



Essays on Development Economics

Iva Trako

► To cite this version:

Iva Trako. Essays on Development Economics. Economics and Finance. Université Paris sciences et lettres, 2018. English. NNT : 2018PSLEH043 . tel-03168276

HAL Id: tel-03168276

<https://theses.hal.science/tel-03168276>

Submitted on 12 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT

de l'Université de recherche Paris Sciences et Lettres
PSL Research University

Préparée à l'Ecole des hautes études en sciences sociales

Essais sur l'économie du développement

Ecole doctorale n°465

ECOLE DOCTORALE ECONOMIE PANTHEON SORBONNE

Spécialité SCIENCES ÉCONOMIQUES

Soutenue par Iva TRAKO
le 17 Mai 2018

Dirigée par **Karen MACOURS**

COMPOSITION DU JURY :

M. BOBONIS Gustavo
University of Toronto, Rapporteur

M. GILES John
The World Bank, Rapporteur

M. VANDEN EYNDE Oliver
EHESS - PSE, Membre du jury

Mme. VAN DE WALLE Dominique
The World Bank, Membre du jury

Mme. MACOURS Karen
EHESS - PSE, Membre du jury

Ecole des Hautes Etudes en Sciences Sociales

ECOLE D'ECONOMIE DE PARIS



THESE

Pour l'obtention du grade de docteur en Sciences Économiques
de l'École des Hautes Études en Sciences Sociales
Discipline: Sciences Economiques - Spécialité: Analyse et Politique Economiques

Présentée et soutenue publiquement à l'École d'Économie de Paris le 17 Mai 2018 par :

Iva TRAKO

Essais sur l'économie du développement

Directrice de thèse : Karen MACOURS

Composition du jury :

<i>Rapporteurs :</i>	Gustavo BOBONIS	Professeur Associé à l'Université de Toronto Chaire de Recherche du Canada en Economie du Développement
	John GILES	Economiste Principal à la Banque Mondiale Development Research Group
<i>Examineurs :</i>	Oliver VANDEN EYNDE	Professeur Associé à l'École d'Économie de Paris Chargé de Recherche au CNRS
	Dominique VAN DE WALLE	Economiste Principale à la Banque Mondiale Development Research Group
<i>Directrice :</i>	Karen MACOURS	Professeur Associé à l'École d'Économie de Paris Chargée de Recherche à l'INRA

Ecole des Hautes Etudes en Sciences Sociales

PARIS SCHOOL OF ECONOMICS



PhD Thesis

Submitted to École des Hautes Études en Sciences Sociales
For the Degree of Doctor of Philosophy in Economics
Field: Economics - Speciality: Analysis and Policy in Economics

Prepared and defended at the Paris School of Economics on May 17th, 2018 by:

Iva TRAKO

Essays on Development Economics

Thesis Advisor: Karen MACOURS

Committee:

<i>Reviewers :</i>	Gustavo BOBONIS	Associate Professor at University of Toronto St. George Campus Canada Research Chair in the Political Economy of Development
	John GILES	Lead Economist at the World Bank Development Research Group
<i>Discussants :</i>	Oliver VANDEN EYNDE	Associate Professor at Paris School of Economics Senior Research Fellow at CNRS
	Dominique VAN DE WALLE	Lead Economist at the World Bank Development Research Group
<i>PhD Advisor :</i>	Karen MACOURS	Associate Professor at Paris School of Economics Senior Research Fellow at INRA

© 2018 Iva Trako
All rights reserved.

Dissertation Advisor:
Professor Karen Macours

Author:
Iva Trako

Essays on Development Economics

Abstract

This thesis consists of three empirical essays that contribute to the literature on development economics. The first chapter exploits the introduction of women's justice centers (WJC) in Peru, a community based approach aimed at reducing violence against women, in order to examine whether improving access to justice for women reduces gender-based violence and consequently improves children's outcomes. The second chapter studies the causal impact of fertility on parental labor-force participation for the case of a developing country in the Balkans (Albania). Lastly, the third chapter uses the 1998-1999 Kosovo war and the following massive displacement of people as a natural experiment in order to estimate the impact of conflict displacement on adult's labor market outcomes and children's schooling outcomes of Kosovars who returned after being forcibly displacement relative to those who stayed.

Contents

Abstract	vi
Abstract	vi
Acknowledgments	xvi
Acknowledgments	xvi
Notice	xix
Notice	xix
Introduction	1
1 Access to Justice, Gender Violence and Children: Evidence from Women's Justice Centers in Peru	10
1.1 Introduction	12
1.2 Background	18
1.2.1 Gender-based violence in Peru	18
1.2.2 Women's Justice Centers Program	20
1.3 The Data	22
1.3.1 Individual and Household Level Data	22
1.3.2 School Level Data	25
1.3.3 District Level Data	26
1.3.4 Measuring Exposure to the WJC Centers	27
1.4 Empirical Strategy	28
1.4.1 Placement of WJC centers	28
1.4.2 Individual Level Specification	32
1.4.3 District Level Specification	33
1.4.4 School Level Specification	34
1.5 Results	36
1.5.1 Impact of WJC Centers on Gender-Based Violence	36
1.5.2 Impact of WJC Centers on Children's School Attendance	38
1.5.3 Impact of WJC Centers on School Enrollment	39
1.6 Discussion: Mechanisms	42

1.7	Robustness Checks	45
1.7.1	Assessing the Internal Validity of the Research Design	45
1.7.2	Accounting for the dynamic impact of WJC centers	48
1.8	Conclusion	50
Appendices		84
1.A	Supplementary Tables	85
2	Fertility and Parental Labor-Force Participation: New Evidence from a Developing Country in the Balkans	100
2.1	Introduction	102
2.2	Background	106
2.2.1	The Evolution of Total Fertility Rate in Albania	106
2.2.2	Labor-Force Participation in Albania	107
2.3	The Data	109
2.3.1	Descriptive Statistics	111
2.4	Empirical Strategy: Instrumental Variables	113
2.4.1	Son Preference in Albania and Siblings Sex-Composition	113
2.4.2	The Relevance Condition	114
2.4.3	The Exclusion Restriction	115
2.4.4	The Econometric Framework	119
2.5	Results	122
2.5.1	First-Stage	122
2.5.2	Main Results: OLS and Second-Stage	124
2.6	Discussion	126
2.6.1	Two plausible mechanisms	126
2.6.2	Heterogeneity Analysis	128
2.7	Conclusion	131
Appendices		154
2.A	Data sources and data construction	155
2.B	LATE without monotonicity	158
2.C	Theoretical Framework	160
2.C.1	Maximization Problem	161
2.D	Supplementary Tables	165
3	Returning Home After Conflict Displacement: Labor Supply and Schooling Outcomes Among Kosovar Households	173
3.1	Introduction	175
3.2	Background	180

