix B.

Number of children: We redo the same analysis for the effect of discrimination-related variables on the number of children of the respondent. An increase in the TFR of origin country at the end of education by 1, raises the number of children of second-generation immigrants in France by 0.181. When we consider age discrimination in this model, the effect of the norm is not altered. We can conclude the same about other types of discrimination related to health, origin, and place of living. These types of discrimination are not significant, the coefficients of the cultural norm and their magnitude do not change. On the other hand, when we consider in this model the feeling of being French, the effect of the norm becomes stronger. The coefficient of the norm increases. The feeling of being French, together with the interaction term TFR are significant. When second-generation female migrants feel French, they have more children. Nevertheless, when TFR in origin country increases by 1, those that feel French lower their number of children by 0.067. When instead we consider discrimination based on religion, the effect of the norm is not altered. This type of discrimination affects negatively the number of children born, but when TFR in the origin country raises by 1, those that feel discriminated have 0.273 more children. This type of discrimination does not alter the persistence of the cultural norm. Regarding gender, when females have felt discriminated, they have more children but when TFR in the origin country raises, they tend to have fewer children. Gender discrimination does not erase the effect of the norm but decreases its power. Females that have felt discriminated in terms of clothes or in terms of skin color have fewer children. Nevertheless this type of discrimination does not alter the norm of the origin country. We also control for age, age of partner, squared age of both respondent and partner, the number of siblings, and education, whose results are shown in Table 7.

In sum, the cultural variable (TFR at the end of education) remains significant throughout the different models, and adding discrimination measures does not modify the effect of the norm. Then we consider the same measures of discrimination for analyzing its effect on the number of children on all women sample , irrespectively of their marital status. The effect of the norm remains significant, the coefficient decreases by 0.04, and the religious discrimination loses significance. The results are shown in Appendix B. Overall the results remain consistent for different samples of our dataset.

Table 3.6. How discrimination affects hours worked

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
LFP in 2008	0.129***	0.130***	0.130***	0.132***	0.129***	0.177***	0.127***	0.131***	0.128***	0.129***
	(0.0294)	(0.0301)	(0.0293)	(0.0302)	(0.0289)	(0.0237)	(0.0273)	(0.0304)	(0.0297)	(0.0319)
Age discrimination		-0.961								
		(4.521)								
Age d * LFP		-0.163								
		(0.133)								
Health discrimination			-0.736							
			(10.23)							
Health d * LFP			-0.121							
			(0.136)							
Origin discrimination				0.0479						
				(1.998)						
Origin d * LFP				-0.0322						
				(0.0522)						
Place of living discrimination					4.018					
					(6.223)					
Place of living d * LFP					-0.307*					
					(0.159)					
Feeling of being French						6.689***				
						(1.419)				
Feeling of being French * LFP						-0.0544**				
						(0.0249)				
Religion discrimination							-3.617			
							(4.350)			
Religion d * LFP							-0.0127			
							(0.0751)			
Gender discrimination								8.349		
								(5.990)		
Gender d *LFP								-0.144		
								(0.144)		
Clothes discrimination									-2.828	
									(3.645)	
Clothes d *LFP									-0.00795	
									(0.112)	
Skin color discrimination										4.336
										(4.578)
Skin Color d * LFP										-0.0470
										(0.0811)
Age	1.998***	1.987***	2.007***	1.988***	2.004***	1.950***	1.993***	1.984***	1.991***	2.008***
	(0.331)	(0.331)	(0.321)	(0.333)	(0.332)	(0.356)	(0.333)	(0.335)	(0.338)	(0.338)
Squared age	-0.0216***	-0.0214***	-0.0217***	-0.0215***	-0.0216***	-0.0212***	-0.0215***	-0.0214***	-0.0215***	-0.0216***
	(0.00357)	(0.00356)	(0.00347)	(0.00360)	(0.00357)	(0.00383)	(0.00361)	(0.00360)	(0.00364)	(0.00360)
Age of Partner	-0.167	-0.175	-0.165	-0.164	-0.185	-0.135	-0.178	-0.153	-0.168	-0.161
	(0.353)	(0.356)	(0.354)	(0.352)	(0.350)	(0.385)	(0.349)	(0.356)	(0.357)	(0.354)
Age of partner Squared	0.000187	0.000266	0.000172	0.000156	0.000354	-0.000215	0.000270	8.51e-05	0.000184	0.000104
	(0.00348)	(0.00352)	(0.00349)	(0.00348)	(0.00345)	(0.00382)	(0.00346)	(0.00349)	(0.00351)	(0.00349)
Siblings	-0.305**	-0.311**	-0.303**	-0.289**	-0.301**	-0.260**	-0.302**	-0.305**	-0.304**	-0.324***
	(0.116)	(0.117)	(0.114)	(0.116)	(0.116)	(0.116)	(0.117)	(0.116)	(0.116)	(0.102)
Education	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Partner's Education	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Mother's Education	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Father's Education	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Region	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	-22.69***	-22.33***	-22.94***	-22.61***	-22.43***	-26.93***	-22.18***	-22.88***	-22.41***	-23.06***
	(5.156)	(5.194)	(5.150)	(5.217)	(5.066)	(4.912)	(5.287)	(5.117)	(5.074)	(5.171)
Observations	2,931	2,931	2,931	2,931	2,931	2,854	2,931	2,931	2,931	2,931
R-squared	0.143	0.145	0.144	0.144	0.144	0.158	0.144	0.144	0.144	0.144

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3.7. How discrimination affects number of children

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
TFR at the end of education	0.181***	0.181***	0.182***	0.184***	0.182***	0.226***	0.178***	0.181***	0.192***
	(0.0125)	(0.0119)	(0.0124)	(0.0115)	(0.0123)	(0.0208)	(0.0132)	(0.0123)	(0.0124)
Age discrimination		-0.567							
		(0.376)							
Age d* TFR		0.0745							
		(0.140)							
Health discrimination			-0.373						
			(0.295)						
Health d *TFR			-0.0568						
			(0.0670)						
Origin discrimination				0.0427					
				(0.146)					
Origin d *TFR				-0.0166					
				(0.0428)					
Place of living discrimination					0.157				
					(0.644)				
Place of living d*TFR					-0.0240				
					(0.189)				
Feeling of being French						0.221**			
						(0.0810)			
Feeling of being French * TFR						-0.0677***			
						(0.0202)			
Religion discrimination							-0.745*		
							(0.392)		
Religion discrimination * TFR							0.273***		
							(0.0849)		
Clothes discrimination								0.655	
								(0.667)	
Clothes d*TFR								-0.0485	
								(0.195)	
Skin color discrimination									-0.119
									(0.311)
Skin d* TFR									-0.0406
									(0.0549)
Age	0.204***	0.203***	0.205***	0.204***	0.204***	0.205***	0.204***	0.205***	0.201***
	(0.0148)	(0.0145)	(0.0141)	(0.0150)	(0.0146)	(0.0135)	(0.0150)	(0.0147)	(0.0152)
Squared age	-0.00222***	-0.00220***	-0.00223***	-0.00222***	-0.00222***	-0.00226***	-0.00222***	-0.00223***	-0.00219***
	(0.000241)	(0.000237)	(0.000232)	(0.000245)	(0.000239)	(0.000226)	(0.000245)	(0.000239)	(0.000244)
Age of Partner	0.0323*	0.0323	0.0319	0.0327*	0.0324*	0.0332*	0.0313	0.0327*	0.0333*
	(0.0186)	(0.0188)	(0.0188)	(0.0187)	(0.0188)	(0.0192)	(0.0186)	(0.0189)	(0.0186)
Age of Partner Squared	-0.000220	-0.000219	-0.000216	-0.000225	-0.000221	-0.000214	-0.000208	-0.000222	-0.000229
	(0.000185)	(0.000185)	(0.000186)	(0.000184)	(0.000189)	(0.000189)	(0.000185)	(0.000186)	(0.000189)
Siblings	0.0159	0.0157	0.0158	0.0161	0.0159	0.0158*	0.0153	0.0157	0.0177
	(0.0107)	(0.0105)	(0.0107)	(0.0108)	(0.0107)	(0.00894)	(0.0110)	(0.0106)	(0.0112)
Education	yes	yes	yes	yes	yes	yes	yes	yes	yes
Partner's Education	yes	yes	yes	yes	yes	yes	yes	yes	yes
Mother's Education	yes	yes	yes	yes	yes	yes	yes	yes	yes
Father's Education	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	-3.402***	-3.382***	-3.409***	-3.421***	-3.410***	-3.597***	-3.374***	-3.447***	-3.388***
	(0.384)	(0.388)	(0.388)	(0.405)	(0.384)	(0.411)	(0.405)	(0.373)	(0.368)
Observations	2,940	2,940	2,940	2,940	2,940	2,863	2,940	2,940	2,940
R-squared	0.303	0.304	0.304	0.303	0.303	0.302	0.305	0.305	0.305

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Moreover we proceed by applying a principal component analysis on the discrimination question variables and create a discrimination index with the respective proportions. This discrimination index does not alter the effect of the cultural norm. The effect of the norm is actually stronger than when we consider discrimination indicators individually. The coefficient of the norm raises from 0.129 to 0.133 for hours worked 0.181 to 0.186 for number of children. Results are shown in Table 8 and in Appendix B we show how we construct the Discrimination measure and the respective weights.

Table 3.8. How discrimination affects numbers of children (left) and hours worked (right)

	(1)	(2)	(3)
LFP in 2008	0.129***	0.127***	0.133***
	(0.0294)	(0.0278)	(0.0305)
Discrimination		-8.381	-2.045
		(5.973)	(9.630)
Discrimination * LFP			-0.180
			(0.240)
Age	1.998***	1.984***	1.978***
	(0.331)	(0.336)	(0.336)
Squared age	-0.0216***	-0.0214***	-0.0214***
	(0.00357)	(0.00361)	(0.00360)
Age of Partner	-0.167	-0.175	-0.169
	(0.353)	(0.349)	(0.351)
Squared Age of Partner	0.000187	0.000252	0.000215
	(0.00348)	(0.00345)	(0.00347)
Siblings	-0.305**	-0.286**	-0.280**
	(0.116)	(0.112)	(0.109)
Education	yes	yes	yes
Partner's Education	yes	yes	yes
Mother's Education	yes	yes	yes
Father's Education	yes	yes	yes
Region	yes	yes	yes
Constant	-22.69***	-21.98***	-22.25***
	(5.156)	(5.130)	(5.229)
Observations	2,931	2,931	2,931
R-squared	0.143	0.144	0.144

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

	(1)	(2)	(3)
TFR at the end of education	0.181***	0.183***	0.186***
	(0.0125)	(0.0128)	(0.0116)
Discrimination		-0.215	0.199
		(0.346)	(0.706)
Discrimination * TFR			-0.116
			(0.210)
Age	0.204***	0.203***	0.204***
	(0.0148)	(0.0147)	(0.0147)
Squared age	-0.00222***	-0.00221***	-0.00222***
	(0.000241)	(0.000240)	(0.000240)
Age of Partner	0.0323*	0.0321	0.0325*
	(0.0186)	(0.0187)	(0.0187)
Squared Age of Partner	-0.000220	-0.000218	-0.000223
	(0.000185)	(0.000186)	(0.000185)
Siblings	0.0159	0.0162	0.0164
	(0.0107)	(0.0108)	(0.0109)
Education	yes	yes	yes
Partner's Education	yes	yes	yes
Mother's Education	yes	yes	yes
Father's Education	yes	yes	yes
Region	yes	yes	yes
Constant	-3.402***	-3.385***	-3.418***
	(0.384)	(0.386)	(0.392)
Observations	2,940	2,940	2,940
R-squared	0.303	0.303	0.303

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

After looking at perceived discrimination, we continue with integration perceptions. The feeling of being French moderates the effect of the cultural norm. Different from discrimination that does not alter the effect of the norm, integration measures can diminish the effect that the norm has on the second generation migrant decision making. [18] Feeling French has a positive and

[18]The importance of distinguishing the Feeling of being French with the discrimination measure is that, while the concept of dual belonging (being French but also feeling another nationality), is accepted in multiculturalist societies like in the United States and Canada, it

significant effect on the hours worked of the respondents. We find that when people feel French, they participate more in the labor market, by increasing the number of hours worked by 4.5 compare to second generation migrants that do not feel the same way. Being originally from a country with a higher LFP, women that feel French work less than those who don't. For fertility, coming from a country with a high TFR, makes woman that feel French have less children compared to those who do not feel French. Results are shown in Table 9.

Table 3.9. How the feeling of being French affects number of children and hours worked

	(1) Worked hours	(2) Worked hours	(3) Worked hours	(4) Number of children	(5) Number of children	(6) Number of children
LFP in 2008	0.129***	0.137***	0.177***			
	(0.0294)	(0.0274)	(0.0237)			
TFR at the end of education				0.181***	0.178***	0.226***
				(0.0125)	(0.0132)	(0.0208)
Feeling of being French		4.525***	6.689***		-0.0204	0.221**
		(1.069)	(1.419)		(0.0372)	(0.0810)
Feeling of being French * TFR						-0.0677***
						(0.0202)
Feeling of being French * LFP			-0.0544**			
			(0.0249)			
Age	1.998***	1.931***	1.950***	0.204***	0.208***	0.205***
	(0.331)	(0.345)	(0.356)	(0.0148)	(0.0136)	(0.0135)
Squared age	-0.0216***	-0.0209***	-0.0212***	-0.00222***	-0.00229***	-0.00226***
	(0.00357)	(0.00371)	(0.00383)	(0.000241)	(0.000228)	(0.000226)
Age of partner	-0.167	-0.114	-0.135	0.0323*	0.0325	0.0332*
	(0.353)	(0.368)	(0.385)	(0.0186)	(0.0193)	(0.0192)
Age of partner Squared	0.000187	-0.000428	-0.000215	-0.000220	-0.000203	-0.000214
	(0.00348)	(0.00364)	(0.00382)	(0.000185)	(0.000190)	(0.000189)
Siblings	-0.305**	-0.260**	-0.260**	0.0159	0.0159*	0.0158*
	(0.116)	(0.115)	(0.116)	(0.0107)	(0.00894)	(0.00894)
Education	yes	yes	yes	yes	yes	yes
Partner's Education	yes	yes	yes	yes	yes	yes
Mother's Education	yes	yes	yes	yes	yes	yes
Father's Education	yes	yes	yes	yes	yes	yes
Region	yes	yes	yes	yes	yes	yes
Constant	-22.69***	-25.48***	-26.93***	-3.402***	-3.452***	-3.597***
	(5.156)	(5.233)	(4.912)	(0.384)	(0.407)	(0.411)
Observations	2,931	2,854	2,854	2,940	2,863	2,863
R-squared	0.143	0.157	0.158	0.303	0.300	0.302

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

---

has been criticized in France, where many perceive identity as a zero-sum game: commitment to a minority culture or a foreign country detracts from the quality of one's commitment to French identity.

3.6.4 Genetic Distance

As feeling French appears an important determinant of second-generation migrants behavior in terms of fertility and labor working time, we further analyze how genetic distance affects the feeling of being French. Genetic distance is an indicator of genetic divergence between populations, measured by the frequency with which recombination events occur between the genes. Doing so, allows to study the hypothesis if coming from further away has a restrictive effect in terms of integration. Distance is a concept that can be explained in terms of geography, genetic composition, behavior and culture. The further away people migrate, the more differences in culture they are expected to face. Social and cultural factors in human activities are reflected in human genetic diversity, considering that these factors are associated with the history of demography in the population.[19] Consistent occurrences can have an impact on demography, and this has an impact on human DNA. Such demographic changes can be observed by analyzing genetic information from current human populations. However, research on ancient DNA offers more details on the migration and demographic shifts that affected population genetic structure. Natural selection, societal structures, and non-biological events like marriage systems have all had a significant impact on human genetic diversity (Torres et al. (2018)). The variations in culture and thought processes can be explained by this genetic variety.