3.2.1	Kosovo War (1998-1999)	180
3.2.2	Forced Displacement from Kosovo	181
3.3	The Data	182
3.3.1	Measuring Conflict Displacement	184
3.3.2	Descriptive Statistics	185
3.4	Empirical Strategy	187
3.4.1	Identification: Instrumental Variables Approach	188
3.4.2	First-Stage Estimation	191
3.4.3	Isolating Plausibly Exogenous Variation	194
3.4.4	Robustness Checks	195
3.5	Results	197
3.5.1	Conflict Displacement and Labor Market Outcomes	198
3.5.2	Conflict Displacement and School Enrollment Outcomes	200
3.6	Mechanisms	201
3.6.1	Channels on Labor Outcomes	201
3.6.2	Channels on Education Outcomes	203
3.7	Conclusion	204
Appendices		222
3.A	Supplementary Figures and Tables	223
3.B	Alternative Identification Strategy	239
3.B.1	Identification: Difference-in-Difference (DID) Strategy	239
3.B.2	Endogeneity in Municipality Displacement	240
3.B.3	Displacement and Schooling Completion Outcomes	241
Bibliography		248

List of Tables

1.1	Women's Descriptive Statistics and WJC Center Exposure - DHS (2006-2014) .	59
1.2	Children's Descriptive Statistics and WJC Center Exposure - DHS (2006-2014) .	60
1.3	School Descriptive Statistics and WJC Center Exposure - School Census (2006-2014)	61
1.4	Placement of WJC Centers in the District	62
1.5	The Effect of WJC centers on Domestic Violence - (2006-2014)	63
1.6	Impact of WJC centers on Emotional Violence - (2006-2014)	64
1.7	WJC centers and Femicides at the District Level - (2009-2015)	65
1.8	WJC centers and Female Deaths due to Aggression (registered cases in hospitals) at the District Level - (2007-2014)	65
1.9	The Effect of WJC centers on Women's Health Outcomes - (2006-2014)	66
1.10	WJC centers and Female Hospitalizations for Mental Health Problems at the District Level - (2006-2016)	67
1.11	The Effect of WJC Centers on Children's Primary School Attendance - (2006-2014)	68
1.12	The Effect of WJC Centers on Children's Secondary School Attendance - (2006-2014)	69
1.13	School Attendance Status and Proximity to a WJC center - (2006-2014)	70
1.14	The Effect of WJC Centers on Primary School Enrollment (2006-2014)	71
1.15	The Effect of WJC Centers on Secondary School Enrollment (2006-2014)	72
1.16	The Effect of WJC Centers on Primary Level 2nd Grade Test Scores - (2006-2014)	73
1.17	Impact of WJC centers on Decision Making and Bargaining Power in the Household - (2006-2014)	74
1.18	Impact of WJC centers on Women's Labor Force Participation - (2006-2014) . .	75
1.19	The Effect of WJC centers on Type of Institution Women Choose to Seek Help from Domestic Violence Events- (2006-2014)	76
1.20	The Effect of WJC Centers on Child Labor - (2006-2014)	77
1.21	CCT <i>Juntos</i> and WJC Centers on Children's Primary School Attendance - (2006-2014)	78
1.22	CCT <i>Juntos</i> and WJC Centers on Primary School Enrollment - (2006-2014) . . .	79
1.A.1	Correlation between WJC center and CCT <i>Juntos</i> program implementation (2005-2014)	86

1.A.2	School Enrollment Effects by Gender and Grade	87
1.A.3	School Enrollment and Children's School Attendance Status Effects by Gender	88
1.A.4	Domestic Violence Effects by Age, Education Level and Type of Domestic Violence - (2006-2014)	89
1.A.5	Domestic Violence by Children's Primary Level School Attendance Status - (2006-2014)	90
1.A.6	The Effect of WJC Centers on Primary Level Attendance Quintiles - (2006-2014)	90
1.A.7	The Effect of WJC Centers on Children's Primary School Attendance - (2006-2014) - Alternative Euclidean Buffers	91
1.A.8	The Effect of WJC Centers on Children's Secondary School Attendance - (2006-2014) - Alternative Euclidean Buffers	92
1.A.9	Heterogeneity by Violent Households	93
1.A.10	Placebo regressions, Impact of Future WJC Centers on Primary School Enrollment	94
1.A.11	Placebo regressions, Impact of Future WJC centers on Secondary School Enrollment	95
1.A.12	Relationship between WJCs within 1km rollout and pre-program school enrollment	96
1.A.13	Relationship between WJCs in the district rollout and pre-program school attendance	97
1.A.14	Relationship between WJCs within 1km rollout and four windows of pre-program standardized test scores (2nd grade - Primary School)	98
1.A.15	Relationship between WJCs in the district and four windows of pre-program domestic violence	99
2.1	Descriptive Statistics for Mother's Fertility in Albania (2002 - 2012)	135
2.2	Descriptive Statistics for Mother's Labor-Supply in Albania (2002 - 2012) . . .	136
2.3	Descriptive Statistics for Father's Labor-Supply in Albania (2002 - 2012)	137
2.4	Fraction of Households that had Another Child by Parity and Sex Composition in Albania (2002- 2012)	138
2.5	Differences in Means for Demographic Characteristics by Sex Composition . .	139
2.6	Differences in HH Budget Shares by Sex Child Composition in Albania (2002 - 2012)	140
2.7	First-Stage Specifications - Albania (2002 - 2012)	141
2.8	OLS and IV Estimates of Mother's Labor-Supply Models in Albania (2002 - 2012)	142
2.9	OLS and IV Estimates of Father's Labor-Supply Models in Albania (2002 - 2012)	143
2.10	OLS and IV Estimates of Parental Labor-Supply - Heterogeneity Analysis by Educational Attainment	144
2.11	OLS and IV Estimates of Parental Labor-Supply - Heterogeneity Analysis by Birth Cohort	145

2.12	OLS and IV Estimates of Parental Labor-Supply - Heterogeneity Analysis by Family Structure	146
2.13	Falsification Exercise: Determinants of Living in an Extended Family - Albania (2002 - 2012)	147
2.14	IV Estimates of the Effect of Fertility on Family Spending per-capita - Heterogeneity Analysis by Expenditure Quintiles and Family Structure - Albania (2002 - 2012)	148
2.D.1	Sample Sizes: LSMS and DHS - Albania (2002-2012)	166
2.D.2	Fraction of Households that had Another Child by Parity and Sex Composition in Albania (2002-2012) - Decomposed by Survey	167
2.D.3	Fraction of Households that had Another Child by Parity and Sex Composition in Albania (2002-2012) - Decomposed by Survey	168
2.D.4	First-Stage Specifications (<i>Albania 2002, 2005 and 2012 LSMS</i>)	169
2.D.5	OLS and IV Estimates of Mother's Labor-Supply Models in Albania (2002 - 2012) (<i>Mothers aged 20 to 49 years old with 2 or more children younger than 13 years old</i>)	170
2.D.6	Secular Additive Effects of Child Gender on Labor Supply - Albania (2002-2012)	171
2.D.7	Secular Additive Effects of Child Gender on Labor Supply - Albania (2002-2012)	172
3.1	Descriptive Statistics of Displacement by Gender and Age Group - Kosovo (1999-2000)	209
3.2	Descriptive Statistics of Displacement Status by Ethnic Group - Kosovo (1999-2000)	209
3.3	First-stage Results - Children Aged (6-19 Years Old) - Kosovo (1999-2000) . . .	210
3.4	First-stage Results - Adults (20-65 Years Old) - Kosovo (1999-2000)	211
3.5	Falsification Test on the First-stage Results - Adults Aged (20-65 Years Old) - Kosovo (1999-2000)	212
3.6	Pre-war Economic Performance and Conflict Incidence (Women) - Kosovo (1999-2000)	213
3.7	Pre-war Economic Performance and Conflict Incidence (Men) - Kosovo (1999-2000)	214
3.8	Pre-War Ethnicity and Conflict Incidence (Women) - Kosovo (1999-2000) . . .	215
3.9	Pre-War Ethnicity and Conflict Incidence (Men) - Kosovo (1999-2000)	216
3.10	The Effect of Conflict Displacement on Women's Labor Market Outcomes - Kosovo (1999-2000)	217
3.11	The Effect of Conflict Displacement on Men's Labor Market Outcomes - Kosovo (1999-2000)	218
3.12	The Effect of Conflict Displacement on Children's Enrollment - Kosovo (1999-2000)	219

3.13	Channels of Conflict Displacement on Labor Market Outcomes - Kosovo (1999-2000)	220
3.14	Channels of Conflict Displacement on Education Outcomes - (IV estimation) - Kosovo (1999)	221
3.A.1	Compliers - Instrument (WCR x Distance to the Albanian Border) -	232
3.A.2	Compliers - Instrument (Bombings x Distance to the Albanian Border) -	233
3.A.3	The Effect of Conflict Displacement on Children's Enrollment - (Wild Bootstrap Inference)	234
3.A.4	The Effect of Conflict Displacement on Women's Labor Market Outcomes - Kosovo (Wild Bootstrap Inference)	235
3.A.5	The Effect of Conflict Displacement on Men's Labor Market Outcomes - Kosovo (Wild Bootstrap Inference)	236
3.A.6	The Effect of Displacement on Labor Market Outcomes by Occupation Type	237
3.A.7	The Effect of Displacement on Labor Market Outcomes by Occupation Type (Only Work off-farm)	238
3.B.1	Identification of Displacement Affected Cohorts and Unaffected Cohorts by Schooling Completion Outcomes	243
3.B.2	First-stage Results for Primary School Completion Outcome - (Difference-in-Difference Estimation)	244
3.B.3	First-stage Results for Secondary School Completion Outcome - (Difference-in-Difference Estimation)	245
3.B.4	The Effect of Municipality Displacement on Female's Schooling Attainment - Kosovo (1999-2000)	246
3.B.5	The Effect of Municipality Displacement on Male's Schooling Attainment - Kosovo	247

List of Figures

1.1	Distribution and Growth of the Opening of the <i>women's justice centers</i> (WJCs) by Year - Peru (1999-2016)	53
1.2	Rollout of the WJCs across Time and Space (1999-2014)	54
1.3	Total Number of Persons Attended in WJC Centers by Year (2002-2016)	55
1.4	Domestic Violence in Peru (2003-2015)	55
1.5	Euclidean Distance Buffers and WJC centers (Schools and DHS Clusters of Households) - Lima and Tumbes	56
1.6	Femicides, Female Deaths due to Aggression and Female Hospitalizations for Mental Health Problems - Peru	57
1.7	WJC center and CCT <i>Juntos</i> presence in the district	58
1.8	WJC center and CCT <i>Juntos</i> entry in the district	58
1.9	Effect of WJC center rollout on changes in pre-program outcomes	80
1.10	Event Study: Pre-WJC and Post-WJC Trends in Women's Outcomes	81
1.11	Event Study: Pre-WJC and Post-WJC Trends in Children's School Attendance	82
1.12	Event Study: Pre-WJC and Post-WJC Trends in School Enrollment and Test Scores	83
2.1	Evolution of the Total Fertility Rate in Albania (1960 - 2012)	149
2.2	Labor-Force Participation and Total Fertility Rate - Albania (1990 - 2012)	150
2.3	Pre-school in Albania (1990 - 2012)	151
2.4	Graphical Version of Fraction of Households that had Another Child by Parity and Sex Composition - Decomposed by Survey	152
2.5	Proportion of parents in extended families living with their married son by their son's birth order	153
2.B.1	LATE without monotonicity - Compliers-Defiers (CD) condition	159
3.1	Total Cumulative Kosovar Refugee, Returned Kosovar Refugee, Serbian Refugee Populations and Survey Timelines, 23 (March-October 1999)	207
3.2	War Casualty Rate- Number of Casualties per 1,000 Inhabitants across Municipalities	208
3.3	Bombing Intensity- NATO bombing Days across Municipalities	208

3.A.1	Displaced populations from Kosovo in neighbouring countries/territories, mid-June 1999	224
3.A.2	Total Cumulative refugee population in Montenegro, Albania, FYR Macedonia, Bosnia and Herzegovina and HEP, (March-October 1999)	225
3.A.3	Total Cumulative Albanian Kosovar Refugee, Returned Albanian Kosovar Refugee and Serbian Refugee Populations (March-October 1999)	225
3.A.4	Daily Returned Refugees (June-October 1999) - UNHCR Estimates	226
3.A.5	Month of First and Last Displacement - 1999 DSHS	226
3.A.6	Proportion of Displaced Individuals at the Municipality Level - 1999 Kosovo DSHS	227
3.A.7	Proportion of Displaced Individuals at the Municipality Level - 2000 Kosovo LSMS	227
3.A.8	Ethnic Majorities across Municipalities in 1991	228
3.A.9	Labor Market Activity Status by Gender and Age Group - 1999 DSHS and 2000 LSMS	229
3.A.10	Enrollment Rates by Gender and Age - 1999 DSHS and 2000 LSMS	230
3.A.11	Estimated Daily Refugee Displacement and Casualties (March - June 1999) . .	231
3.A.12	Estimated Daily Refugee Displacement and NATO Bombings (March - June 1999)	231

Acknowledgments

This PhD thesis is the product of nearly four years of research at Paris School of Economics. I would not have been able to complete this dissertation without the help and support of many people.

Foremost, I am grateful to my advisor Karen Macours for her continuous support and guidance, for sharing her knowledge and experience, for always finding time to meet, for encouraging my research ideas and, most importantly, for projecting tranquility and ensuring that I always remained calm along this PhD experience. Her encouragement for exploring the relevant, fundamental questions in development, her creativity when searching for novel answers, and her insightful critique have been very valuable and helped me to stay curious and rigorous.

I would like to thank Oliver Vanden Eynde for following my work from the very beginning, giving me extremely insightful feedback and comments, agreeing to be a jury member, taking part in the pre-defense and writing a very detailed report which has been very useful for my progress.

I would like to extend my profound gratitude to Gustavo Bobonis, John Giles and Dominique Van De Walle for agreeing to be external members of the jury. I greatly appreciate their valuable time and effort, and I am sure that their suggestions and comments have greatly improved my work.

I am also indebted to my co-authors Maria Micaela Sviatschi and Guadalupe Kavanaugh for having accepted to investigate with me the inter-generational benefits of improving access to justice for women using evidence from Peru. I have learned a lot working with them and I have enjoyed very much our debates on how to improve our paper.

While the list of people that has contributed to the chapters of this thesis with their comments is long, I would like to acknowledge in particular the faculty at Paris School of Economics who have helped me during this PhD with many insightful discussions, in particular Ekaterina Zhuravskaya, Sylvie Lambert, Marc Gurgand, Andrew Clark, Akiko Suwa-Eisenmann, François Libois, Jeremie Gignoux and Denis Cogneau. I am also grateful to the many researchers in various institutions with whom I have had fruitful and helpful research conversations: Pascaline Dupas, Suresh Naidu, Pamela Jakiela, Cristian Pop-Eleches, Christian

Fons-Rosen, Martin Ravallion, Dominique Van De Walle and Marta Rubio Codina.