To test this hypothesis, we consider a linear regression model of hours worked or number of children on cultural norms and other explanatory variables, together with an endogenous binary treatment. Our treat is the Feeling of being French modeled by age and genetic distance. We consider as genetic distance data of the migratory distance, from Ashraf and Galor (2013). Our results show that genetic distance has a negative effect and is significant in explaining the 'feeling of being French'. The more the

[19] Key et al. (2016) show that genetic diversity within a population is affected by demography as a result of subsistent styles in humans.

individuals are distant genetically, the less French they feel. With age, people start feeling more French. Moreover, the feeling of being French is significant in explaining hours worked, which is in line with Table 9. It has a negative effect on the number of children, which is in the same direction as Table 9 but here the coefficient is significant, which is not the case previously. The effects of the norms are not altered and they are still strongly significant in explaining the fertility and labor working time of second-generation migrants in France. Results are shown below in Table 10.

Table 3.10. LRM by treating feeling of being French on age and genetic distance

	(1)	(1b)	(1c)	(2)	(2b)	(2c)
	Model 1			**Model 2**		
VARIABLES	Hours Worked	Feeling French		Number of children	Feeling French	
LFP in 2008	0.101***			-0.00671***		
	(0.0284)			(0.00101)		
TFR at the end of education	-0.0161			0.158***		
	(0.204)			(0.0200)		
Age Dummies	yes			yes		
Squared age	-0.000167			-0.000115		
	(0.00108)			(0.000133)		
Age of Partner	0.268			0.0589***		
	(0.290)			(0.0165)		
Squared Age of Partner	-0.00352			-0.000488**		
	(0.00296)			(0.000191)		
Siblings	-0.146			0.0144*		
	(0.115)			(0.00823)		
Education variables	yes			yes		
Region	yes			yes		
Feeling of being French = 1	47.86***			-1.143***		
	(1.760)			(0.291)		
Age		0.00443**			0.00454**	
		(0.00183)			(0.00191)	
Genetic distance		-4.990***			-7.611*	
		(1.711)			(4.114)	
athrho			-1.966***			0.593***
			(0.150)			(0.168)
lnsigma			3.256***			0.248***
			(0.0257)			(0.0442)
Constant	-20.71***	0.366***		1.985***	0.450***	
	(7.977)	(0.0821)		(0.671)	(0.0917)	
Observations	2,836	2,836	2,836	2,836	2,836	2,836

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

3.6.5 Robustness checks

We test the robustness of our estimates by considering different estimation strategies, specifications, and samples. We consider a Poisson Model for studying fertility and a Hurdle model for studying labor force participation. We then consider different samples in terms of age, marital status, and income. Since in our main analysis we consider as origin country the mother's ancestry, we extend our study by considering as a cultural proxy the origin country of the father. The common result of these robustness checks is that culture matters for the number of children and labor working time and that the norm gains strength when measured later in time compared to the time of birth of the respondent.

A Poisson Regression Model for studying Fertility

For studying fertility we consider a Poisson regression which is a non-linear multiple regression model. We assume that the fertility of the respondent has a Poisson distribution considering that it is a discrete count variable. We then analyze the determinants that affect the number of children of the respondent. As shown in the histogram of Appendix E there is no excess of 0 in the fertility data, for this we do not consider a Zero Inflated Poisson (ZIP) model. The results imply that the marginal effect of the TFR at birth is not independent of the other characteristics of the respondent. The similarity of the results with the OLS model stands in the characteristic that the effect of the norm becomes stronger with time. Like in the OLS model we control for age, partner's age, education of respondent, partner and parents, the number of siblings, and the location of their residence. The results of the effects of origin TFR on the fertility of second-generation immigrants in France when we consider the actual fertility of the respondent following a Poisson distribution, are shown in Table 8.

Each additional TFR in the origin country, depending on when it was measured, is associated with 0.03, 0.06, and 0.09 more children. A second-generation immigrant that comes from a country with a TFR of 2 is associated with 0.03 to 0.09 more children on average. Like with the OLS Model we see that the cultural norm's effect becomes stronger the closer we come to the actual year the respondent took the survey. The cultural norm measured later in time has higher coefficients compared to the norms measured earlier. Regarding the personal, partner, and parents' characteristics the association of these variables with fertility are in the same direction as with the OLS results.

Table 3.11. Poisson Model: How TFR in origin country affects fertility of second generation immigrant in France

	(1)	(2)	(3)
TFR at birth	0.0346***		
	(0.00932)		
TFR at the end of education		0.0660***	
		(0.00765)	
TFR in 2008			0.0912***
			(0.0172)
LFP in 2008	-0.00158**	-0.00247***	-0.00465***
	(0.000678)	(0.000434)	(0.000402)
Age			
18-20	-1.017***	-1.184***	-1.287***
	(0.275)	(0.255)	(0.262)
21-25	-0.586***	-0.699***	-0.791***
	(0.0990)	(0.0922)	(0.0965)
26-30	-0.371***	-0.422***	-0.500***
	(0.0998)	(0.0721)	(0.0776)
31-35	-0.0981	-0.107**	-0.161***
	(0.0643)	(0.0495)	(0.0543)
36-40	-0.0689	-0.0516	-0.0796*
	(0.0437)	(0.0395)	(0.0412)
More than 50		0.0319	0.0457
		(0.0458)	(0.0431)
Squared age	2.18e-05	-4.74e-05	-9.90e-06
	(8.77e-05)	(5.18e-05)	(4.58e-05)
Age of Partner	0.0835***	0.0385***	0.0380***
	(0.0108)	(0.0101)	(0.0105)
Age of partner Squared	-0.000853***	-0.000344***	-0.000336***
	(0.000112)	(8.94e-05)	(9.37e-05)
Siblings	0.00718**	0.00629*	0.00732
	(0.00337)	(0.00372)	(0.00501)
Education			
Primary	0.0973**	0.0600**	0.0362
	(0.0403)	(0.0265)	(0.0267)
Lower Secondary	-0.0710*	-0.0536	-0.0806**
	(0.0393)	(0.0383)	(0.0347)
Upper-Secondary	-0.174***	-0.128***	-0.159***
	(0.0284)	(0.0196)	(0.0198)
College	-0.321***	-0.226***	-0.282***
	(0.0472)	(0.0380)	(0.0353)
Partner's Education	yes	yes	yes
Mother's Education	yes	yes	yes
Father's Education	yes	yes	yes
Region	yes	yes	yes
Constant	-1.021***	-0.0393	0.0768
	(0.261)	(0.304)	(0.321)
Pseudo R2	0.0775	0.0832	0.0821
Observations	2,282	2,921	2,931

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.12. Factor changes in the Poisson Regression Model

	b	z	P>|z|	e^b	$e^b Stdx$	SD of X
Model 1						
TFR at birth	0.0348	3.794	0.000	1.035	1.070	1.951
LFP 2008	-0.0015	-2.221	0.026	0.998	0.971	19.140
Model 2						
TFR at the end of education	0.0661	8.710	0.000	1.068	1.133	1.882
LFP 2008	-0.0025	-5.743	0.000	0.998	0.954	19.082
Model 3						
TFR in 2008	0.0917	5.375	0.000	1.096	1.109	1.125
LFP 2008	-0.0047	-12.003	0.000	0.995	0.915	19.072

In Table 12 we show the factor changes in the Poisson regression model. The results show that having a parent originally from a country with 1 unit higher TFR at birth, increases the expected number of children by 1.035, holding other variables constant. For a standard deviation increase in TFR at birth in the origin country, roughly 1.9, the average number of children increases by a factor of 1.07, holding other variables constant. An increase in LFP in the country of origin by 1, increases the expected number of children by 0.998. When we consider as the cultural norm the proxy measured by TFR at the end of education: Being originally from a country with one unit higher TFR at the end of education, raises the expected number of children by 1.068. For a standard deviation increase in TFR at birth in the origin country, roughly 1.8, the average number of children increases by a factor of 1.13, holding other variables constant. We get similar values when we measure the cultural norm by the TFR in 2008.

Hurdle Model for studying LFP

For studying the impact of the cultural norm of the origin country on labor working time, we consider a Hurdle Model due to an excess of zero's in the labor force participation data, which is not sufficiently accounted for in the previous model. A histogram with the distribution of hours worked, showing an excess of zeros is shown in Appendix E. We model LFP using two parts, the first which is the probability of attaining value 0, and the second part models the probability of the non-zero values. With this model, we fit a linear hurdle model for a bounded dependent variable of labor force participation. The hurdle model combines a selection model that determines the boundary points of the LFP with an outcome model that determines its non-bounded values.

We find that only labor force participation in 2008 is strongly significant and has a positive effect on worked hours. An increase in the labor force participation in 2008 by 1 unit, increases the difference in the logs of expected hours worked by 0.09. LFP at birth and LFP at end of education have similar effects as in the OLS. The number of children continues to be important for the decision of labor working time. In Table 13 we show a summary of the average marginal effects of our explanatory variables on hours worked and in Table 14 we show the results of the Hurdle Model. In Appendix J we also show a horse-raced version of the Hurdle Model with the three measures of the cultural norm in the same regression. The results show that LFP at the end of the education is the only cultural norm significant in explaining hours worked.

Table 3.13. Summary of average marginal effects on hours worked

Average marginal Effects	Delta-method - dy/dx	Std. Err.	z	P>|z|	95% Conf. Interval
Model 1					
LFP at birth	-.0201475	.033332	-0.60	0.546	-.085477 : .0451821
TFR at the end of education	-.5728112	.4293196	-1.33	0.182	-1.414262 ; .2686398
Number of children	-3.371828	.772365	-4.37	0.000	-4.885636 ; -1.858021
Model 2					
LFP at the end of education	.0256237	.0185785	1.38	0.168	-.0107894 ; .0620369
TFR at the end of education	-.2913659	.272855	-1.07	0.286	-.8261519 ; .2434202
Number of children	-2.437398	.5052246	-4.82	0.000	-3.42762 ; -1.447176
Model 3					
LFP in 2008	.0938626	.0303762	3.09	0.002	.0343264 ; .1533989
TFR at the end of education	.0440708	.2703646	0.16	0.871	-.4858341 ; .5739757
Number of children	-2.181078	.4244298	-5.14	0.000	-3.012945 ; -1.349211

Table 3.14. Hurdle Model: : How LFP in origin country affects FLFP of second generation immigrant in France

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Model 1			**Model 2**			**Model 3**		
Variables	Outcome	Single Selection	lnsigma	Outcome	Single Selection	lnsigma	Outcome	Single Selection	lnsigma
LFP at birth	0.0493	-0.00389***							
	(0.0482)	(0.00146)							
LFP at the end of education				-0.0173	0.00279**				
				(0.0233)	(0.00134)				
LFP in 2008							-0.0272	0.00901***	
							(0.0269)	(0.00215)	
TFR at the end of education	-0.333	-0.0378		-0.573	-0.00363		-0.447	0.0206	
	(0.376)	(0.0383)		(0.653)	(0.00964)		(0.532)	(0.0181)	
Number of children		-0.305***			-0.205***			-0.185***	
		(0.0334)			(0.0212)			(0.0235)	
Age									
18-20	-5.813	-0.753		-6.107	-1.634***		-7.665	-1.559***	
	(14.39)	(0.580)		(13.39)	(0.462)		(11.27)	(0.427)	
21-25	-2.748	-0.223		-3.180	-0.966***		-4.128	-0.936***	
	(9.910)	(0.488)		(8.970)	(0.313)		(7.430)	(0.277)	
26-30	-0.282	-0.0739		0.0811	-0.691***		-0.741	-0.636***	
	(9.623)	(0.372)		(8.729)	(0.240)		(7.342)	(0.204)	
31-35	0.984	-0.0135		1.233	-0.493***		0.657	-0.461***	
	(9.294)	(0.267)		(8.447)	(0.182)		(7.309)	(0.148)	
36-40	1.079	0.0371		0.558	-0.283**		0.112	-0.246**	
	(6.548)	(0.164)		(5.695)	(0.132)		(5.004)	(0.114)	
More than 50				-1.149	-0.145		0.693	-0.217	
				(5.882)	(0.199)		(4.714)	(0.173)	
Squared age	0.00326	0.000652**		0.00339	2.05e-05		0.00231	1.52e-05	
	(0.00883)	(0.000281)		(0.00766)	(0.000163)		(0.00634)	(0.000144)	
Age of Partner	-0.191	0.0388		-0.0609	0.0339		-0.0385	0.0261	
	(0.521)	(0.0360)		(0.401)	(0.0295)		(0.365)	(0.0290)	
Squared Age of Partner	0.000942	-0.000418		-0.000204	-0.000423		-0.000396	-0.000362	
	(0.00703)	(0.000407)		(0.00431)	(0.000310)		(0.00390)	(0.000311)	
Siblings	-0.251	-0.0167		-0.200	-0.0109*		-0.257	-0.0144*	
	(0.410)	(0.0123)		(0.258)	(0.00611)		(0.223)	(0.00783)	
Education									
Primary	0.416	-0.0670		-0.155	0.0224		-0.205	0.110	
	(2.024)	(0.164)		(3.473)	(0.107)		(3.028)	(0.0937)	
Lower Secondary	6.392	0.267**		2.403	0.334***		2.019	0.341***	
	(6.293)	(0.127)		(2.778)	(0.0602)		(2.644)	(0.0641)	
Upper-Secondary	4.979**	0.621***		2.193	0.559***		1.858	0.571***	
	(2.274)	(0.0876)		(1.945)	(0.0577)		(2.047)	(0.0492)	
College	5.281***	0.586***		1.798	0.665***		1.981	0.695***	
	(1.744)	(0.0784)		(2.308)	(0.0547)		(2.327)	(0.0570)	
Partner's Education	yes	yes		yes	yes		yes	yes	
Mother's Education	yes	yes		yes	yes		yes	yes	
Father's Education	yes	yes		yes	yes		yes	yes	
Region	yes	yes		yes	yes		yes	yes	
Constant	30.24	-0.708	3.018***	29.65	0.222	3.141***	32.04	0.0610	3.108***
	(29.05)	(1.161)	(0.408)	(22.87)	(0.961)	(0.393)	(20.09)	(0.921)	(0.378)
Observations	1,860	1,860	1,860	2,756	2,756	2,756	2,938	2,938	2,938
Pseudo R2	0.0485			0.0351			0.0367		

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Sample extensions

Our results are robust to alternative sample selection criteria, estimation techniques, and exchanging the origin of the mother with that of the father when creating the variables of cultural norms. We furthermore control for the education of the respondent, partner and parents, the number of siblings, income and age to account for the different timing of the cultural norm and the time of the answer of the respondent.

When considering different samples of our dataset, results remain similar. A sample that we consider when studying fertility is female respondents that have finished their fertility cycle. For this, we limit our sample to women above the age of 40. The effect of the norm is still positive and significant. Again, the effect of the norm is stronger the later the norm is measured in the origin country. Tables with results are shown in Appendix A.1. Then we consider all marital status females and the results hold for both effects on fertility and labor working time. For fertility the three norms are significant while for hours worked, only the two norms measured at a later stage. The results are consistent with the sample of married women. Tables with results are shown in Appendix A.2. Moreover, we redo our study by considering father's country of origin in terms of cultural transmission of norms, instead of the mother's like previously. This is important especially in cases when parents have different origin countries. For fertility estimates, coefficients are in a similar range for the three measures of norms. We see the same for labor force participation. Model 2 and 3 have similar values and they remain significant and at the same magnitude. This happens because in our sample of second-generation migrants we have 62% whose parents are originally from the same country. Only 38 % of first-generation migrants have married someone from a different country of origin than theirs. As a result, we see the same strength of coefficients of cultural norms for the countries of origins of both parents. Tables with results are shown in Appendix C. In Appendix F we study how cultural norms affect the number of children and hours worked by accounting for income. Results hold as previously, nevertheless they should be interpreted with caution considering the endogeneity of income. The measure is a categorical variable based on a survey question asking if their income ranges from 'sufficient' to 'hard times'.