In addition, I appreciate the help of the administrative staff at Paris School of Economics, notably Weronika Leduc, Eliane Madelaine and Véronique Guillotin, and their guidance in surviving French bureaucracy. I also appreciate the help of the IT team, in particular Radja Aroquiaradja and Stéphane Brice.

Moreover, I would like to acknowledge the professors at Universidad de San Andrés, especially Martin Rossi and Walter Sosa-Escudero who first guided me into research several years ago, who trusted in my potential and who gave me all the support I needed. In particular, Martin took me under his wing as a research assistant at Universidad de San Andrés and he introduced me to the fascinating world of research in economics. His support was fundamental for me to carry out my graduate studies in Paris.

My time at Paris School of Economics would not have been as enjoyable without my fellow PhD colleagues. I will miss the many coffee and lunch breaks, the PhD drinks and the different debates, especially with Luis, Oscar, Hector, Sebastian, Roberto, Emma, Sulin, Vanda, Anne, Quhyn, Lourdes, Pepe, Clara, Marc, Juni, Cristina, Jaime x2, Caroline, Simon, Anthony, Sacha and Alessandro. Some of them have become very good friends and I am grateful for their support. I am especially thankful to my friends Luis, Oscar, Sebastian, Roberto and Hector for their unconditional friendship and for making my years at PSE more fun and less serious. Without our adventures, our long talks, our travels and our parties, these years in Paris would not have been the same.

In Paris, I also found a very special friend – Ivan. Even though we knew each other since the Master in Argentina, our paths crossed again when we started the PhD together in France. Our travels around the world, our slow walks discovering Paris, our long phone calls and whatsapps and our love for ice-cream have made us true friends. I am especially thankful to my dear friend Ivan for all his guidance, support and for everything he has done for me, particularly in my more difficult periods, and for our time together in Buenos Aires, Lisbon, Paris, Albania and Greece.

I also have to recognise the support of all my Argentinean friends who have visited me in Paris during these years and have made me feel as I had never left Buenos Aires. I am especially thankful to my dear friends Laura, Flor, Mica, Mili, Jose, Meli, Sofi, Vir and Lausa with whom I share many years of friendship and I consider as the sisters I never had.

Most importantly, I would like to thank my partner, Bruno, for his constant support and belief in me, for always giving me the necessary strength to look forward and hope for the better but, especially for always trying to find a way to make me happy. None of this would have been possible without his love and support during these years. I also want to thank Bruno's family for all their help and support during these years and for offering me many moments of joy, tranquility and great food in Montpellier.

Finally, I am most grateful to my family. My brother, Xhoni, thank you for always being there and making me laugh. But above all I thank my parents, Edmond and Iris, who always taught me the value of education and provided me unconditional support to be able to pursue my dreams. This dissertation is dedicated to them.

Paris, January 2018

Iva Trako

Notice

The three chapters of this dissertation are self-contained research articles, therefore the terms “paper” or “article” are used, and parts of the content may be repeated. Chapters 1 is co-authored, which explains the use of the “we” pronoun. The Tables and Figures are presented at the end of each Chapter.

Introduction

Development economics is a fascinating field because it attempts to explore the economic challenges peculiar to low income countries in the world and the contrasting experience of success and failure in the economies of different regions of the world. This field has developed theories and methods that inform policies and practices promoting economic growth and welfare for populations in developing and emerging economies. Usually, development economics touches so many different topics: income growth, welfare economics (including the study of poverty and inequality), agricultural economics, economic demography, labor economics (education, health, conditions in the work-place etc), the study of markets for goods, services, inputs, outputs, credit and insurance, public economics, institutions, corruption, conflict, natural resources and the environment etc. What is economic development *not* about, one might ask?

Economists, especially modern development economists have long been considered to invade other social sciences and throughout this path they have accepted the challenge to understand matters that lay beyond their original realm. The general consensus today among most development economists is that economic development is not just equivalent to growth. Much of the focus in this field has shifted to the micro-level and to project evaluation. Basically, development economics research involves using experimental and quasi-experimental methods to learn about people's economic behavior and evaluate the impact of policy intervention on welfare outcomes.

Following the fashion of economics PhD theses, this dissertation is a collection of essays that try to make a marginal contribution to the vast literature in development economics. Each essay constitutes a chapter of this dissertation and is concerned with a particular topic in development economics. These chapters are stand-alone chapters, in the sense that each has its own introduction, method, datasets, results and discussion. The choice of countries in each

chapter is driven by a combination of data availability and suitability for the research question. Finding and collecting reliable data with a wide coverage for a developing country can be a difficult challenge. Although the chapters of this dissertation cover a wide variety of topics and use data from different developing countries, they all share a common ground in that their research questions pertain to development economics. In this introduction, I will begin by reviewing the literature that is relevant for each chapter of this thesis, followed by a brief summary of the different chapters.

Chapter 1, “**Access to Justice, Gender Violence and Children: Evidence from Women’s Justice Centers in Peru**”, evaluates the impact of a policy intervention in Peru aimed at improving access to justice and reducing violence against women. There are two main bodies of literature on gender-based violence. The first one focuses on the risk factors for gender-based violence, while the second one focuses on the effects of gender-based violence on women’s outcomes, including those for children living in households with domestic violence. While most fundamental studies on the causes and effects of gender-based violence center in developed countries, especially the U.S., a new wave of literature has expanded the scope of study to developing countries due to this form of violence’s perceived obstacle to the broader development agenda.

There is a growing literature on causal channels that impinge on the prevalence of gender-based violence. One type of this literature focuses on the intra-household bargaining channels that affect domestic violence through improvements in women’s outside options. For example, [Aizer \(2011\)](#) shows that a decline in the gender wage gap reduces violence against women in California. The author’s interpretation is that a relative improvement in female income reduces her exposure to spousal violence by increasing her bargaining power. [Stevenson and Wolfers \(2006\)](#) find that the adoption of unilateral divorce laws in the United States resulted into a drop in female homicide and domestic violence. In a more recent study, [Brassiolo \(2016\)](#) finds a decline in spousal conflict and in extreme partner violence in response to introducing less stringent divorce legislation in Spain. Using victimization data from the US, [Miller and Segal \(2016\)](#) find that as female representation increases among police officers in an area, violent crimes against women in that area, and especially domestic violence, are reported to the police at significantly higher rates. They also show that increases in female officer representation are followed by significant declines in intimate partner homicide rates and in rates of repeated

domestic abuse.

Another strand of this literature focuses on the heterogeneous effects of conditional cash transfer programs on gender-based violence. [Bobonis, González-Brenes and Castro \(2013\)](#) analyze the effect of the Mexican program *Oportunidades* on domestic violence and find that beneficiary women are less likely to be victims of physical violence but are more likely to receive threats of violence. Using the same randomized evaluation, [Angelucci \(2008\)](#) finds that among households that received small transfers, alcohol-related gender-based violence declined, whereas in households that received large transfers, the level of spousal abuse from husbands with particularly low levels of education increased.

In addition, recent research has observed that in many contexts, increased autonomy and women's entry into the formal labor market is often associated with a higher likelihood of experiencing violence in Colombia ([Friedemann-Sánchez and Lovatón, 2012](#)), Bangladesh ([Heath, 2012](#); [Rahman, Hoque and Makinoda, 2011](#)) and India ([Eswaran and Malhotra, 2011](#)). Indeed, the paid employment or the non-labor income of a female intimate partner may be threatening for some men, especially those who are unemployed. Abusive partners may perceive a loss of status and power and use violence or coercion to regain control.

However, less literature has been written on the consequences of gender-based violence on children's outcomes, especially in developing countries. Previous research has shown that children exposed to domestic violence are associated with a number of health, emotional and behavioural problems including, low birthweight, aggressive behaviour, bullying, depression, violence in adulthood and also diminishing academic performance.¹ With respect to children's education outcomes, studies conducted in the United States have found lower reading levels among teenagers who have been exposed to domestic violence ([Thompson and Whimper, 2010](#)), lower academic achievement in math and reading for children in elementary and middle school ([Kiesel, Piescher and Edleson, 2011](#)), lower scores on standardized tests for children ages 6 to 17 - especially for girls and children younger than 12 years old ([Peek-Asa et al., 2007](#)) - and more grade repetition and truancy among children 6 to 15 years old ([Emery, 2011](#)). Moreover, [Carrell and Hoekstra \(2010\)](#) show that exposure to school peers from troubled families significantly decreases reading and math test scores and increases misbehaviour in the classroom.

Among the scattering studies conducted in developing countries, [Jayasinghe, Jayawardena](#)

¹See [Edleson \(1999\)](#); [Wolfe et al. \(2003\)](#); [Pollak \(2004\)](#); [Fantuzzo et al. \(1997\)](#); [Koenen et al. \(2003\)](#); [Holt, Buckley and Whelan \(2008\)](#); [Baldry \(2003\)](#); [Carlson \(2000\)](#); [Currie \(2006\)](#); [Black, Sussman and Unger \(2010\)](#); [Aizer \(2011\)](#).

and Perera (2009) show that children who were directly or indirectly exposed to domestic violence at home had poor school attendance and lower academic achievement on average. Similarly, Durand et al. (2011) find that Brazilian children 5 to 12 years old who lived with mothers exposed to psychological, physical and sexual domestic violence were more likely to be among those dropping out of school or failing a school year.

What is perhaps most striking about this literature is that rigorous studies attempting to evaluate the effectiveness of various intervention strategies aimed at curbing gender-based violence are quite scarce. This is mainly due to the difficulties and ethical considerations on collecting reliable data on gender-based violence. Another difficulty is dealing with the endogeneity problem. Randomized experiments, for instance, are extremely rare. In addition, even though WJC centers are one such intervention that has been gaining popularity, little attention has been paid on to the actual effectiveness of such centers on eradicating violence against women and, particularly, there is very little evidence on the extent of spillovers on their children. Two exceptions are the studies of Agüero (2013) and Perova and Reynolds (2017), which exploit the variation stemming from the gradual municipality/district -level rollout of the WPS/WJC centers in Peru and Brazil, respectively. We complement these papers by providing causal estimates at a more disaggregated level and on a wider number of outcomes that allow us to disentangle mechanisms and study spillover effects on children.

Many developing countries have unequal access to justice, especially for women. What are the implications for gender-based violence, intra-household bargaining and investments in children? Chapter 1 provides quasi-experimental evidence on women's justice centers (WJCs) a community based approach aimed at reducing violence against women in Peru. Using administrative data from health providers and district attorney offices, we exploit the gradual rollout of these centers across all over Peru and find a reduction in domestic violence, female deaths due to aggression, femicides and mental health problems. Moreover, we find that the WJCs substantially increase human capital investments in children, increasing enrollment, attendance, test scores, while decreasing child labor. These results are consistent with a bargaining model in which the threat point is determined by access to justice. In sum, the evidence in this paper implies that providing access to justice for women is not only important for addressing gender-based violence, but also generates intergenerational benefits. The research to date has outlined much of the gender-based violence problem and

provided some fundamental understanding of its causes and consequences, but has left policy makers with little on which to build effective interventions. In this light, our paper contributes to the literature on domestic violence by focusing on an unexplored empowerment channel for women which is better access to justice and the role of women's justice centers (WJC) in breaking the cycle of violence and generating a spillover effect on their children's outcomes.

Chapter 2, "**Fertility and Parental Labor-Force Participation: New Evidence from a Developing Country in the Balkans**", examines fertility and labor supply decisions of Albanian households with at least two children. There is an extensive empirical literature attempting to explain the effect of fertility on labor supply. While earlier studies considered fertility to be an exogenous determinant of female labor supply, recent papers have recognized the endogeneity problem by aiming to establish a causal relationship between fertility and labor supply.² To address this endogeneity problem, economists have used different natural experiments to exploit exogenous variation in family size: twinning at first birth, siblings sex-composition and fertility shocks (Clarke, 2017).

There is considerable empirical evidence from developed countries, especially the U.S, supporting a negative impact of fertility on female labor-force participation, though not universally so. Examples include Rosenzweig and Wolpin (1980); Bronars and Grogger (1994); Jacobsen, Pearce and Rosenbloom (2001), who use twinning at first birth as a source of exogenous variation in order to estimate the effect of having a second child. Angrist and Evans (1998) exploit parental preferences for mixed-sex siblings in order to estimate the effect of a third or higher order child. They find that in the U.S. fertility reduces female labour supply significantly but that there is no significant change in male labor supply. Chun and Oh (2002) argue that while U.S. households prefer balancing the sex composition of their children, Korean families prefer sons. They use the first child's sex as an instrument for fertility and find

²Some of the earlier papers that have tried to establish a relationship between fertility and female labor supply can be categorized in several groups according to how the authors have tried to tackle the problem of endogeneity. The first group is illustrated by the studies of Gronau (1973); Heckman (1974) and Heckman and Willis (1977) who assume that fertility is exogenous and established a strong negative correlation between female labor supply and fertility. A second group of studies (Cain and Dooley, 1976; Schultz, 1978; Fleisher and Rhodes, 1979) acknowledge the endogeneity of the fertility decision and tried to deal with the problem by estimating simultaneous equation models. These studies find a much smaller estimate when treating fertility as an endogenous variable. The biggest challenge of this approach is that it is quite difficult to find plausible exclusion restrictions that could identify the underlying structural parameters. A third group of studies incorporates actual fertility as a regressor but adds the lagged dependent variable (i.e. labor supply) to control for unobserved heterogeneity across women. Nakamura and Nakamura (1992) recommended this approach, and it has been used by a number of authors (Even, 1987; Lehrer, 1992). Although adding the lagged dependent variable can help control the unobserved heterogeneity, it still does not address accurately the problem of the endogeneity of the fertility decision.

that having children reduces the labor force participation of married Korean women. More recent papers use different instruments that rely on childless mothers undergoing infertility treatments ([Cristia, 2008](#); [Lundborg et al., 2017](#)) or miscarriage in a women's first pregnancy ([Hotz, McElroy and Sanders, 2005](#)) or natural experiments like the use of the contraceptive pill at state level ([Bailey, 2013](#)) or changes in abortion legislation ([Bloom et al., 2009](#); [Angrist and Evans, 1996](#)) similarly conclude that fertility has a negative impact on mother's labor force-participation or earnings.