Regarding the perceived discrimination and integration analysis, we identify the effect of culture on actual decisions by extending our sample to that of all men and women. Results hold as in the previous analysis. The table with results is shown in Appendix B.

3.7 Conclusion

Our study concludes that cultural norms affects fertility and labor working time for both men and women. By exploiting the rich dataset of TeO for the year 2008, we find that cultural norms in the origin country affect the number of children and hours worked per week of second-generation migrants in France. Having a parent that has migrated to France, but that has origins from a country with a high TFR, makes second-generation migrants have 0.2 more children. Second-generation migrants work 0.12 hours per week more when their parents come from a country with higher LFP. This article supports the overall importance of cultural norms' effect on decisions and behavior of women in such a personal decisions like the number of children, as well as in the economic markets by looking at labor working time. Our findings are in line with the existing literature that suggests that culture is an important aspect of people's decisions. We confirm the findings of Fernandez and Fogli (2009) for the second-generation migrants in the US to hold also for the French context.

Moreover, we find that the timing of the norm is crucial. Recent cultural norms are stronger in explaining decisions compared to norms measured earlier in time. Cultural norms of the origin country at the time of the interview and at the end of education explain better fertility and labor market behavior compared to norms measured at the time of birth. This shows that the socialization of peers has a stronger effect than that of parents, stressing the importance of horizontal cultural transmission. This result adds to the findings of the vertical and horizontal transmission in second-generation migrants by investigating which type of cultural transmission has the strongest effect on the decision-making of second-generation migrants in France.

Not only do cultural norms affect behavior but also perceived discrimination and integration aspects, like the feeling of being French, are two important perceptions that affect fertility decisions and labor working time. We find that feeling French initiates individuals work between 4.5 to 6.6 hours more per week, while among those that feel French, coming from a country with high TFR, makes them have 0.06 fewer children. Perceived discrimination does not alter the effect that the cultural norms have on second-generation immigrants' behavior. Integration perceptions like feeling French, moderate the effect that the norm has on immigrants' decisions. When comparing the mother's and father's country of origin and their importance in the transmission of cultural norms, we do not find a difference in the power of cultural norm transmission. The effect of the norm from both parents' country of origin is very similar because 62% of first-generation migrants have married someone from their country of origin. We show that culture plays an important role in explaining the variability in women's fertility and labor time. TFR in the origin country affects the actual number of children of the respondent. LFP in the origin country affects the hours worked of second-generation migrants in France. Age, education, number of siblings, and partner characteristics are also very important drivers that explain preferences and decision making.

Our study leaves open avenues for future research. Culture's impact is restricted to the variables available in our only wave of the TeO dataset. Further possible research could focus on the length of parental residence in France or other measures of cultural norms. Moreover, this study is relevant for policy implications. Regarding labor market participation, a policy for successful integration and inclusion is essential considering our results. A policy that promotes inclusion for all and relaxes the barriers that can hinder the participation of people with a migrant background. Immigration has a very long history in France since the late nineteenth century and it is a delicate and powerful analysis pushing toward different types of policies. Historically, the French immigration policy has pointed at two objectives: i) meeting the needs of the labor market by introducing migrant workers, ii) compensating French demographic deficits by favoring the immigration of foreign families and pushing for their integration into the national body. Nevertheless, diversity comes with challenges. We provide evidence that the culture of origin country and perceived integration explain important lifetime decisions, elements that should be taken into account while drafting integrative policies.

3.8 Appendix

3.8.1 Women respondents that have finished their fertility cycle

Table 3.15. **OLS Model**: How TFR in origin country affects number of children of second generation immigrants in France, for women that have finished their fertility cycle (> 40)

	(1)	(2)	(3)
TFR at birth	0.130***		
	(0.0429)		
TFR at the end of education		0.235***	
		(0.0208)	
TFR in 2008			0.377***
			(0.0338)
LFP in 2008	-0.00125	-0.00353**	-0.0129***
	(0.00277)	(0.00159)	(0.00117)
Age	-0.131	-0.164	-0.0532
	(0.454)	(0.145)	(0.145)
Squared age	0.00131	0.00152	0.000591
	(0.00519)	(0.00150)	(0.00149)
Partner's Age	0.252***	0.0637	0.0583
	(0.0500)	(0.0501)	(0.0516)
Age of Partner Squared	-0.00236***	-0.000470	-0.000432
	(0.000494)	(0.000406)	(0.000421)
Siblings	0.0255	0.0269**	0.0323**
	(0.0200)	(0.0108)	(0.0150)
Education			
Primary	0.00766	0.0228	-0.0667
	(0.228)	(0.0849)	(0.0890)
Lower Secondary	-0.288	-0.247	-0.354**
	(0.199)	(0.175)	(0.157)
Upper-Secondary	-0.600***	-0.389***	-0.504***
	(0.118)	(0.0865)	(0.0901)
College	-0.847***	-0.430***	-0.649***
	(0.155)	(0.110)	(0.111)
Partner's Education			
Primary	-0.672**	-0.311**	-0.340**
	(0.241)	(0.143)	(0.148)
Lower Secondary	-0.175	-0.371*	-0.382*
	(0.226)	(0.198)	(0.207)
Upper-Secondary	-0.221	-0.227*	-0.274**
	(0.146)	(0.121)	(0.127)
College	-0.214	-0.239	-0.349**
	(0.179)	(0.149)	(0.150)
Mother's Education	yes	yes	yes
Father's Education	yes	yes	yes
Region	yes	yes	yes
Constant	-0.992	4.291	2.028
	(9.101)	(3.174)	(3.228)
Observations	771	1,419	1,429
R-squared	0.189	0.235	0.227

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.16. Poisson Model: How TFR in origin country affects fertility of second generation immigrants in France, for women that have finished their fertility cycle (> 40)

	(1)	(2)	(3)
TFR at birth	0.0520***		
	(0.0157)		
TFR at the end of education		0.0875***	
		(0.00659)	
TFR in 2008			0.129***
			(0.0109)
LFP in 2008	-0.000368	-0.00101*	-0.00492***
	(0.00106)	(0.000592)	(0.000474)
Age	-0.0567	-0.0700	-0.0253
	(0.174)	(0.0509)	(0.0510)
Squared age	0.000575	0.000661	0.000281
	(0.00199)	(0.000523)	(0.000518)
Age of partner	0.108***	0.0298	0.0269
	(0.0237)	(0.0195)	(0.0201)
Age of partner Squared	-0.00101***	-0.000236	-0.000213
	(0.000211)	(0.000151)	(0.000158)
Siblings	0.00958	0.00916***	0.0107**
	(0.00694)	(0.00345)	(0.00501)
Education			
Primary	0.00698	0.0137	-0.0230
	(0.0715)	(0.0253)	(0.0278)
Lower Secondary	-0.0926	-0.0779	-0.118**
	(0.0685)	(0.0607)	(0.0544)
Upper-Secondary	-0.216***	-0.140***	-0.183***
	(0.0418)	(0.0298)	(0.0308)
College	-0.330***	-0.171***	-0.259***
	(0.0583)	(0.0436)	(0.0431)
Partner's Education			
Primary	-0.234**	-0.0929*	-0.106**
	(0.0921)	(0.0490)	(0.0505)
Lower Secondary	-0.0567	-0.127*	-0.131*
	(0.0793)	(0.0743)	(0.0783)
Upper-Secondary	-0.0751	-0.0728*	-0.0910**
	(0.0504)	(0.0412)	(0.0434)
College	-0.0733	-0.0744	-0.115**
	(0.0651)	(0.0534)	(0.0540)
Mother's Education	yes	yes	yes
Father's Education	yes	yes	yes
Region	yes	yes	yes
Constant	-0.609	1.611	0.735
	(3.488)	(1.173)	(1.167)
Observations	771	1,419	1,429

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

3.8.2 All Marital Status Woman

Table 3.17. OLS Model: How TFR in origin country affects the number of second generation immigrants in France, sample **all marital status**

	(1)	(2)	(3)
TFR at birth	0.0585***		
	(0.0169)		
TFR at the end of education		0.141***	
		(0.0142)	
TFR in 2008			0.151***
			(0.0311)
LFP in 2008	-0.00221	-0.00521***	-0.00790***
	(0.00138)	(0.00102)	(0.000841)
Age			
18-20	-1.141***	-1.468***	-1.613***
	(0.265)	(0.229)	(0.239)
21-25	-0.898***	-1.104***	-1.243***
	(0.190)	(0.163)	(0.169)
26-30	-0.646***	-0.741***	-0.858***
	(0.182)	(0.134)	(0.137)
31-35	-0.242*	-0.257**	-0.341***
	(0.132)	(0.103)	(0.107)
36-40	-0.124	-0.100	-0.148*
	(0.0893)	(0.0806)	(0.0821)
More than 50		0.0472	0.0655
		(0.154)	(0.156)
Squared age	0.000108	-6.26e-05	-9.84e-06
	(0.000169)	(0.000130)	(0.000122)
Age of Partner	0.148***	0.0829***	0.0791***
	(0.0140)	(0.0123)	(0.0128)
Squared Age of Partner	-0.00149***	-0.000693***	-0.000635***
	(0.000217)	(0.000175)	(0.000174)
Siblings	0.0214**	0.0164	0.0236*
	(0.00875)	(0.0104)	(0.0133)
Education			
Primary	0.305**	0.142*	0.101
	(0.114)	(0.0780)	(0.0785)
Lower Secondary	-0.165**	-0.145**	-0.186***
	(0.0725)	(0.0690)	(0.0631)
Upper-Secondary	-0.443***	-0.362***	-0.432***
	(0.0751)	(0.0533)	(0.0554)
College	-0.721***	-0.565***	-0.671***
	(0.0838)	(0.0651)	(0.0644)
Partner's Education	yes	yes	yes
Mother's Education	yes	yes	yes
Father's Education	yes	yes	yes
Region	yes	yes	yes
Constant	-1.020**	0.364	0.680
	(0.369)	(0.378)	(0.412)
Observations	3,153	3,867	3,877
R-squared	0.371	0.375	0.366

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3.18. OLS Model: How LFP in origin country affects the number of hours worked of second generation migrants in France, sample all marital status **woman**

	(1)	(2)	(3)
LFP at birth	0.00558		
	(0.0277)		
LFP at the end of education		0.0593***	
		(0.0171)	
LFP in 2008			0.116***
			(0.0235)
TFR at the end of education	-0.603**	-0.500**	-0.261
	(0.221)	(0.197)	(0.266)
Age			
18-20	-5.145	-13.41***	-14.33***
	(6.444)	(3.706)	(3.609)
21-25	-1.216	-7.552**	-8.206**
	(5.859)	(3.586)	(3.368)
26-30	-0.0286	-5.659**	-5.890***
	(4.191)	(2.221)	(1.998)
31-35	0.774	-3.663**	-3.868**
	(3.490)	(1.707)	(1.546)
36-40	0.392	-2.963*	-2.788*
	(2.521)	(1.515)	(1.368)
More than 50		-3.173	-3.300
		(2.600)	(2.079)
Squared age	0.00666	0.00210	0.00131
	(0.00401)	(0.00184)	(0.00170)
Age of Partner	-0.634*	-0.221	-0.242
	(0.330)	(0.223)	(0.237)
Squared Age of Partner	0.00677*	0.000774	0.000748
	(0.00387)	(0.00217)	(0.00246)
Siblings	-0.388	-0.293*	-0.307
	(0.315)	(0.166)	(0.180)
Education			
Primary	-2.143	-0.0215	1.306
	(1.870)	(2.077)	(1.885)
Lower Secondary	5.329**	4.654***	4.620***
	(2.293)	(1.132)	(1.096)
Upper-Secondary	9.661***	8.331***	8.209***
	(0.965)	(0.951)	(0.864)
College	12.43***	11.24***	11.39***
	(0.988)	(1.078)	(1.050)
Partner's Education	yes	yes	yes
Mother's Education	yes	yes	yes
Father's Education	yes	yes	yes
Region	yes	yes	yes
Constant	24.85**	26.92***	25.59***
	(10.96)	(7.892)	(8.153)
Observations	2,551	3,637	3,867
R-squared	0.113	0.087	0.093

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

3.8.3 Labor Force Participation for both man and woman

Table 3.19. OLS Model: How LFP in origin country affects the number of hours worked of second generation migrants in France, sample **all women and man**

	(1)	(2)	(3)
LFP at birth	0.0371		
	(0.0229)		
LFP at the end of education		0.0492***	
		(0.0119)	
LFP in 2008			0.0903***
			(0.0126)
TFR at the end of education	-0.807***	-0.779***	-0.607***
	(0.245)	(0.127)	(0.168)
Age			
18-20	-11.94***	-23.80***	-25.31***
	(3.815)	(2.119)	(2.013)
21-25	-2.644	-13.52***	-14.79***
	(4.157)	(1.553)	(1.501)
26-30	0.841	-8.572***	-9.412***
	(2.804)	(1.384)	(1.125)
31-35	1.119	-5.798***	-6.387***
	(2.068)	(0.796)	(0.726)
36-40	1.750	-2.924***	-3.177***
	(1.374)	(0.701)	(0.652)
More than 50		-3.711***	-3.480***
		(1.207)	(1.180)
Squared age	0.0142***	0.00475***	0.00349***
	(0.00290)	(0.000728)	(0.000694)
Age of Partner	-1.310***	-0.747***	-0.745***
	(0.272)	(0.126)	(0.124)
Squared Age of Partner	0.00848**	0.00190	0.00176
	(0.00319)	(0.00161)	(0.00157)
Siblings	-0.169	-0.0765	-0.0855
	(0.140)	(0.0833)	(0.0884)
Education			
Primary	-3.203**	0.994	1.951
	(1.373)	(1.264)	(1.293)
Lower Secondary	2.794	2.250*	2.486**
	(2.078)	(1.172)	(1.135)
Upper-Secondary	4.890***	5.071***	5.133***
	(1.101)	(0.674)	(0.580)
College	6.524***	6.644***	6.770***
	(1.115)	(0.721)	(0.644)
Partner's Education	yes	yes	yes
Mother's Education	yes	yes	yes
Father's Education	yes	yes	yes
Region	yes	yes	yes
Constant	43.44***	50.20***	50.09***
	(7.890)	(3.604)	(3.089)
Observations	4,789	7,082	7,552
R-squared	0.088	0.079	0.082

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

3.8.4 Discrimination

Table 3.20. OLS Model: How discrimination affects hours worked of second generation immigrants in France, for both man and woman

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
LFP in 2008	0.0978***	0.100***	0.0976***	0.108***	0.0969***	0.141***	0.0964***	0.0999***	0.0953***	0.106***
	(0.0146)	(0.0149)	(0.0145)	(0.0227)	(0.0151)	(0.0149)	(0.0158)	(0.0151)	(0.0149)	(0.0116)
Age discrimination		-1.438								
		(2.963)								
Age discrimination * LFP		-0.150***								
		(0.0453)								
Health discrimination			-10.30							
			(9.261)							
Health discrimination * LFP			0.105							
			(0.147)							
Origin discrimination				1.598						
				(2.341)						
Origin d *LFP				-0.0777						
				(0.0534)						
Place of living discrimination					-4.401					
					(3.883)					
Place of living discrimination * LFP					0.0124					
					(0.0852)					
Feeling of being French						7.003***				
						(0.699)				
Feeling of being French * LFP						-0.0588***				
						(0.0166)				
Religion discrimination							-2.409			
							(5.811)			
Religion d * LFP							0.00313			
							(0.177)			
Gender discrimination								5.506		
								(5.261)		
Gender d * LFP								-0.153*		
								(0.0795)		
Clothes discrimination									-10.20**	
									(4.102)	
Clothes d * LFP									0.0655	
									(0.108)	
Skin color discrimination										2.274
										(3.439)
Skin d * LFP										-0.0791
										(0.0785)
Age dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Squared age	0.00294***	0.00302***	0.00296***	0.00287***	0.00294***	0.00265***	0.00292***	0.00294***	0.00293***	0.00292***
	(0.000730)	(0.000739)	(0.000734)	(0.000723)	(0.000734)	(0.000685)	(0.000734)	(0.000723)	(0.000738)	(0.000731)
Age of Partner	-0.722***	-0.725***	-0.723***	-0.724***	-0.738***	-0.590***	-0.725***	-0.727***	-0.714***	-0.725***
	(0.126)	(0.128)	(0.126)	(0.121)	(0.125)	(0.116)	(0.127)	(0.129)	(0.125)	(0.128)
Squared Age of Partner	0.00151	0.00154	0.00152	0.00152	0.00167	0.000288	0.00155	0.00158	0.00144	0.00154
	(0.00159)	(0.00163)	(0.00158)	(0.00156)	(0.00159)	(0.00161)	(0.00159)	(0.00162)	(0.00158)	(0.00160)
Siblings	-0.205**	-0.203**	-0.205**	-0.189**	-0.204**	-0.211**	-0.203**	-0.207**	-0.201**	-0.177*
	(0.0787)	(0.0780)	(0.0788)	(0.0803)	(0.0781)	(0.0946)	(0.0783)	(0.0784)	(0.0784)	(0.0880)
Education	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Partner's Education	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Mother's Education	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Father's Education	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Region	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	47.97***	47.78***	48.01***	47.76***	48.42***	40.61***	48.14***	47.99***	47.89***	47.68***
	(3.047)	(3.068)	(3.061)	(2.827)	(3.071)	(2.896)	(3.058)	(3.092)	(3.033)	(3.294)
Observations	7,578	7,578	7,578	7,578	7,578	7,419	7,578	7,578	7,578	7,578
R-squared	0.081	0.082	0.081	0.082	0.081	0.088	0.081	0.081	0.082	0.081

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3.21. OLS Model: How discrimination affects the number of children of second generation migrants in France, for all woman (married and unmarried)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
TFR at the end of education	0.148***	0.148***	0.149***	0.149***	0.148***	0.202***	0.145***	0.148***	0.162***
	(0.0109)	(0.0105)	(0.0108)	(0.0109)	(0.0110)	(0.0159)	(0.0115)	(0.0109)	(0.0123)
Age discrimination		-0.299							
		(0.204)							
Age discrimination * TFR		0.0203							
		(0.0731)							
Health discrimination			-0.213						
			(0.270)						
Health discrimination * TFR			-0.0533*						
			(0.0300)						
Origin discrimination				-0.0958					
				(0.122)					
Origin discrimination * TFR				0.00854					
				(0.0348)					
Place of living discrimination					-0.0781				
					(0.463)				
Place of living discrimination * TFR					-0.00872				
					(0.176)				
Feeling of being French						0.206***			
						(0.0671)			
Feeling of being French * TFR						-0.0776***			
						(0.0208)			
Religion discrimination							-0.517		
							(0.384)		
Religion discrimination * TFR							0.220**		
							(0.0903)		
Clothes discrimination								0.400	
								(0.524)	
Clothes discrimination * TFR								0.0230	
								(0.171)	
Skin color discrimination									0.000867
									(0.178)
Skin discrimination * TFR									-0.0613*
									(0.0327)
Age Dummies	yes	yes	yes	yes	yes	yes	yes	yes	yes
Education	yes	yes	yes	yes	yes	yes	yes	yes	yes
Partner's Education	yes	yes	yes	yes	yes	yes	yes	yes	yes
Mother's Education	yes	yes	yes	yes	yes	yes	yes	yes	yes
Father's Education	yes	yes	yes	yes	yes	yes	yes	yes	yes
Region	yes	yes	yes	yes	yes	yes	yes	yes	yes
Constant	0.374	0.371	0.377	0.393	0.382	0.231	0.367	0.378	0.320
	(0.390)	(0.388)	(0.390)	(0.391)	(0.395)	(0.392)	(0.395)	(0.387)	(0.401)
Observations	3,893	3,893	3,893	3,893	3,893	3,807	3,893	3,893	3,893
R-squared	0.371	0.371	0.371	0.371	0.371	0.371	0.372	0.372	0.373

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Principal Component Analysis of the Discrimination Index:

We construct a discrimination index based on Age, Health, Origin, Place of living, Religion, Outfit and Skin Color discrimination following a principal component analysis. The idea is to create a summary indice considering all types of discrimination that second generation immigrants face. When studying the effect of cultural norms in the origin country on the fertility of the respondent, the index of discrimination is constructed by the following equation:

Discrimination 1 = 0.23 * Age discrimination + 0.16 * Health Discrimination + 0.14 * Origin discrimination + 0.13 * Place of living discrimination + 0.12 * Religion discrimination + 0.10 * Outfit discrimination + 0.08 * Skin color discrimination

When we study the effect of cultural norms on labor working time, we also account for gender discrimination when creating the discrimination index. We calculate the index with the following equation:

Discrimination 2 = 0.21 * Age discrimination + 0.15 * Health discrimination+ 0.13 * Origin discrimination + 0.12 * Place of living discrimination + 0.11 * Religion discrimination + 0.10 * Gender discrimination + 0.08 * Outfit discrimination + 0.07 * Skin color discrimination

3.8.5 Father's country of origin

Table 3.22. OLS Model: How TFR in Father's origin country affect number of children of second generation migrants in France

	(1)	(1.b)	(2)	(2.b)	(3)	(3.b)
TFR at birth (father)	0.0737***	0.0880***				
	(0.0160)	(0.0136)				
TFR at the end of education (father)			0.165***	0.182***		
			(0.0140)	(0.0127)		
TFR in 2008 (father)					0.211***	0.197***
					(0.0398)	(0.0293)
LFP in 2008 (father)	-0.00339**		-0.00678***		-0.0110***	
	(0.00128)		(0.000857)		(0.000854)	
Age						
18-20	-1.343***	-1.319***	-1.723***	-1.728***	-1.976***	-1.956***
	(0.344)	(0.366)	(0.311)	(0.308)	(0.334)	(0.324)
21-25	-0.948***	-0.951***	-1.209***	-1.245***	-1.428***	-1.461***
	(0.213)	(0.221)	(0.192)	(0.186)	(0.206)	(0.200)
26-30	-0.690***	-0.694***	-0.834***	-0.856***	-1.009***	-1.029***
	(0.201)	(0.211)	(0.158)	(0.153)	(0.170)	(0.168)
31-35	-0.261*	-0.261	-0.310**	-0.325**	-0.427***	-0.445***
	(0.151)	(0.155)	(0.123)	(0.119)	(0.136)	(0.134)
36-40	-0.174	-0.177	-0.158	-0.162*	-0.220**	-0.224**
	(0.110)	(0.110)	(0.0983)	(0.0928)	(0.105)	(0.101)
More than 50			0.0843	0.0885	0.130	0.126
			(0.155)	(0.166)	(0.147)	(0.167)
Squared age	0.000123	0.000104	-7.84e-05	-0.000136	-2.02e-05	-0.000103
	(0.000195)	(0.000210)	(0.000137)	(0.000132)	(0.000129)	(0.000128)
Age of Partner	0.135***	0.131***	0.0601***	0.0480**	0.0573***	0.0428**
	(0.0170)	(0.0172)	(0.0179)	(0.0187)	(0.0179)	(0.0186)
Squared Age of Partner	-0.00140***	-0.00136***	-0.000522**	-0.000384*	-0.000480**	-0.000303
	(0.000230)	(0.000240)	(0.000193)	(0.000188)	(0.000191)	(0.000183)
Siblings	0.0173*	0.0186**	0.0133	0.0151	0.0185	0.0265*
	(0.00868)	(0.00800)	(0.00908)	(0.00913)	(0.0135)	(0.0137)
Education						
Primary	0.223*	0.267**	0.135*	0.146*	0.0844	0.0857
	(0.127)	(0.107)	(0.0754)	(0.0747)	(0.0761)	(0.0801)
Lower Secondary	-0.175*	-0.160*	-0.148*	-0.146*	-0.215**	-0.215**
	(0.0894)	(0.0880)	(0.0858)	(0.0827)	(0.0812)	(0.0837)
Upper-Secondary	-0.392***	-0.397***	-0.295***	-0.308***	-0.376***	-0.410***
	(0.0626)	(0.0661)	(0.0439)	(0.0435)	(0.0488)	(0.0518)
College	-0.645***	-0.654***	-0.447***	-0.451***	-0.579***	-0.607***
	(0.0879)	(0.0914)	(0.0734)	(0.0743)	(0.0745)	(0.0816)
Partner's Education	yes	yes	yes	yes	yes	yes
Mother's Education	yes	yes	yes	yes	yes	yes
Father's Education	yes	yes	yes	yes	yes	yes
Region	yes	yes	yes	yes	yes	yes
Constant	-0.585	-0.686	1.089*	1.110*	1.428**	1.443**
	(0.427)	(0.428)	(0.568)	(0.587)	(0.607)	(0.611)
Observations	2,293	2,309	2,938	2,958	2,948	2,968
R-squared	0.292	0.293	0.302	0.299	0.293	0.280

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.23. OLS Model: How LFP in Father's origin country affect number of hours worked of second generation woman migrants in France

	(1)	(2)	(3)	(4)	(5)	(6)
LFP at birth (father)	-0.0158	-0.0126				
	(0.0256)	(0.0232)				
LFP at the end of education (father)			0.0542**	0.0653***		
			(0.0199)	(0.0226)		
LFP in 2008 (father)					0.124***	0.131***
					(0.0398)	(0.0402)
TFR at the end of education (father)	-0.901**		-0.721**		-0.343	
	(0.393)		(0.282)		(0.276)	
Age						
18-20	-3.357	-2.793	-15.39**	-14.76**	-15.32**	-15.10**
	(9.766)	(9.778)	(5.948)	(6.157)	(5.436)	(5.423)
21-25	0.738	1.208	-9.634*	-9.229*	-9.757*	-9.641*
	(8.612)	(8.533)	(5.144)	(5.223)	(4.772)	(4.708)
26-30	2.125	2.463	-6.411*	-6.078*	-6.177**	-6.078**
	(6.464)	(6.375)	(3.086)	(3.120)	(2.729)	(2.719)
31-35	1.745	1.936	-4.337	-4.124	-4.310*	-4.263*
	(5.234)	(5.146)	(2.679)	(2.698)	(2.371)	(2.416)
36-40	1.186	1.183	-2.807	-2.663	-2.446	-2.420
	(3.386)	(3.327)	(1.647)	(1.660)	(1.500)	(1.511)
More than 50			-3.563	-3.469	-3.220	-3.223
			(3.457)	(3.449)	(2.628)	(2.631)
Squared age	0.0103*	0.0108**	0.00240	0.00260	0.00159	0.00161
	(0.00518)	(0.00508)	(0.00239)	(0.00241)	(0.00199)	(0.00198)
Age of Partner	-0.167	-0.200	0.0731	0.0562	0.0523	0.0542
	(0.465)	(0.457)	(0.420)	(0.420)	(0.377)	(0.380)
Squared Age of Partner	0.00119	0.00116	-0.00192	-0.00210	-0.00208	-0.00231
	(0.00484)	(0.00477)	(0.00401)	(0.00399)	(0.00367)	(0.00361)
Siblings	-0.336	-0.481	-0.256*	-0.378**	-0.304**	-0.364**
	(0.279)	(0.331)	(0.131)	(0.172)	(0.141)	(0.159)
Education						
Primary	-1.994	-2.249	-0.202	-0.0754	1.052	1.010
	(1.693)	(1.718)	(2.049)	(2.071)	(1.924)	(1.928)
Lower Secondary	6.439**	6.592**	5.272***	5.415***	5.262***	5.333***
	(2.568)	(2.520)	(1.231)	(1.261)	(1.305)	(1.305)
Upper-Secondary	10.67***	11.22***	8.368***	8.770***	8.228***	8.425***
	(1.063)	(1.097)	(0.830)	(0.822)	(0.738)	(0.680)
College	11.18***	12.05***	9.540***	10.19***	10.07***	10.40***
	(0.909)	(1.087)	(1.163)	(1.058)	(1.204)	(1.083)
Partner's Education	yes	yes	yes	yes	yes	yes
Mother's Education	yes	yes	yes	yes	yes	yes
Father's Education	yes	yes	yes	yes	yes	yes
Region	yes	yes	yes	yes	yes	yes
Constant	10.15	7.863	19.93	17.59	18.12	16.98
	(15.60)	(15.77)	(13.38)	(13.87)	(12.75)	(12.43)
Observations	1,790	1,790	2,746	2,746	2,938	2,948
R-squared	0.135	0.132	0.093	0.092	0.097	0.097

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

3.8.6 "Horse race" of the cultural norm variables

Table 3.24. OLS Model: How TFR in origin country affect number of children of second generation woman migrants in France

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
TFR at birth	0.0728***				-0.0722***	0.00973	-0.0727***
	(0.0161)				(0.0217)	(0.0350)	(0.0220)
TFR at the end of education		0.171***		0.163***	0.189***		0.172***
		(0.0155)		(0.0502)	(0.0383)		(0.0491)
TFR in 2008			0.210***	0.0156		0.142	0.0307
			(0.0426)	(0.115)		(0.0893)	(0.120)
LFP in 2008	-0.00348**	-0.00635***	-0.0105***	-0.00657***	-0.00788***	-0.00717***	-0.00828***
	(0.00123)	(0.000881)	(0.000818)	(0.00156)	(0.000843)	(0.00186)	(0.00184)
Age							
18-20	-1.411***	-1.734***	-1.980***	-1.747***	-1.568***	-1.522***	-1.577***
	(0.380)	(0.310)	(0.329)	(0.349)	(0.372)	(0.425)	(0.396)
21-25	-1.005***	-1.230***	-1.445***	-1.242***	-1.090***	-1.085***	-1.100***
	(0.227)	(0.186)	(0.198)	(0.195)	(0.223)	(0.254)	(0.240)
26-30	-0.754***	-0.860***	-1.037***	-0.869***	-0.762***	-0.792***	-0.769***
	(0.212)	(0.149)	(0.160)	(0.168)	(0.198)	(0.223)	(0.215)
31-35	-0.283*	-0.307**	-0.431***	-0.313**	-0.278*	-0.302*	-0.283*
	(0.152)	(0.110)	(0.121)	(0.122)	(0.141)	(0.159)	(0.151)
36-40	-0.184*	-0.155	-0.220**	-0.158	-0.177*	-0.188*	-0.179*
	(0.106)	(0.0929)	(0.0972)	(0.0932)	(0.0977)	(0.106)	(0.0995)
More than 50		0.0959	0.130	0.0969			
		(0.147)	(0.141)	(0.141)			
Squared age	7.32e-05	-0.000105	-3.83e-05	-9.81e-05	-5.09e-05	0.000128	-2.77e-05
	(0.000207)	(0.000127)	(0.000118)	(0.000126)	(0.000207)	(0.000175)	(0.000162)
Age of Partner	0.130***	0.0573***	0.0559***	0.0570***	0.140***	0.133***	0.139***
	(0.0187)	(0.0173)	(0.0176)	(0.0174)	(0.0182)	(0.0181)	(0.0180)
Squared Age of Partner	-0.00134***	-0.000487**	-0.000457**	-0.000485**	-0.00147***	-0.00140***	-0.00147***
	(0.000256)	(0.000194)	(0.000192)	(0.000198)	(0.000240)	(0.000225)	(0.000236)
Siblings	0.0192**	0.0145	0.0205	0.0140	0.0112	0.0119	0.0104
	(0.00830)	(0.00990)	(0.0140)	(0.0126)	(0.00973)	(0.0112)	(0.0117)
Education							
Partner's Education	yes	yes	yes	yes	yes	yes	yes
Mother's Education	yes	yes	yes	yes	yes	yes	yes
Father's Education	yes	yes	yes	yes	yes	yes	yes
Region	yes	yes	yes	yes	yes	yes	yes
Constant	-0.418	1.155**	1.471**	1.164**	-0.171	-0.342	-0.177
	(0.466)	(0.553)	(0.588)	(0.545)	(0.471)	(0.468)	(0.469)
Observations	2,290	2,938	2,948	2,938	2,290	2,290	2,290
R-squared	0.299	0.311	0.299	0.311	0.310	0.304	0.311