The empirical literature from developing countries is relatively small and the evidence found is not unified. Some studies in developing countries have also found negative or no impact. [Cruces and Galiani \(2007\)](#) generalize the results for the U.S. found by [Angrist and Evans \(1998\)](#) to the populations of two Latin American countries (Argentina and Mexico) and find a negative effect of fertility on female labor force participation. Using data from a social experiment in Bangladesh, [Schultz \(2009\)](#) also finds a negative effect of family planning programs on female labor-force participation. Using son-preference as an instrument, [Lee \(2002\)](#) finds no significant effect of fertility on rural female labor supply in China. [Ebenstein \(2009\)](#) also uses son-preference but he reports a negative effect of fertility on maternal labor force participation in Taiwan. A more recent study by [Agüero and Marks \(2011\)](#) uses self-reported infertility as an instrument for family size. Based on data from 26 low- and middle-income countries, their estimation results show that the presence of children affects neither the likelihood of work nor its intensity for women.

However, there is also a scattering of papers in developing countries which show mixed or even positive effects of fertility on labor supply. Based on data from Demographic and Health Surveys covering 59 developing countries, [Porter and King \(2012\)](#) report that while many women in developing countries are less likely to work when they have more children, some are more likely to work due to greater financial costs of feeding more children. [Bloom et al. \(2009\)](#) combine data from Demographic and Health Surveys with abortion legislation laws in each country, and also find mixed results of fertility. There is also a micro study from [Priebe \(2010\)](#), who uses data from Indonesia and finds a positive effect of fertility on female labor supply, which is mostly driven by women which are poorer, less educated and who live in rural areas.

All these studies report LATEs that hold consistent results for alternative populations

and compliers, and so do not provide external validity for inference in other populations. However, a more recent literature suggests that the negative effect could be observed in other contexts. External validity of the labor supply-fertility local average treatment effect is extensively examined by [Dehejia, Pop-Eleches and Samii \(2015\)](#) and [Bisbee et al. \(2015\)](#), who find that quasi-experimental evidence generalizes more readily to countries which share closer geographical, education, time and labor-force participation characteristics. There is also a novel strand of literature which documents how the impact of childbearing on female labor-force participation has evolved over time (the last two centuries) and across countries. Using twin births and siblings sex composition, [Aaronson et al. \(2017\)](#) find that the effect of fertility on labor supply is large and negative at high levels of income, but small and often indistinguishable from zero at low levels of income. These effects are consistent both across time looking at the historical time series of currently developed countries and at a contemporary cross section of developing countries.

Chapter 2 examines the effect of fertility on parental labor-force participation in a developing country in the Balkans, with particular attention to the intervening role of childcare provided by grandparents in extended families. In order to address the potential endogeneity in the fertility decision, I exploit Albanian parental preference for having sons combined with the siblings sex-composition instrument as an exogenous source of variation. Using a repeated cross-section of parents with at least two children, I find a positive and statistically significant effect of fertility on parental labor supply for those parents who are more likely to be younger, less educated or live in extended families. In particular, IV estimates for mothers show that they increase labor supply, especially in terms of hours worked per week and the likelihood of working off-farm. Similarly, father's likelihood of working off-farm and having a second occupation increase as a consequence of further childbearing. The heterogeneity analysis suggests that this positive effect might be the result of two plausible mechanisms: childcare provided by non-parental adults in extended families and greater financial costs of maintaining more children.

Although most recent research suggests a negative effect of fertility on female labor supply, the contribution of this paper is that the effect of fertility on parental labor supply can be positive in the context of a developing country, similar to [Porter and King \(2012\)](#); [Bloom et al. \(2009\)](#) and [Priebe \(2010\)](#). Moreover, this paper presents a comparative analysis on

the relationship between fertility and labor supply based on the type of instrument. By decomposing the same-sex instrument, the results found in the paper suggest that in a context characterized by a strong son preference, all the relevant effect is coming from the *two girls* siblings sex composition. This finding might provide external validity implications for other similar settings.

Chapter 3, “**Returning Home After Conflict Displacement: Labor Supply and Schooling Outcomes Among Kosovar Households**”, analyzes the effect of forced displacement on adult’s labor market outcomes and children’s schooling in the context of the post-war Kosovo. Conflict displacement has surfaced as one of the most pressing humanitarian and development issues of our days. There are not only more displaced individuals around the world compared to any time since World War II, but also its rapid acceleration is disconcerting ([UNHCR, 2017](#)). There is already an extensive economic literature on the impacts of voluntary migration and the impacts of war and violence, but the literature on the economics of forced displacement is still in its early stages ([Ruiz and Vargas-Silva, 2013](#)).³ Some of the main reasons for this limited number of studies are lack of reliable data on conflict areas and methodological difficulties in separating the impact of war and violence from the impact of forced migration.

Despite this, the literature on the economics of forced migration is starting to gain attention in the last years as the micro data sets on conflict areas are becoming more available. Some of the seminal studies in this field include [Cortes \(2004\)](#), who compares refugees to economic immigrants in the United States, [Ibáñez and Moya \(2006\)](#), who make a simple comparison of asset loss and consumption changes between displaced and non-displaced Colombians, and [Ssewanyana, Younger and Kasirye \(2007\)](#), who attempt to estimate the effect of displacement on consumption levels in northern Uganda. However, these studies fail to account for the potential endogeneity problem by providing credible counterfactuals.

There are also a few examples of quantitative estimates of the effect of displacement and the consequences seem to be mixed. For instance, [Sarvimäki, Uusitalo and Jantti \(2009\)](#) find increased mobility among displaced Finns due to WWII and consequently higher long-run incomes. Nevertheless, most of the previous literature suggests that there are serious

³[Ruiz and Vargas-Silva \(2013\)](#) summarize some of the literature on the economics of forced migration for both developed and developing countries and they divide it into two categories: (i) the impacts of forced migration on the forced migrants and (ii) the impacts of forced migration on the host communities. In this dissertation, I am going to focus on the first group since Chapter 3 is particularly relevant for this body of literature.

negative consequences of forced displacement for those forced to migrate. [Fiala \(2015\)](#) finds a sizeable reduction in consumption smoothing for displaced households in Uganda. [Eder \(2014\)](#) analyzing post-war Bosnia, shows that displaced individuals invest less on their children's education. [Kondylis \(2010\)](#) also using data from post-war Bosnia, finds higher unemployment for men and lower labor force participation for women. [Bauer, Braun and Kvasnicka \(2013\)](#), analyzing the integration of Germans from Easter Europe, conclude that the first generation of migrants has lower incomes and ownership rates. [Abdel-Rahim, Jaimovich and Ylönen \(2015\)](#), studying displacement in Nuba Mountains of Sudan, conclude that displaced households hold fewer assets and are less involved in production. [Verwimp and Muñoz-Mora \(2018\)](#) investigate the food security and nutritional status of formerly displaced households in Burundi and they find that individuals who remain much longer in a displacement status are worse off compared to those who returned earlier.