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.25. OLS Model: How LFP in origin country affects hours worked of second generation immigrant in France

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
LFP at birth	-0.0226				-0.193***	-0.114***	-0.111*
	(0.0241)				(0.0406)	(0.0264)	(0.0619)
LFP at the end of education		0.0380*		-0.145***	0.237***		-0.00870
		(0.0193)		(0.0404)	(0.0363)		(0.110)
LFP in 2008			0.109**	0.239***		0.198***	0.203**
			(0.0385)	(0.0649)		(0.0231)	(0.0725)
TFR at the end of education	-1.023**	-0.726***	-0.370	-0.328	-0.655*	-0.506*	-0.506**
	(0.409)	(0.254)	(0.278)	(0.209)	(0.321)	(0.244)	(0.243)
Age							
18-20	-5.850	-15.83**	-15.66***	-16.69***	-3.883	-5.653	-5.720
	(8.516)	(5.845)	(5.390)	(5.609)	(8.444)	(8.356)	(7.861)
21-25	-2.222	-10.36*	-10.55*	-11.28*	-1.043	-2.509	-2.560
	(7.940)	(5.501)	(5.280)	(5.421)	(8.012)	(7.880)	(7.421)
26-30	0.0485	-6.719**	-6.529**	-7.167**	0.992	0.111	0.0778
	(5.888)	(3.185)	(2.977)	(3.225)	(6.001)	(5.952)	(5.656)
31-35	0.116	-5.005*	-4.894*	-5.351*	0.633	0.178	0.161
	(4.826)	(2.755)	(2.506)	(2.793)	(4.817)	(4.830)	(4.665)
36-40	0.682	-3.125*	-2.808*	-3.114	0.777	0.883	0.884
	(3.194)	(1.733)	(1.566)	(1.822)	(3.213)	(3.156)	(3.170)
More than 50		-2.377	-2.474	-2.403			
		(3.362)	(2.776)	(3.213)			
Squared age	0.00754	0.00166	0.00111	6.99e-05	0.00723	0.00577	0.00574
	(0.00458)	(0.00246)	(0.00220)	(0.00215)	(0.00477)	(0.00459)	(0.00428)
Age of Partner	-0.240	0.187	0.124	0.123	-0.412	-0.370	-0.367
	(0.470)	(0.414)	(0.384)	(0.411)	(0.476)	(0.484)	(0.487)
Squared Age of Partner	0.00249	-0.00329	-0.00287	-0.00288	0.00417	0.00359	0.00356
	(0.00500)	(0.00391)	(0.00374)	(0.00397)	(0.00519)	(0.00527)	(0.00533)
Siblings	-0.370	-0.263*	-0.333*	-0.333*	-0.334	-0.344	-0.344
	(0.297)	(0.148)	(0.161)	(0.174)	(0.302)	(0.301)	(0.307)
Education							
Primary	-1.920	-0.212	0.860	0.0171	-1.478	-0.887	-0.877
	(1.618)	(2.110)	(1.961)	(1.968)	(1.565)	(1.402)	(1.366)
Lower Secondary	5.894**	5.293***	5.204***	5.424***	6.097**	6.129**	6.128**
	(2.526)	(1.207)	(1.297)	(1.110)	(2.617)	(2.568)	(2.552)
Upper-Secondary	10.08***	8.332***	8.323***	8.381***	9.730***	9.686***	9.688***
	(0.999)	(0.861)	(0.728)	(0.763)	(1.045)	(1.018)	(1.036)
College	10.70***	9.749***	10.18***	10.04***	10.17***	10.36***	10.37***
	(0.695)	(1.186)	(1.199)	(1.195)	(0.766)	(0.729)	(0.780)
Partner's Education	yes	yes	yes	yes	yes	yes	yes
Mother's Education	yes	yes	yes	yes	yes	yes	yes
Father's Education	yes	yes	yes	yes	yes	yes	yes
Region	yes	yes	yes	yes	yes	yes	yes
Constant	16.76	19.82	18.42	19.90	16.31	16.04	16.04
	(14.94)	(13.89)	(13.61)	(14.33)	(15.10)	(15.28)	(15.30)
Observations	1,860	2,756	2,938	2,745	1,860	1,860	1,860
R-squared	0.132	0.089	0.095	0.099	0.142	0.146	0.146

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

3.8.7 Descriptive Statistics

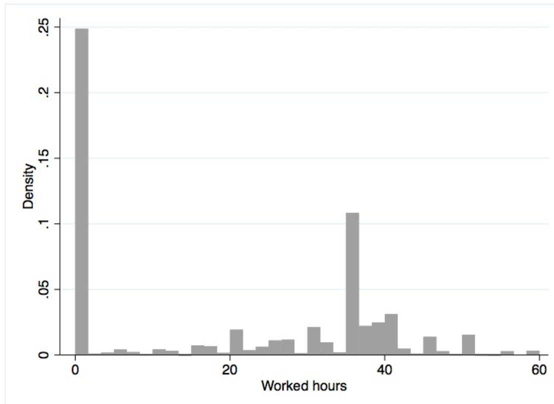


Figure 3.1. Hours of work distribution

Note: Excess of zeros, for this reason we use a Hurdle Model in studying the effect on cultural norms on labor work time

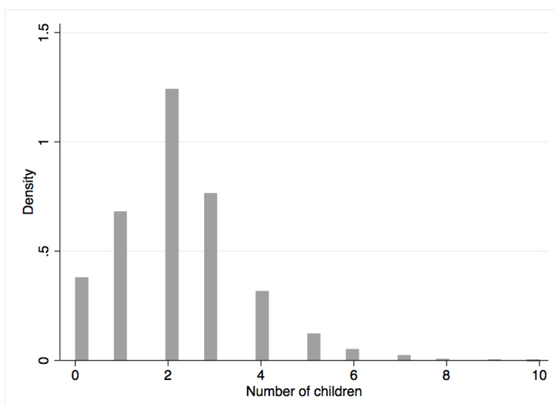


Figure 3.2. Number of children distribution

Note: No excess of zeros, we proceed with a Poisson Model as a robustness check in studying whether cultural norms have an effect on the number of children.

3.8.8 Results considering income

	(1)	(1.b)	(2)	(2.b)	(3)	(3.b)
TFR at the end of education			0.156***	0.168***		
			(0.0132)	(0.0105)		
TFR at birth	0.0645***	0.0798***				
	(0.0143)	(0.0114)				
TFR in 2008					0.187***	0.165***
					(0.0419)	(0.0306)
LFP in 2008	-0.00351***		-0.00599***		-0.00971***	
	(0.00115)		(0.000923)		(0.000886)	
Age						
18-20	-1.501***	-1.459***	-1.842***	-1.872***	-2.053***	-2.072***
	(0.433)	(0.446)	(0.354)	(0.347)	(0.371)	(0.356)
21-25	-1.041***	-1.025***	-1.283***	-1.325***	-1.470***	-1.511***
	(0.249)	(0.249)	(0.180)	(0.180)	(0.195)	(0.192)
26-30	-0.785***	-0.771***	-0.903***	-0.925***	-1.056***	-1.075***
	(0.239)	(0.238)	(0.153)	(0.148)	(0.165)	(0.159)
31-35	-0.297	-0.287	-0.335**	-0.352***	-0.442***	-0.461***
	(0.177)	(0.174)	(0.118)	(0.118)	(0.131)	(0.131)
36-40	-0.202*	-0.197*	-0.184*	-0.193**	-0.240**	-0.251**
	(0.113)	(0.112)	(0.0906)	(0.0877)	(0.0948)	(0.0929)
More than 50			0.0984	0.0982	0.127	0.117
			(0.130)	(0.144)	(0.126)	(0.148)
Squared age	7.51e-05	7.41e-05	-0.000103	-0.000151	-3.83e-05	-0.000114
	(0.000216)	(0.000220)	(0.000126)	(0.000119)	(0.000113)	(0.000114)
Age of partner	0.129***	0.124***	0.0588***	0.0464**	0.0579***	0.0436**
	(0.0186)	(0.0190)	(0.0158)	(0.0163)	(0.0162)	(0.0161)
Age of partner Squared	-0.00133***	-0.00127***	-0.000517**	-0.000378**	-0.000492**	-0.000321*
	(0.000261)	(0.000270)	(0.000186)	(0.000175)	(0.000185)	(0.000169)
Siblings	0.0176**	0.0173**	0.0126	0.0132	0.0185	0.0246*
	(0.00800)	(0.00718)	(0.00964)	(0.00986)	(0.0135)	(0.0137)
Income						
Uncomfortable	-0.367*	-0.350*	-0.281	-0.275	-0.322**	-0.355**
	(0.188)	(0.186)	(0.167)	(0.168)	(0.147)	(0.153)
Sufficient	-0.479***	-0.477***	-0.380**	-0.389***	-0.396***	-0.436***
	(0.136)	(0.134)	(0.135)	(0.135)	(0.125)	(0.126)
Tight	-0.331***	-0.325***	-0.198*	-0.204*	-0.200**	-0.232**
	(0.114)	(0.113)	(0.0995)	(0.0989)	(0.0918)	(0.0927)
Hard time	-0.146	-0.119	0.0936	0.119	0.0993	0.116
	(0.111)	(0.108)	(0.108)	(0.112)	(0.105)	(0.107)
Education						
Primary	0.259**	0.263**	0.140*	0.134*	0.0906	0.0789
	(0.108)	(0.112)	(0.0732)	(0.0724)	(0.0749)	(0.0800)
Lower Secondary	-0.166	-0.146	-0.140	-0.129	-0.189**	-0.176*
	(0.0968)	(0.0967)	(0.0887)	(0.0863)	(0.0838)	(0.0869)
Upper-Secondary	-0.390***	-0.388***	-0.296***	-0.301***	-0.369***	-0.389***
	(0.0607)	(0.0644)	(0.0417)	(0.0398)	(0.0434)	(0.0427)
College	-0.604***	-0.607***	-0.409***	-0.408***	-0.528***	-0.543***
	(0.0939)	(0.0945)	(0.0760)	(0.0731)	(0.0697)	(0.0710)
Partner's Education	yes	yes	yes	yes	yes	yes
Mother's Education	yes	yes	yes	yes	yes	yes
Father's Education	yes	yes	yes	yes	yes	yes
Region	yes	yes	yes	yes	yes	yes
Constant	-0.00454	-0.130	1.379**	1.446**	1.658***	1.754***
	(0.568)	(0.584)	(0.555)	(0.561)	(0.582)	(0.570)
Observations	2,277	2,292	2,916	2,935	2,926	2,945
R-squared	0.309	0.308	0.322	0.317	0.311	0.299

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3.26. OLS Model: How fertility in origin country affects fertility of second generation immigrant in France

	(1)	(1.b)	(2)	(2.b)	(3)	(3.b)
LFP at birth	-0.0256	-0.0228				
	(0.0255)	(0.0233)				
LFP in 2008					0.118***	0.121***
					(0.0318)	(0.0281)
LFP at the end of education			0.0544***	0.0609***		
			(0.0180)	(0.0166)		
TFR at the end of education	-0.691**		-0.463**		-0.122	
	(0.304)		(0.206)		(0.340)	
Age						
18-20	-7.011	-6.639	-14.83**	-14.38**	-14.28***	-14.26**
	(6.863)	(6.894)	(5.766)	(5.918)	(4.985)	(5.074)
21-25	-3.524	-3.260	-9.841**	-9.562**	-9.713***	-9.716***
	(5.179)	(5.140)	(3.711)	(3.705)	(3.343)	(3.182)
26-30	-1.205	-1.034	-6.402**	-6.175**	-5.909**	-5.911**
	(3.703)	(3.688)	(2.596)	(2.627)	(2.102)	(2.100)
31-35	-1.563	-1.451	-5.411**	-5.240**	-5.054***	-5.070***
	(2.729)	(2.717)	(2.174)	(2.221)	(1.666)	(1.750)
36-40	-0.404	-0.447	-3.162*	-3.057*	-2.672*	-2.685*
	(1.822)	(1.830)	(1.615)	(1.647)	(1.328)	(1.365)
More than 50			-2.363	-2.328	-2.636	-2.626
			(2.308)	(2.309)	(1.835)	(1.818)
Squared age	0.00465	0.00493	0.000150	0.000268	-0.000223	-0.000252
	(0.00292)	(0.00291)	(0.00194)	(0.00200)	(0.00161)	(0.00161)
Age of Partner	-0.259	-0.285	0.139	0.125	0.0601	0.0653
	(0.480)	(0.470)	(0.296)	(0.298)	(0.277)	(0.279)
Age of partner Squared	0.00341	0.00340	-0.00229	-0.00235	-0.00161	-0.00175
	(0.00543)	(0.00534)	(0.00285)	(0.00281)	(0.00273)	(0.00267)
Siblings	-0.284	-0.391	-0.145	-0.220*	-0.206	-0.227**
	(0.236)	(0.249)	(0.118)	(0.107)	(0.128)	(0.103)
Income						
Uncomfortable	7.681***	8.292***	4.917**	5.404**	5.224**	5.350**
	(2.661)	(2.750)	(2.155)	(2.143)	(2.025)	(2.009)
Sufficient	7.796**	8.257***	7.075***	7.406***	7.559***	7.642***
	(2.756)	(2.852)	(1.729)	(1.786)	(1.517)	(1.590)
Tight	4.828	5.192	4.575**	4.826***	5.128***	5.161***
	(3.056)	(3.145)	(1.648)	(1.691)	(1.537)	(1.613)
ard time	-0.0663	0.100	-0.643	-0.551	0.302	0.329
	(2.654)	(2.710)	(1.642)	(1.681)	(1.617)	(1.631)
Education						
Primary	-1.601	-1.745	-0.0527	0.0418	1.054	0.967
	(1.521)	(1.496)	(1.313)	(1.318)	(1.378)	(1.439)
Lower Secondary	3.803**	3.911***	4.698***	4.740***	4.813***	4.827***
	(1.348)	(1.348)	(1.033)	(1.062)	(0.867)	(0.877)
Upper-Secondary	9.410***	9.805***	8.360***	8.592***	8.487***	8.554***
	(0.939)	(0.889)	(0.725)	(0.711)	(0.552)	(0.545)
College	9.925***	10.50***	9.835***	10.20***	10.41***	10.52***
	(0.744)	(0.678)	(0.668)	(0.709)	(0.729)	(0.793)
Partner's Education	yes	yes	yes	yes	yes	yes
Mother's Education	yes	yes	yes	yes	yes	yes
Father's Education	yes	yes	yes	yes	yes	yes
Region	yes	yes	yes	yes	yes	yes
Constant	15.52	13.57	16.28	14.64	14.21	13.79
	(14.02)	(14.11)	(10.84)	(10.72)	(9.936)	(9.409)
Observations	1,851	1,851	2,735	2,735	2,916	2,926
R-squared	0.189	0.187	0.149	0.148	0.157	0.158

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3.27. OLS Model: How LFP in origin country affects LFP of second generation immigrant in France