Chapter 3 uses the 1998-1999 Kosovo war and the following massive displacement of people as a natural experiment in order to estimate the impact of conflict displacement on labor market and education outcomes of Kosovars after they returned from exile. I exploit the interaction of the spatial variation in conflict intensity -as measured by casualties and bombings- and distance to the Albanian border as a source of exogenous variation in the displacement status. Results indicate that displaced Kosovar men are less likely to be employed in the agricultural sector and to work on their own account, while displaced Kosovar women are more likely to be inactive. Loss of assets (e.g. land, livestock) in an agrarian skill-based economy and also loss of social networks in an informal labor market might have further decreased the probability to find employment relative to stayers. However, shortly after the return home, the results also indicate that displaced Kosovar men and women are more likely to be working off-farm, especially in the construction and public administration sectors, which indicates a relatively quick recovery. In addition, displaced Kosovar girls are more likely to be enrolled in primary school, but I find no effect on education for boys. The refugee camp experience might have provided better conditions to young Kosovar girls compared to the the precarious pre-war "parallel" education system.

This paper contributes to the growing literature on the economics of forced displacement at the microeconomic level by accounting for potential selection issues through a novel interaction-based instrument involving conflict intensity and distance. Moreover, Kosovo constitutes an

interesting case study for this analysis as it is one of only a small number of countries for which detailed conflict intensity and conflict displacement information is available for the immediate period after the conflict.

Chapter 1

Access to Justice, Gender Violence and Children: Evidence from Women's Justice Centers in Peru¹

¹This Chapter is based on a joint work with Maria Micaela Sviatschi (Princeton University) and Guadalupe Kavanaugh (Rutgers University). We are very grateful for the abundance of support, guidance, and helpful comments from Karen Macours. We also gratefully acknowledge all the helpful comments and suggestions from Oliver Vanden Eynde, Gustavo Bobonis, John Giles, Martin Ravallion, Dominique van de Walle, Denis Cogneau, Sylvie Lambert, Zach Brown, Suresh Naidu, Cristian Pop-Eleches, Miguel Urquiola, Bentley MacLeod, Jonas Hjort, Fabiola Alba and Maria Fernanda Rosales and all the participants of the Development Seminar at Paris School of Economics. We are responsible for all remaining errors.

Abstract

Many developing countries have unequal access to justice, especially for women. What are the implications for gender-based violence, intra-household bargaining and investments in children? This paper provides quasi-experimental evidence on women's justice centers (WJCs) a community based approach aimed at reducing violence against women in Peru. Using administrative data from health providers and district attorney offices, we examine the gradual rollout of these centers and find a reduction in domestic violence, female deaths due to aggression, femicides and mental health problems. Moreover, we find that the WJCs substantially increase human capital investments in children, increasing enrollment, attendance, test scores, while decreasing child labor. These results are consistent with a bargaining model in which the threat point is determined by access to justice. In sum, the evidence in this paper implies that providing access to justice for women is not only important for addressing gender-based violence, but it also generates inter-generational benefits.

JEL Classification: J12, J16, I25, K38

Keywords: gender-based violence, access to justice, education, household bargaining.

1.1 Introduction

An accessible and fair justice system is thought to be important for economic development, so much so that it was recently added as a United Nations Sustainable Development Goal. However, in developing countries, the high rates of under-reporting for gender-based violence and low convictions rates for crimes against women implies unequal access to the law, particularly for women. In particular, women are often unable to seek justice for domestic violence or receive equitable treatment during a divorce (Duflo, 2012; Revilla, 1999). Evidence from India finds that only 3% of women have ever had contact with the police, despite the fact that violence is quite high (Banerjee et al., 2012). This may have to do with the fact that women do not trust formal institutions enough to report violence. In some countries, the police regularly ignore domestic violence complaints, saying that “domestic disputes” are not a police matter (e.g. Jubb et al., 2010; Boesten, 2012).²

At the same time, gender-based violence remains a worldwide social problem, affecting 30% of women each year (WHO, 2013).³ Research and policy in this area has mainly focused on addressing these issues through economically empowering women, but in some cases this can actually increase gender-based violence (e.g. Bobonis, González-Brenes and Castro, 2013; Angelucci, 2008; Aizer, 2010). There is little evidence on the role of state capacity to deter gender-based violence by improving access to justice. This is of critical importance given the paradoxical finding that better economic conditions can lead to a backlash effects, increasing conflict over household resources. A potential, and previously unconsidered, explanation for this is the fact that deterrence of crimes committed against women may be perceived to be low, allowing perpetrators to threaten violence without repercussions. In other words, when civic participation and access to justice for women are low, women cannot rely on the justice system as a credible threat to prevent violence.

In this paper, we examine whether improving access to justice for women reduces gender-

²Furthermore, in cases of family violence in rural Peruvian communities, women are often assumed to have a certain level of blame in a conflict (Revilla, 1999). Traditional methods of justice based on local customs also are often discriminatory towards women and rarely impartial (Franco and González, 2009).

³Women who suffer from abusive in the household are more likely to report physical, mental, sexual and reproductive health problems (Campbell, 2002). Domestic violence may also limit their ability to take care of children. A growing literature on domestic violence finds that childhood exposure to domestic violence is associated with a number of emotional and behavioral problems (e.g. Pollak, 2004; Carlson, 2000; Huth-Bocks, Levendosky and Semel, 2001; Koenen et al., 2003; Carrell and Hoekstra, 2010).

based violence and consequently improves children's outcomes in Peru. First, we study how women's access to justice affects the incidence of gender-based violence as measured by self-reported domestic violence. We also use administrative data from hospitals and district attorneys to examine the effect on femicides, female deaths due to aggression, and mental health. Second, we examine the intergenerational effects, focusing on investments in children's human capital. This provides insight into whether household investments in children become more aligned with women's preferences when violence against women declines. In particular, whether these effects are consistent with a bargaining model in which the threat point is determined by access to justice.

To isolate causal channels, we exploit the gradual rollout of *women's justice centers* (WJC) across Peru during the period 2006-2014. The WJC centers are specialized institutions whose main purpose is to reduce gender-based violence by providing legal and psychological support. Violence against women is particularly important in Peru where the rate of intimate partner violence has been among the highest in the world, consistent with the fact that women in Peru have little bargaining power (e.g. [Garcia-Moreno et al., 2006](#); [Bott et al., 2012](#)). As a response to this endemic problem in the country, the *Peruvian Ministry for Women and Vulnerable Populations* decided to create the WJC in 1999 as part of the National Program against Sexual and Family Violence. During the period of analysis, the number of WJC centers has grown from 13 in the first year to 226 by the end of 2014, covering 100% of the 24 regions of Peru and 96% of the provinces. Even though WJC centers are one such intervention that has been gaining popularity in developing countries, little attention has been paid to the actual effectiveness of such centers.

Given this setting, we use a difference-in-differences strategy which exploits variation created by the differential timing in the opening of the WJC centers and spatial variation in the exposure of a school/household to a WJC center, together with province-by-year fixed effects. We geo-match schools and households with detailed data on WJC's locations and founding years in order to construct two different measures of exposure to the WJC center: presence of WJC center within 1 kilometer from the household/school and presence of a WJC center in household's/school's district. This empirical strategy allows us to compare changes in outcomes of households, including women and their children, and schools already residing in the proximity of a WJC center ("treatment households/schools") to those not yet reached by the program ("control households/schools").