3.8.9 Magnitude of coefficients

Table 3.28. Magnitude of the coefficients in our smaple

TFR OLS Model	Size	LFP OLS Model	Size
TFR at birth	0.14	LFP at birth	0.48
TFR end of education	0.32	LFP end of education	1.05
TFR in 2008	0.25	LFP 2008	2.41
LFP in 2008	0.64-1.9	TFR 2008	0.38-1.19
Siblings	0.06-0.09	Siblings	0.59-1.03
Age of partner	0.3-1.3	Age of partner	0.93-2.8
Age		Age	
18-20	0.97-1.3	18-20	5.58-11.28
21-25	1.28-1.84	21-25	5.01-13.37
26-30	1.04-1.44	26-30	2.07-9.6
31-35	0.40-0.62	31-35	2.62-8.3
36-40	0.25-0.30	36-400.82-4.89	0.82-4.89

3.8.10 Summary Statistics by County

Table 3.29. Summary statistics by Country of Origin

Country	Number of children	TFR at birth	TFR at end of educ	TFR 2008	Hours worked	LFP at birth	LFP at end of educ	LFP 2008
Algeria	1.26	6.94	4.17	2.72	23.17	4.37	8.20	13.13
Austria	1.31	2.2	1.46	1.42	33.68	40.46	45.50	53.38
Belgium	1.35	2.07	1.62	1.85	26.10	29.74	38.36	46.91
Benin	0.86	6.82	6.24	5.46	20.65	53.08	61.21	75.68
Bulgaria	1.05	2.13	1.54	1.56	20.94	59.44	50.64	48.31
Burkina Faso	1.24	6.76	6.73	6.03	30.55	51.82	73.78	81.17
Cambodia	0.87	6.18	4.49	3.01	25.02	75.57	77.38	76.04
Cameroon	1.22	6.33	6	5.21	21.55	45.20	56.14	76.55
Central African	1.33	5.91	5.55	5.28	19.06	62.06	64.57	.
Chad	0 .68	6.86	7.23	6.80	24.25	23.35	55.58	55.43
Congo, Dem, Rep,	1.71	6.40	6.68	6.62	18.95	.	69.13	67.50
Congo, Rep,	1.62	5.93	5.05	4.79	19.92	52.89	58.57	51.66
Cote d'Ivoire	1.28	7.50	6.27	5.22	22.26	54.09	49.06	58.45
Czech Republic	1.16	2.10	1.50	1.51	34.16	.	51.97	49.32
Denmark	1.3	2.04	1.67	1.89	32.6	48.85	59.46	59.90
Equatorial Guine	2.40	5.78	5.90	5.36	23	.	56.53	.
Finland	1	1.94	1.73	1.85	147	46.92	58.82	57.38
France	1.03	2.23	1.84	2.01	27.94	43.46	46.64	51.10
Gabon	0.77	5.49	4.59	4.18	14.06	52.28	51.86	45.33
Gambia	2.2	6.23	6.02	5.64	15.6	58.24	56.15	49.26
Germany	1.18	2.03	1.36	1.38	31.83	41.71	45.82	52.52
Ghana	1.92	6.62	5.96	4.35	26.42	58.20	54.25	68.36
Greece	1.15	2.16	1.47	1.5	25.94	30.17	36.02	42.42
Guinea	2.30	6.36	6.27	5.49	15.08	44.40	45.16	.
Guinea-Bissau	2.05	6.24	6.33	5.17	21.63	2.13	2.21	.
Hungary	1.29	1.99	1.64	1.35	23.94	44.77	49.89	42.73
Ireland	2.2	3.54	2.33	2.06	5.2	30.23	35.31	56.82
Italy	1.26	2.20	1.41	1.45	31.86	27.71	34.19	38.66
Lao PDR	1.02	6.16	4.83	3.32	24.36	.	80.60	72.51
Latvia	1	1.92	1.56	1.58	20	64.06	53.54	55.00
Lithuania	1	2.13	1.39	1.45	17.50	.	52.75	50.51
Luxembourg	1.13	1.78	1.60	1.61	26.73	29.26	38.60	48.05
Mali	1.50	7.13	6.93	6.67	19.12	22.73	40.25	68.32
Mauritania	2.03	6.45	5.52	5.05	22.54	17.67	26.14	28.18
Morocco	1.13	5.78	3.46	2.54	21.10	15.99	24.29	26.56
Netherlands	1.41	2.43	1.59	1.77	24.32	28.20	42.91	58.33
Nigeria	1.33	6.55	6.20	5.90	20.75	34.03	45.27	53.27
Poland	0 .92	2.36	1.75	1.39	32.9	59.44	55.06	46.61
Portugal	1.28	2.64	1.72	1.39	31.13	33.14	48.74	55.83
Romania	0.97	2.59	1.57	1.6	25.44	59.71	54.31	46.80
Senegal	1.31	7.08	5.83	5.09	19.51	35.42	41.13	38.09
Slovak Republic	1.83	2.48	1.70	1.34	30	.	53.09	51.00
Slovenia	1.6	2.20	1.52	1.53	25	.	51.41	53.53
Spain	1.27	2.62	1.57	1.45	29.74	21.62	33.49	50.35
Sweden	1.11	1.93	1.72	1.91	32.66	49.40	64.86	59.80
Togo	1.04	6.91	5.82	5.00	26.60	46.32	68.25	79.20
Tunisia	1.31	5.65	3.17	2.05	24.10	17.58	21.51	24.73
Turkey	1.53	4.61	2.98	2.18	20.75	46.55	32.46	24.01
United Kingdom	1.37	2.38	1.76	1.91	24.03	41.51	49.46	55.97
Vietnam	1.05	5.46	3.16	1.91	25.09	73.51	70.72	71.01

3.8.11 Alternative Age Controls

Table 3.30. OLS Model: How fertility in origin country affects number of children of second generation woman immigrant in France

	(1)	(1.b)	(2)	(2.b)	(3)	(3.b)
TFR at birth	0.0732***	0.0881***				
	(0.0166)	(0.0133)				
TFR at the end of education			0.175***	0.189***		
			(0.0173)	(0.0133)		
TFR in 2008					0.216***	0.199***
					(0.0443)	(0.0319)
LFP in 2008	-0.00339**		-0.00620***		-0.0104***	
	(0.00122)		(0.000855)		(0.000864)	
Age dummies						
18-20	-1.606***	-1.557***	-1.555***	-1.480***	-1.921***	-1.792***
	(0.211)	(0.216)	(0.226)	(0.229)	(0.262)	(0.248)
21-25	-1.164***	-1.147***	-1.072***	-1.034***	-1.395***	-1.323***
	(0.108)	(0.111)	(0.0930)	(0.0979)	(0.104)	(0.106)
26-30	-0.899***	-0.885***	-0.801***	-0.766***	-1.067***	-1.005***
	(0.0671)	(0.0705)	(0.0590)	(0.0639)	(0.0692)	(0.0718)
31-35	-0.406***	-0.394***	-0.323***	-0.298***	-0.513***	-0.472***
	(0.0676)	(0.0724)	(0.0625)	(0.0660)	(0.0647)	(0.0713)
36-40	-0.249***	-0.242***	-0.188***	-0.171***	-0.293***	-0.266***
	(0.0721)	(0.0699)	(0.0586)	(0.0569)	(0.0574)	(0.0597)
More than 50			0.0603	0.0400	0.156	0.109
			(0.158)	(0.167)	(0.142)	(0.162)
Partner's Age dummies						
18-20	-0.333	-0.281	-0.235	-0.157	-0.274	-0.190
	(0.387)	(0.393)	(0.391)	(0.410)	(0.387)	(0.416)
21-25	-0.750***	-0.737***	-0.687***	-0.665***	-0.693***	-0.682***
	(0.168)	(0.173)	(0.162)	(0.173)	(0.157)	(0.176)
26-30	-0.512***	-0.490***	-0.454***	-0.418***	-0.468***	-0.429***
	(0.0703)	(0.0719)	(0.0728)	(0.0730)	(0.0764)	(0.0744)
31-35	-0.172*	-0.162	-0.118	-0.0924	-0.126	-0.0940
	(0.0887)	(0.0943)	(0.0894)	(0.0954)	(0.0869)	(0.0977)
36-40	-0.0173	-0.0173	0.0258	0.0334	0.0148	0.0210
	(0.0603)	(0.0590)	(0.0610)	(0.0609)	(0.0610)	(0.0605)
More than 50	0.0403	0.0545	-0.0396	-0.0374	-0.00436	0.00892
	(0.185)	(0.189)	(0.126)	(0.130)	(0.114)	(0.124)
Siblings	0.0191**	0.0194**	0.0158	0.0172	0.0216	0.0292*
	(0.00840)	(0.00758)	(0.00972)	(0.00999)	(0.0139)	(0.0141)
Education						
Primary	0.275**	0.279**	0.147*	0.138*	0.0974	0.0802
	(0.107)	(0.111)	(0.0771)	(0.0743)	(0.0766)	(0.0801)
Lower Secondary	-0.188*	-0.167*	-0.164	-0.154	-0.222**	-0.211**
	(0.0938)	(0.0919)	(0.0964)	(0.0915)	(0.0866)	(0.0855)
Upper-Secondary	-0.404***	-0.403***	-0.314***	-0.320***	-0.398***	-0.424***
	(0.0597)	(0.0628)	(0.0466)	(0.0459)	(0.0506)	(0.0518)
College	-0.667***	-0.675***	-0.477***	-0.480***	-0.616***	-0.643***
	(0.0860)	(0.0857)	(0.0813)	(0.0804)	(0.0799)	(0.0842)
Partner's Education	yes	yes	yes	yes	yes	yes
Mother's Education	yes	yes	yes	yes	yes	yes
Father's Education	yes	yes	yes	yes	yes	yes
Region	yes	yes	yes	yes	yes	yes
Constant	2.828***	2.596***	2.545***	2.220***	3.001***	2.589***
	(0.143)	(0.112)	(0.0892)	(0.0822)	(0.0909)	(0.0701)
Observations	2,289	2,304	2,937	2,956	2,947	2,966
R-squared	0.297	0.296	0.314	0.308	0.302	0.286

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3.31. OLS Model: How LFP in origin country affects hours worked of second generation immigrant in France

	(1)	(1.b)	(2)	(2.b)	(3)	(3.b)
LFP at birth	-0.0256	-0.0203				
	(0.0237)	(0.0206)				
LFP at the end of education			0.0402*	0.0524**		
			(0.0212)	(0.0235)		
LFP in 2008					0.111**	0.120***
					(0.0409)	(0.0419)
TFR at the end of education	-1.121***		-0.811***		-0.465	
	(0.373)		(0.281)		(0.301)	
Age dummies						
18-20	-17.17***	-17.11***	-18.70***	-18.08***	-17.50***	-17.09***
	(4.191)	(4.236)	(4.222)	(4.213)	(4.130)	(4.181)
21-25	-12.44***	-12.50***	-12.96***	-12.63***	-12.25***	-12.02***
	(3.203)	(3.264)	(2.534)	(2.556)	(2.606)	(2.621)
26-30	-8.666***	-8.792***	-9.008***	-8.750***	-8.047***	-7.830***
	(1.955)	(1.975)	(1.345)	(1.263)	(1.211)	(1.188)
31-35	-6.139***	-6.305***	-6.403***	-6.247***	-5.588***	-5.458***
	(1.917)	(1.901)	(1.524)	(1.498)	(1.471)	(1.487)
36-40	-2.422	-2.734	-3.306***	-3.232***	-2.462**	-2.374**
	(1.724)	(1.678)	(1.112)	(1.124)	(1.073)	(1.087)
More than 50			-2.801	-2.586	-3.641*	-3.716*
			(2.278)	(2.291)	(1.764)	(1.832)
Partner's Age dummies						
18-20	0.499	1.649	0.805	1.874	2.238	2.912
	(6.096)	(6.167)	(6.614)	(6.779)	(6.956)	(7.433)
21-25	1.178	1.874	1.776	2.205	2.554	2.891
	(3.479)	(3.434)	(2.834)	(2.749)	(2.700)	(2.661)
26-30	0.881	1.399	1.474	1.836	2.187	2.439
	(1.839)	(1.795)	(1.732)	(1.743)	(1.645)	(1.694)
31-35	2.372	2.633	3.164**	3.391**	3.621**	3.802***
	(1.598)	(1.587)	(1.438)	(1.440)	(1.275)	(1.300)
36-40	-0.497	-0.163	0.156	0.355	0.190	0.321
	(1.031)	(1.043)	(1.157)	(1.153)	(1.025)	(1.062)
More than 50	2.546	1.863	0.735	0.273	0.853	0.563
	(1.716)	(1.737)	(2.352)	(2.190)	(2.212)	(2.070)
Siblings	-0.367	-0.547	-0.262*	-0.404**	-0.339*	-0.422**
	(0.301)	(0.348)	(0.151)	(0.187)	(0.164)	(0.177)
Education						
Primary	-1.686	-1.937	-0.200	-0.0448	0.844	0.807
	(1.655)	(1.632)	(2.061)	(2.082)	(1.894)	(1.911)
Lower Secondary	5.743**	5.941**	5.183***	5.277***	5.149***	5.202***
	(2.529)	(2.482)	(1.249)	(1.272)	(1.347)	(1.352)
Upper-Secondary	9.991***	10.69***	8.255***	8.714***	8.285***	8.559***
	(0.966)	(1.012)	(0.911)	(0.886)	(0.771)	(0.700)
College	10.65***	11.71***	9.661***	10.42***	10.16***	10.62***
	(0.736)	(0.928)	(1.208)	(1.088)	(1.223)	(1.068)
Partner's Education	yes	yes	yes	yes	yes	yes
Mother's Education	yes	yes	yes	yes	yes	yes
Father's Education	yes	yes	yes	yes	yes	yes
Region	yes	yes	yes	yes	yes	yes
Constant	25.60***	21.93***	24.20***	20.78***	19.55***	17.47***
	(2.137)	(1.381)	(1.591)	(1.588)	(2.262)	(2.054)
Observations	1,859	1,859	2,755	2,755	2,937	2,947
R-squared	0.132	0.129	0.089	0.086	0.095	0.094

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

3.8.12 Hurdle Model

Table 3.32. Hurdle Model: : How LFP in origin country affects FLFP of second generation immigrant in France, horse-race version

Variables	Outcome: Hours Worked	Single Selection	lnsigma
LFP at birth	0.00538	-0.0117***	
	(0.0334)	(0.00297)	
LFP at the end of education	0.0889**	-0.00235	
	(0.0385)	(0.00618)	
LFP in 2008	-0.0815***	0.0202***	
	(0.0313)	(0.00499)	
TFR at the end of education	-0.315	0.00285	
	(0.260)	(0.0281)	
Number of children		-0.276***	
		(0.0380)	
Age			
18-20	-8.685*	-0.745	
	(5.088)	(0.566)	
21-25	-4.791	-0.273	
	(3.218)	(0.481)	
26-30	-2.965	-0.0803	
	(2.885)	(0.379)	
31-35	-2.600	-0.0235	
	(2.245)	(0.263)	
36-40	-1.685	0.0494	
	(1.381)	(0.157)	
Squared age	-0.00145	0.000468*	
	(0.00287)	(0.000279)	
Age of partner	-0.240	0.0205	
	(0.330)	(0.0367)	
Age of partner Squared	0.00322	-0.000246	
	(0.00412)	(0.000415)	
Siblings	-0.107	-0.0137	
	(0.246)	(0.0134)	
Education			
Primary	-0.172	0.0219	
	(1.709)	(0.140)	
Lower Secondary	1.501	0.298**	
	(1.634)	(0.133)	
Upper-Secondary	3.009***	0.607***	
	(0.939)	(0.0912)	
College	4.343***	0.589***	
	(1.090)	(0.0861)	
Partner's Education			
Primary	3.830***	0.436***	
	(1.069)	(0.154)	
Lower Secondary	-0.307	-0.0174	
	(1.626)	(0.134)	
Upper-Secondary	0.452	-0.0647	
	(1.028)	(0.101)	
College	-0.115	-0.176*	
	(1.039)	(0.102)	
Mother's Education	yes	yes	
Father's Education	yes	yes	
Region	yes	yes	
Constant	42.04***	-0.661	2.313***
	(8.015)	(1.147)	(0.0337)
Observations	1,855	1,855	1,855

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 3.33. Average marginal effects on hours worked, horse-race version

Average marginal Effects	Delta-method - dy/dx	Std. Err.	z	P>|z|	95% Conf. Interval
LFP at birth	-0.1205242	0.0327524	-3.68	0.000	-0.1847177 : 0.0563307
LFP at the end of education	0.0271247	0.0630629	0.43	0.667	-0.0964764 : 0.1507258
LFP in 2008	0.1652409	0.0506003	3.27	0.001	0.0660662 : 0.2644156
TFR at the end of education	-0.1539793	0.2840591	-0.54	0.588	-0.7107248 : 0.4027663
Number of children	-2.909889	0.3878679	-7.50	0.000	-3.670096 : -2.149682

3.8.13 The interlink between culture and economic decisions

Weber studied culture and economic evolution and he argued that the transcendental premise of any science of culture is in the circumstance that we are people of culture. His full statement is: *"We are men of culture, endowed with the capacity and the will of assuming a conscious position in the face of the world conferring sense to it"*. He also came up with the idea of subculture, where groups like race, ethnicity, religion, region, occupation, gender, and sexual preference live a certain lifestyle based on different values and norms. Others like Durkeim held the belief that culture has logical, functional, and historical relationships to society. Other thoughts on culture are from Marx who focused on the struggle between groups of people with diverging economic interests whose economy determined the cultural structure of values and ideologies. He believed that the economic aspect affects the cultural one, different from Hegel who believed that culture affects the material aspect.

The literature has established the relevance and power of culture in behavior, decision-making, beliefs, and outcomes, but there is still a gap in exploiting the heterogeneity of cultural norms. This depends of course on how we define culture, as its definition can be ambiguous in different contexts. The main cultural traits used in economics are values and beliefs.[20] There are

[20] Alesina and Giuliano (2015) distinguishes between formal and informal institutions by reviewing work to assess the presence of a two-way causal effect between culture and institutions. The first one is the legal system and regulations of a country, while the second can be viewed as trust and interaction between people. These two are strongly interlinked

different ways along with problematics in measuring culture. Surveys are a common way of studying culture although they come with causality problematics. Nevertheless, studying people from heterogenous backgrounds, in the same institutional environment, allows capturing the cultural traits that influence people's decisions and economic outcomes. Cultural values can be so strong as to influence not only migrants but also the second-generation descendants in the host country.

3.8.14 French Institutional Context

France had an excellent reputation for welcoming immigrants during the onset of the Industrial Revolution. This was related to a significant migrant wave between 1830 and 1850. Six million people died in France during the First World War. Immigration was unrestricted up to the conclusion of World War I, when 4.3 million foreigners entered France between 1850 and 1920. 6 % of the population, or little under 3 million people, were immigrants from French colonies. Italy, Spain, Belgium, and Switzerland were among the bordering nations who immigrated, and all four soon merged into the French culture and population.[21]

France experienced a labor deficit following World War II, and in order to make up for this shortfall, the French government hired immigrants from all over the world. After Vietnam gained its independence from France and Algeria's war, there was a significant influx of immigrants from both countries in 1954. Immigrants were more prevalent during the course of the 20th century. Between the 1920s and 1960s, there were around 300,000 new arrivals per year. Only in 1970 did the migratory growth pattern start to slow and become more steady. Immigration from Europe began to decline in the second part of the century, while African and Asian immigrants rose. The integration and assimilation procedures that immigrants and second-generation immigrants had

and they both affect one another. They are complementary and interact.

[21] According to INSEE data published in 2008 the migration composition was: Italians (35 %) were followed by Poles (20 %), Spanish (15 %), Belgians (10 %), and a smaller number of people from central and eastern European countries.

to go through in France and beyond are described in Barou (2018) and Scullion (1995). The majority of immigrants to France were still from European nations like Portugal, Italy, and Spain, but more recent migratory flows occurred from North Africa, particularly Algeria (which was a part of France until 1962) and the former protectorates of Morocco and Tunisia. French or former French holdings in Central Africa, Asia, and the Americas came after them. Due to this significant number of immigrants, France became a favorite country for refugees seeking asylum. Immigrants worked in the steel, construction, mining, agricultural, and automobile industries.

During the 1970s, France had an economic crisis, and in 1974 it began formulating immigration policies that included limiting immigration from its former colonies. Moreover, France and other European Union nations withdrew from accepting asylum seekers in the latter half of the 20th century. Racial prejudice and social antagonism were prevalent in political climates during economic downturns.From the middle of the 1980s onward, France adopted an integration program, allocating government funds to organizations that urged immigrants to assimilate and follow French law while preserving their own cultures and customs. Since the 1970s, the issue of immigrant assimilation and societal cohesiveness has gained attention once more. Social injustices and calls for cultural appreciation posed an economic threat to the French republican paradigm. These reasons made racial discrimination more pervasive, increased pressure from the populist extreme right and ongoing assimilation, and pushed for an integrative strategy prior to 1970. In conclusion, the history of immigration in France is rather complex. Baudin et al. (2021) and provide a thorough overview of it.

4

Conclusions

This thesis contributes to the literature of population and family economics, more specifically to topics of sex selection due to son preference, migration incentives and cultural transmission of migrants. It exploits key drivers of decision making related to these topics by relying both on theoretical and empirical investigations. The main results are that i) Gender empowerment and equal investment in children cannot lower biased sex ratios down to their biological values. Cultural norms can be important aspects of abortion decisions explaining this residual ii) Culture matters and determines behavior and decision making when it comes to fertility and labor market outcomes of migrants iii) Soil quality affects people's decision to migrate through its underlying mechanism of agricultural productivity. These three chapters rely on the richness of datasets like DHS, LSMS, Mosaic and TeO. It is only through the power of interdisciplinary research and data collection on fertility, gender specific differentials, investment in children, climate data, soil quality data and track of second generation migrants that we can put these topics together and study drivers of behavior and decision making regarding fertility and migration outcomes.

Bibliography

Abramitzky, Ran, Adeline Delavande, and Luis Vasconcelos, "Marrying Up: The Role of Sex Ratio in Assortative Matching," *American Economic Journal: Applied Economics*, 2011, *3*, 124–157.

Abrevaya, J, "Are There Missing Girls in the United States? Evidence from Birth Data," *American Economic Journal: Applied Economics*, 2009, *1* (2), 1–34.

Afulani, Patience and Joseph Asunka, "Socialization, Adaptation, Transnationalism, and the Reproductive Behavior of Sub-Saharan African Migrants in France," *Population Research and Policy Review*, 2015, *34*, 561–592.

Alesina, Alberto and Paola Giuliano, "Culture and Institutions," *Journal of Economic Literature*, 2015, *53* (4), 898–944.

_ , _ , and Nathan Nunn, "On the Origins of Gender Roles: Women and the Plough," *The Quarterly Journal of Economics*, 2013, *128* (2), 469–530.

Almond, Douglas, Lena Edlund, and K Milligan, "Son Preference and the Persis- tence of Culture: Evidence from South and East Asian Immigrants to Canada," *Population and Development Review*, 2013, *39* (1), 75–95.

Ambrosetti, E, L Ortensi, C Ciniza, and A Marina, "Sex imbalances at birth in migratory context: Evidence from Italy," *Genus*, 2015, *71*, 29–51.

Anglewicz, Philip, Mark VanLandingham, Lucinda Manda-Taylor, and Hans-Peter Kohler, "Cohort profile: internal migration in sub-Saharan Africa—the Migration and Health in Malawi (MHM) study," *BMJ Open*, 2017, *7* (e014799).

Ashraf, Quamrul and Oded Galor, "The 'Out of Africa Hypothesis', Human Genetic Diversity, and Comparative Economic Development," *The American Economic Review*, 2013, *103* (1), 1–46.

Barber, Nigel, "The Sex Ratio as a Predictor of Cross-National Variation in Violent Crime.," *Cross-Cultural Research*, 2000, *34* (3), 264–282.

Barbier, Edward and Jacob P. Hochard, "Does Land Degradation Increase Poverty in Developing Countries?," 2016.

Barou, Jacques, "Integration of immigrants in France: a historical perspective," *Identities*, 2018, *21* (6), 642–657.

Barrios, Salvador, Luisito Bertinelli, and Eric Strobl, "Climatic change and rural-urban migration: The case of sub-Saharan Africa," *Journal of Urban Economics*, 2006, *60* (3).

Basu, Deepankar and Robert De Jong, "Son Targeting Fertility Behavior: Some Consequences and Determinants," *Demography*, 2 2010, *47*, 521–536.

Baudin, Thomas, Simone Morriconi, and Dylan Planard, "Discrimination, work and fertility among second generation migrants in France," *Working Paper*, 2021.

Becker, Gary, *Economics of the Family: Marriage, Children, and Human Capital*, National Bureau of Economic Research, 1974.

Behrman, Julia and Abigail Weitzman, "Point of Reference: A Multi-sited Exploration of African Migration and Fertility in France," *Working Paper*, 2019.

Beine, Michel and Christopher Parsons, "Climatic Factors as Determinants of International Migration," *Scandinavian Journal of Economics*, 2015, *117* (2), 723–767.

Ben-Porath, Y and F Welch, "Do sex preferences really matter?," *The Quarterly Journal of Economics*, 1976, *90* (2), 285–307.

Berson, Clemence, "Private vs. Public Sector: Discrimination against Second-Generation Immigrants in France," *Documents de travail du Centre d'Economie de la Sorbonne 09059*, 2009.

Bertoli, Simone and Francesca Marchetta, "Migration, Remittances and Poverty in Ecuador," *The Journal of Development Studies*, 2012, *50* (8), 1067–1089.

Bethman, Dirk and Michael Kvasnicka, "World War II, Missing Men and Out of Wedlock Childbearing," *Economic Journal*, 2013, *123*, 162–194.

Blau, Francine D., "Immigrants and gender roles: assimilation vs. culture," *IZA Journal of Migration*, 2015, *4* (23).

Bohra-Mishra, Pratikshya, Michael Oppenheimer, and Solomon M. Hsiang, "Nonlinear permanent migration response to climatic variations but minimal response to disasters," *National Academy of Sciences*, 2014, *111* (27), 9780–9785.

Brainerd, Elizabeth, "The lasting effect of sex ratio imbalance on marriage and family: Evidence from World War II in Russia," *The Review of Economics and Statistics*, 2017, *99* (2), 229–242.

Brausmann, Alexandra, Moritz Flubacher, and Filippo Lechthaler, "Valuing meteorological services in resource-constrained settings: Application to smallholder farmers in the Peruvian Altiplano," *Working Paper, CER-ETH - Center of Economic Research (CER-ETH) at ETH Zurich*, 2019.

Call, Maia and Clark Gray, "Climate anomalies, land degradation, and rural out-migration in Uganda," *Population and Environment*, 2020, *41*, 507–528.

Chabé-Ferret, Bastien, “Adherence to cultural norms and economic incentives: Evidence from fertility timing decisions,” *Journal of Economic Behaviour and Organization*, 2019, *162* (24-48).

Chahnazarian, A, “Determinants of the sex ratio at birth: Review of recent literature,” *Social Biology*, 1988, *35* (3-4), 214–235.

Chen, Lanyan, *Gender and Chinese Development:Towards an equitable society*, London and New York: Routledge, 2008.

Chort, Isabelle, Philippe De Vreyer, and Thomas Zuber, “Enduring Gendered Mobility Patterns in Contemporary Senegal,” *Working Paper*, 2018.

Coniglio, Nicola and Giovanni Pesce, “Climate variability and international migration: an empirical analysis,” *Environment and Development Economics*, 2015, *20*, 434–468.

Dallmann, Ingrid and Kathrin Millock, “Climate Variability and Inter-State Migration in India,” *CESifo Economic Studies*, 2017, *63* (4), 560–594.

Das-Gupta, M, “Selective Discrimination against Female Children in Rural Punjab, India,” *Population and Development Review*, 1987, *13* (1), 77–100.

_ , J Zhenghua, X Li-Bohua, W Chung, and H Bae, “Why is Son Preference so Persistent in East and South Asia? A Cross-Country Study of China, India, and the Republic of Korea.,” *World Bank Policy Research Working Paper*, 2003, (2942).

_ , W Chung, and S Li, “Evidence for an Incipient Decline in Numbers of Missing Girls in China and India,” *Population and Development Review*, 2009, *35* (2), 401–416.

Das-Gupta, Monica, “Explaining Asia’s "Missing Women": A New Look at the Data,” *Population and Development Review*, 2005, *31* (3), 529–535.

_ , “”Missing Girls” in the South Caucasus Countries Trends, Possible Causes, and Policy Options,” *Policy Research Working Paper*, 2015, (7236).

Dell, Melissa, Benjamin Jones, and Benjamin Olken, "Temperature and Income: Reconciling new cross-sectional and panel estimates," *American Economic Review: Papers and Proceedings*, 2009, *99* (2), 198–204.

_ , _ , and _ , "Temperature Shocks and Economic Growth: Evidence from the Last Half Century," *American Economic Journal: Macroeconomics*, 2012, *4* (3), 66–95.

_ , _ , and _ , "What Do We Learn from the Weather? The New Climate-Economy Literature," *Journal of Economic Literature*, 2014, *52* (3), 740–798.

Deschênes, Olivier and Michael Greenstone, "The Economic Impacts of Climate Change: Evidence from Agricultural Output and Random Fluctuations in Weather," *The American Economic Review*, 2007, *97* (1), 354–385.

Dillon, Andrew and Lakshman Nagraj Rao, "Land Measurement Bias: Comparisons from Global Positioning System, Self-Reports, and Satellite Data," *ADB Economics Working Paper Series*, 2018.

_ , Sydney Gourlay, Kevin McGee, and Gbemisola Oseni, "Land Measurement Bias and its Empirical Implications: Evidence from a Validation Exercise," *Policy Research Working Papers*, 2016.

Docquier, Frédéric, Giovanni Peri, and Ilse Ruyssen, "The Cross-country Determinants of Potential and Actual Migration," *International Migration Review*, 2014, *48* (1), S37–S99.

Dubuc, Sylvie and Koleman, "An Increase in the Sex Ratio of Births to India-born Mothers in England and Wales: Evidence for Sex-Selective Abortion," *Population and Development Reviewopulation and Development Review*, 2007, *33* (2), 383–400.

Dustmann, Christian and Anna Okatenko, "Out-migration, wealth constraints, and the quality of local amenities," *Journal of Development Economics*, 2014, *110*, 52–63.

Edlund, L, I Hongbin, J Yi, and J Zhang, "Sex Ratios and Crime: Evidence from China," *The Review of Economics and Statistics*, 2013, *95* (5), 1520–1534.

Edlund, Lena and Douglas Almond, "Son-biased sex ratios in the 2000 United States Census," *Proceedings of the National Academy of Sciences*, 2008, *105* (15), 5681–2.

Eitzinger, Anton, Claudia R. Binder, and Markus A. Meyer, "Risk perception and decision-making: do farmers consider risks from climate change?," *Climatic Change*, 2018, *151*, 507–524.

Englund, Harri, "Ethnography after Globalism: Migration and Emplacement in Malawi," *American Ethnologist*, 2002, *29* (2), 261–286.

Feng, Shuaizhang, Alan B. Krueger, and Michael Oppenheimer, "Linkages among climate change, crop yields and Mexico–US cross-border migration," *Proceedings of the National Academy of Sciences*, 2010, *107* (32), 14257–14262.

Fernandez, Raquel, "Cultural Change as Learning: The Evolution of Female Labor Force Participation over a Century," *American Economic Review*, 2013, *103* (1), 472–500.

_ and Alessandra Fogli, "Fertility: The Role of Culture and Family Experience," *Journal of the European Economic Association*, 2006, *4* (2-3), 552–561.

_ and _ , "Culture: An Empirical Investigation of Beliefs, Work, and Fertility," *American Economic Journal: Macroeconomics*, 2009, *1* (1), 146–177.

Ferrari, Giulia and Ariane Pailhé, "Transition to adulthood in France: do children of immigrants differ from natives ?," *Working Paper*, 2016.

Fogli, Alessandra and Laura Veldkamp, "Nature or Nurture? Learning and the Geography of Female Labor Force Participation," *Econometrica*, 2011, *79* (4), 1103–1138.

Frank, Kristyn and Feng Hou, “Beyond culture: source country female labour force participation and the earnings of immigrant women,” *Work, Employment and Society*, 2016, *30* (3), 410–435.

Gilligan, C, *n a Different Voice: Women’s Conceptions of Self and of Morality*, Harvard University Press, 1982.

Gleditsch, NP, Idean Salehyan, and Ragnhild Nordas, “Climate Change and Conflict: The Migration Link,” *International Peace Academy, New York*, 2007.

Gommes, R and F Petrassi, “Rainfall variability and drought in sub-Saharan Africa since 1960,” *FAO, Research and Technology Development Division, Agrometeorology Group*, 1996.

Gonzalez, Libertad, “Sex Selection and Health at Birth among Indian Immigrants,” 2014.

Goodkind, D, “The claim that China’s fertility restrictions contributed to the use of prenatal sex selection: A skeptical reappraisal,” *Population Studies*, 2015, *69* (3), 263–279.

Gorodzeisky, Anastasia and Moshe Semyonov, “Labor force participation, unemployment and occupational attainment among immigrants in West European countries,” *Plos One*, 2017, *12* (5), 1–22.

Gray, Clark, “Environment, land and rural out-migration in the southern Ecuadorian Andes.,” *World Development*, 2009, *37*, 457–468.

_ , “Soil quality and human migration in Kenya and Uganda,” 2011.

_ **and Valerie Mueller**, “Drought and Population Mobility in Rural Ethiopia,” *World Development*, 2012, *40* (1), 134–145.

_ **and** _ , “Drought and Population Mobility in Rural Ethiopia,” *World Development*, 2012, *40* (1), 134–145.

— **and** — , "Natural disasters and population mobility in Bangladesh," *Proceedings of the National Academy of Sciences*, 2012, *109* (16), 6000–6005.

Green, Erik, "Modern Agricultural History in Malawi: Perspectives on Policy-Choice Explanations," *African Studies Review*, 2007, *50* (3), 115–133.

Gröger, André and Yanos Zylberberg, "Internal Labor Migration as a Shock Coping Strategy: Evidence from a Typhoon," *American Economic Review*, 2016, *8* (2), 123–53.

Grossjean, Pauline and Rose Khatar, "It's Raining Men! Hallelujah? The long-run Consequences of male-biased sex ratios," *The Review of Economic Studies*, 2018, *86* (2), 723–754.

Guilmoto, Christophe, "High Sex Ratio at Birth in Eastern Europe," 2010.

— , "The masculinization of births. Overview and current knowledge," *Population*, 2015, *70*, 183–243.

— **and G Duthe**, "Masculinization of births in Eastern Europe," *Population and Societies*, 2013, (506), 4.

— **, Nora Dudwick, Arjan Gjonca, and Laura Rahm**, "How do demographic trends change? The onset of birth masculinization in Albania, Georgia, and Vietnam 1990–2005," *Population and Development Review*, 2018, *44* (1), 37–61.

— **, Valentine Becquet, and Catriona Dutreuilh**, "Sex Imbalane at Birth in Vietnam: Rapid Increase Followed by Stabilization," *Population*, 2018, *73* (3), 519–544.

— **, Xuyên Hoàng, and Toan Ngo Van**, "Recent increase in sex ratio at birth in Viet Nam," *Plos One*, 2010, *2*, 4.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "Does Culture Affect Economic Outcomes?," *Journal of Economic Perspectives*, 2006, *20* (2), 23–48.

Gutmann, Myron, Glenn Deane, Nathan Lauster, and Andrés Peri, “Two population-environment regimes in the Great Plains of the United States, 1930-1990,” *Population and Evironment*, 2005, *27* (2), 191–225.

Hazarika, Gautam, Chandan Kumar, and Jha Sudipta Sarangi, “Ancestral ecological endowments and missing women,” *Journal of Population Economics*, 2019, *32* (4), 1101–1123.

Hermans-Neumann, Kathleen and Henk Hilderink, “Opportunities and Challenges for Investigating the Environment-Migration Nexus,” *Human Ecology Interdisciplinary Journal*, 2015, *43* (2), 309–322.

_ , Joerg Priess, and Martin Herold, “Human migration, climate variability, and land degradation: hotspots of socio-ecological pressure in Ethiopia,” *Regional Environmental Change*, 2017, *17* (5), 1479–1492.

Himmelweit, Susan, Cristina Santos, Almudena Sevilla, and Catherine Sofer, “Sharing of Resources Within the Family and the Economics of Household Decision Making,” *Journal of Marriage and Family*, 06 2013, *75.*

Hirvonen, Kalle, “Temperature Changes, Household Consumption, and Internal Migration: Evidence from Tanzania,” *American Journal of Agricultural Economics*, 2016.

Homer-Dixon, Thomas, “On the Threshold: Environmental Changes as Causes of Acute Conflict,” *Internation Security*, 1991, *16* (2), 76–116.

_ , “Environmental scarcities and violent conflicst: Evidence from cases,” *Internation Security*, 1994, *19* (1), 5–40.

Hornbeck, Richard, “The Enduring Impact of the American Dust Bowl: Short- and Long-Run Adjustments to Environmental Catastrophe,” *American Economic Review*, 2012, *102* (4), 1477–1507.

Hsiang, Solomon M., “Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America,” *Proceedings of the National Academy of Sciences*, 2010, *107* (35), 15367–15372.

Hudson, Valerie and Andrea den Boer, *Bare Branches: The Security Implications of Asia's Surplus Male Population*, Cambridge, MA, USA: MIT Press, 2004.

IOM, "Migration in Malawi, a Country profile - International Organization for Migration," 2014.

IPCC, "The intergovernmental Panel on Climate Change," 2014.

Jacobsen, R, H Moller, and A Mouritsen, "Natural variation in the human sex ratio," *Human Reproduction*, 1999, *14* (12), 3120–5.

Jayet, Cyril, "Feeling French and Feeling being Seen as French The Relations between Two Dimensions of National Belonging," *Sociologie*, 2016, *7*, 113–132.

Kahn, Joan R., "Immigrant Selectivity and Fertility Adaptation in the United States," *Social Forces*, 1988, *67* (1), 108–128.

_ , "Immigrant and Native Fertility during the 1980s: Adaptation and Expectations for the Future," *The International Migration Review*, 1994, *28* (3), 501–519.

Kaser, Karl, *Patriarchy after patriarchy: gender relations in Turkey and in the Balkans, 1500-2000*, LIT Verlag Munster, 2008.

_ **, J Halpern, and R Wagner**, "Patriarchy in the Balkans: Temporal and Cross-Cultural Approaches. The History of the Family," *An International Quarterly*, 1996, *1* (4), 425–442.

Key, Felix, Qiaomei Fu, Frederic Romagne, and Michael Lachmann Aida Andres, "Human adaptation and population differentiation in the light of ancient genomes," *Nature Communications*, 2016, *10775* (7).

Kok, Suzanne, Nicole Bosch, Anja Deelen, and Rob Euwals, "Migrant Women on the Labour Market: On the Role of Home-and Host-Country Participation," *IZA Discussion Papers*, 2011, *5817.*

Kubik, Zaneta and Mathilde Maurel, "Weather Shocks, Agricultural Production and Migration: Evidence from Tanzania," *The Journal of Development Studies*, 2016, *52* (5), 665–680.

LeBlanc, David and Romain Perez, "The relationship between rainfall and human density and its implications for future water stress in sub-Saharan Africa," *Ecological Economics*, 2008, *66* (2-3), 319–336.

Lin, Kuo-Chuan, Steven P Hamburg, Lixin Wang, Chin-Tzer Duh, Chu-Mei Huang, Chung-Te Chang, and Teng-Chiu Lin, "Impacts of increasing typhoons on the structure and function of a subtropical forest: reflections of a changing climate," *Scientific reports*, 2017, *7* (1), 1–10.

Lochmann, Alexia, Hilel Rapoport, and Biagio Speciale, "The Effect of Language Training on Immigrants' Economic Integration: Empirical Evidence from France," *IZA Discussion Papers*, 2018.

Lynch, Katherine, "Why weren't (many) European women 'missing'?," *The History of the Family*, 3 2011, *16*, 250–266.

Malpede, Maurizio and Mario Percoco, "Long-term Economic Effects of Aridification on the GDPs of Africa and Asia," 2021.

Marchiori, Luca and Ingmar Schumacher, "When nature rebels: international migration, climate change, and inequality," *Journal of Population Economics*, 2011, *24* (2), 569–600.

_ , Jean-François Maystadt, and Ingmar Schumacher, "The impact of weather anomalies on migration in sub-Saharan Africa," *Journal of Environmental Economics and Management*, 2012, *63* (3), 355–374.

McKenzie, David and Hilel Rapoport, "Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico," *Journal of Development Economics*, 2007, *84* (1), 1–24.

Methorstab, Ron, Dirk Roep, Frans Verheesc, and Jos Verstegend, "Differences in farmers' perception of opportunities for farm development," *NJAS - Wageningen Journal of Life Sciences*, 2017, *81*, 9–18.

Millock, Katrin, "Migration and Environment," *Annual Review of Resource Economics*, 2015, *7* (35-60).

Mountford, Andrew and Hillel Rapoport, "The Brain Drain and the World Distribution of Income," *Journal of Development Economics*, 2011, *95* (1), 4–17.

_ **and** _ , "Globalization, Brain Drain, and Development," *Journal of Economic Literature*, 2012, *50* (3), 681–730.

Myers, Norman, "Environmental Refugees," *Population and Evironment*, 1997, *19* (2), 167–182.

Neuman, Emma, "Source country culture and labor market assimilation of immigrant women in Sweden: evidence from longitudinal data," *Review of Economics of the Household*, 2018, *16* (585-627).

Olsson, Lennart, Humberto Barbosa, Suruchi Bhadwal, Annete Cowie, Kenel Delusca, Dulce Flores-Renteria, Kathleen Hermans, Esteban Jobbagy, Werner Kurz, Diqiang Li et al., "Land Degradation: IPCC Special Report on Climate Change, Desertification, Land 5 Degradation, Sustainable Land Management, Food Security, and 6 Greenhouse gas fluxes in Terrestrial Ecosystems," 2019.

Pailhé, Ariane, "The convergence of second-generation immigrants' fertility patterns in France: The role of sociocultural distance between parents' and host country," *Demographic Research*, 2017, *36* (45), 1361–1398.

Pergament, E, Pb Toydemir, and M Fiddler, "Sex Ratio: A biological perspective of 'Sex and the City'," *Reprod Biomed Online*, 2002, *5* (1), 43–46.

Read, Jen'nan Ghazal, "Cultural Influences on Immigrant Women's Labor Force Participation: The Arab-American Case," *The International Migration Review*, 2004, *38* (1), 52–77.

Rubin, E, "The sex ratio at birth," *PubMed*, 1976, *21* (4), 45–48.

Schlenker, Wolfram and David B Lobell, "Robust negative impacts of climate change on African agriculture," *Environmental Research Letters*, 2010, *5* (1).

_ **and Michael Roberts**, "Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change.," *Proc Natl Acas Sci USA*, 2009, *106* (37), 15594–15598.

Scullion, Rosemarie, "Vicious Circles: Immigration and National Identity in Twentieth-Century France," *SubStance*, 1995, *24* (1/2), 30–48.

Selenica, Teki Emin, "L'Albanie en 1927," 1927.

Sills, E. and G. Palermo, "Preimplantation genetic diagnosis for elective sex selection, the IVF market economy, and the child–another long day's journey into night?," *Journal of assisted reproduction and genetics*, 2002, *19* (9), 433–7.

Silverman, Ay, Sr Stephens, Mt Drouin, Rg Zack, J Osborne, and Sa Ericsson, "Female sex selection using clomiphene citrate and albumin separation of human sperm," *Hum Reprod*, 2002, *17* (5), 1254–1256.

Simon, Patrick, "France and the Unknown Second Generation: Preliminary Results on Social Mobility," *The International Migration Review*, 2003, *37* (4), 1091–1119.

Singh, Pripp, Brekk, and Stray-Pederson, "Different sex ratios of children born to Indian and Pakistani immigrants in Norway," *BMC Pregnancy and Childbirth*, 2010, *10* (1), 40.

Stark, Oded and E. David Bloom, "The New Economics of Labor Migration," *The American Economic Review*, 1985, *75* (2), 173–178.

Stern, Nicholas, *The economics of climate change. The Stern Review*, UK Treasury, 2006.

Tapia, Francisco J. Beltrán, "Sex ratios and missing girls in late-19th-century Europe," *EHES Working Paper*, 2019, (160).

Thornton, Philip, Gopal Alagarswamy, Peter Jones, and Jeff Anderson, "Spatial variation of crop yield response to climate change in East Africa," *Global Environmental Change*, 2009, *19* (1), 54–65.

Torres, Marcello, Richard Howitt, and Lineu Rodrigues, "Analyzing rainfall effects on agricultural income: Why timing matters," *EconomiA*, 2019, *20* (1), 1–14.

Torres, Raul, Zachary Szpiech, and Ryan Hernandez, "Human demographic history has amplified the effects of background selection across the genome," *PLoS Genet.*, 2018, *6* (14), e1007387.

Tse, Chun-Wing, "Do Natural Disasters Lead to More Migration? Evidence from Indonesia," *Working Paper*, 2011.

Tylor, Edward Burnett, *Primitive Culture*, Cambridge University Press, 1871.

UNFPA, "UNFPA Annual Report," 2011.

_ , "UNFPA Annual Report," 2016.

Varma, R, "Technological Fix: Sex Determination in India," *Bulletin of Science, Technology and Society*, 2002, *22* (1), 21–30.

Wang, Wendy, "Son preference and educational opportunities of children in China— "I wish you were a boy!"," *Gender Issues*, 2005, *22* (2), 3–30.

Weil, David N., *Economic growth*, Addison-Wesley, 2004.

Wert, G De and W Dondorp, "Preconception sex selection for non-medical and intermediate reasons: ethical reflections.," *Facts, views and vision in ObGyn*, 2010, *2* (4), 267–277.

Wisner, Ben, Piers Blaikie, Terry Cannon, and Ian Davis, *At Risk: Natural Hazards, People's Vulnerability, and Disasters*, New York, NY: Routledge, 2004, 2nd edition, 2004.

World-Bank, "World Development Report 2010: Development and Climate Change," *The World Bank Group, Washington, DC*, 2009.

_ , “World Development Indicators database,” *Malawi - Second Agriculture Sector Wide Approach Support Project*, 2017.

Xing, Zh and Th Hesketh, “Abnormal sex ratios in human populations: Causes and consequences,” *Proceedings of the National Academy of Sciences*, 2006, *103* (36), 13271–13275.

Zhang, X and S Wei, “The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China,” *Journal of Political Economy*, 2011, *119* (3), 511–564.