To observe these sources of variation, we build a detailed panel using multiple geo-coded datasets: individual and household-level data, school level data and administrative data on WJC centers, femicides, female mortality due to aggression and female hospitalizations for mental health problems which allow us to analyze the effects at a very disaggregated level. First, our individual and household-level data comes from the *Peruvian Demographic Health Survey (DHS)*, which is a nationally representative cross-sectional survey that contains rich information on demographic and socioeconomic characteristics of the household members, as well as a detailed domestic violence module for married or cohabiting women. The *Peruvian DHS* covers the period 2000-2014 and is geo-coded at the cluster level. Second, our school level data comes from the *Peruvian School Census*, which is a large geo-coded panel dataset on primary and secondary school enrollment that covers the universe of schools in Peru during the period 1998 to 2014. Third, the administrative data on WJC centers comes from the *Peruvian Ministry for Women and Vulnerable Populations (MIMP)* and consists of a geo-coded directory of WJC centers and their founding dates across all over Peru from 1999 till 2014. Data on femicides at the district level comes from the *Peruvian Crime Observatory at the Ministry of Public Affairs*. Finally, data on female deaths due to aggression and female hospitalizations for mental health problems at the district level come from hospital records obtained from the *Peruvian Ministry of Health*.

Our first finding is that improving access to justice for women reduces self-reported domestic violence, femicides, female deaths due to aggression and improves mental and physical health. In particular, we find that after the opening women who reside in the proximity of a WJC center are significantly less likely to suffer from physical and emotional violence by their spouse. At the same time, the presence of a WJC center in the district is associated with 2-7% reduction in the number of femicides and female deaths due to aggression. Moreover, we find evidence that mental health hospitalizations decline by 20%. Finally, we also find some evidence that after the WJC opened, women are more likely to utilize formal institutions in cases of violence, suggesting an increase in trust in state institution. These results may also lead to an improvement in women's intra-household bargaining power.

We show that the impact of WJCs is not limited to the direct recipients of gender-based violence. We show evidence that after the opening of a WJC, children in primary and secondary school living in household's located near a WJC center are significantly more likely to attend

school, to pass a grade and less likely to drop out of school. More specifically, the probability that children reached by the WJC centers attended school and passed a grade increases by approximately 2 percentage points, while drop-outs decrease by 1.8 percentage points. These results are robust to using different datasets measuring schooling outcomes. Using the school panel data, we find that the introduction of a WJC center within 1km of a school is associated with a 2% to 3% increase in the total number of children enrolled in primary schools in the year after the center introduction. Finally, schools located near a WJC center have better national test scores. Consistent with the education results, we also find that young girls are less likely to be working after the opening of the WJC centers.⁴

Furthermore, to better understand why empowering women would help promote children's education in the context of Peru, we analyze the impact of WJC center by households where the grandmother was subject to violence. We find evidence that most of the results for children are driven by these households, suggesting that when justice increases, there is a lower credible threat of violence, increasing the ability of women to guide household resources towards their preferred investments.

The next focus of this paper is to examine the mechanisms driving these results. Access to justice may allow women to credibly threaten to involve police or decrease the incentive for offenders to use violence given the higher probability of criminal penalties. Several economic theories of household bargaining power suggest that policies aimed at increasing women's outside options when in an abusive relationship may also affect within-household distribution through changes in their relative bargaining positions (Farmer and Tiefenthaler, 1996; McElroy and Horney, 1981; Manser and Brown, 1980). For instance, economic empowerment of women is often considered a major tool in the fight against intimate partner violence.⁵ Similarly,

⁴These results are in line with previous research in developed countries showing that children exposed to domestic violence are associated with a number of health, emotional and behavioural problems and also diminishing academic performance. See Edleson (1999); Wolfe et al. (2003); Pollak (2004); Fantuzzo et al. (1997); Koenen et al. (2003); Holt, Buckley and Whelan (2008); Baldry (2003); Carlson (2000); Currie (2006); Black, Sussman and Unger (2010); Aizer (2011). For evidence on the role of domestic violence on children's health outcomes in Latin America see Agüero (2013).

⁵On the one hand, employment opportunities such as conditional cash transfers or access to welfare services may empower women by increasing their resources within the household; improve their outside options and bargaining status in their relationships; and decrease their exposure to violence (Farmer and Tiefenthaler, 1996; Stevenson and Wolfers, 2006; Aizer, 2010; Hidrobo and Fernald, 2013). On the other hand, an increase in the resources available to women may strengthen the incentives of men to use violence or threats of violence in order to control these newly obtained resources or to regain decision-making power within the household. As a result, women may become more vulnerable to mistreatment (Bobonis, González-Brenes and Castro, 2013; Eswaran and Malhotra, 2011; Bloch, Rao and Desai, 2004).

women's threat point increases when they have access to justice and when support services are more helpful. Relatedly, the introduction of WJC centers may have helped break the silence regarding violence against women, turning it into a public issue. The awareness campaigns may have contributed to building trust among women for the justice system.

Consistent with this mechanism, we find suggestive evidence of an improvement in the bargaining power of women in the household. In particular, we find that women living near a WJC center are more likely to make joint decisions with their husband, less likely to earn less than their husband and more likely to earn as much as their husband. We also find that institutional trust is reinforced in the areas close to the WJC center. Both mechanisms lead to the conclusion that WJC's intervention in households with abuse may change the behavior of offenders and victims by improving the situation of the woman within the household and, consequently, their ability to take care for their children.⁶

The main threat to our identification strategy is time-varying unobservables that are correlated to both the timing of the WJC center introduction and changes in the prevalence of domestic violence and education outcomes. To ensure that our results are not driven by selection or time-varying unobservables, we perform several falsification exercises and robustness checks. First, in order to control for the nonrandom placement of the WJC centers, we also include a province-by-year fixed effect which controls for any characteristics that may vary at the province and year level. By using province-by-year fixed effects, our identification assumption is that treatment schools/households would otherwise have changed similarly, on average, to control schools/households within their same province. Second, we focus our analysis in the middle of the rollout period for which identifying assumptions are likely to hold. In particular, we show that schools/households reached by the WJC centers from 2006 till 2014 had similar pre-program trends. Third, we show that WJC center placement was not anticipated by changes in gender-based violence and schooling.⁷ Lastly, we limit the samples to areas most comparable to the those with WJC center presence: urban schools and urban clusters of households, since the WJC centers were more likely to be located in more densely

⁶This mechanism may operate by changing women bargaining power within the household and also by improving their health since now they are exposed to less violence. While we cannot disentangle both effects, we find evidence that both mechanisms may be important.

⁷A central issue in our analysis is the fact that WJC centers are not placed randomly. Conversations with policymakers and WJC center managers suggest they choose where to locate primarily based on population density, the level of infrastructure and proximity to several institutions, but there was no mention of locating based on anticipated increases in violence and schooling or previous years increases.

populated areas. We further examine the results by limiting the sample to districts which ever had a WJC center.

To the best of our knowledge, this is the first quantitative analysis that attempts to explore the impact of an unexamined dimension of institutional intervention, which provides better access to justice for women, on the incidence of gender-based violence against women and its spillover effects on children's human capital. This study does not only provide evidence of the effectiveness of an important component of Peru's public policy aimed at curbing gender-based violence, but it also contributes to the literature on gender, crime and development by providing a new insight on gender-based violence against women, women's empowerment in developing countries and its indirect effect on children's education outcomes.

This paper is closely related to recent papers exploring the effect of police stations run by women in India and the US ([Miller and Segal, 2016](#); [Amaral, Nishith and Bhalotra, 2018](#)). Both papers find that as female representation increases among police officers in an area, violent crimes against women in that area are reported to the police at significantly higher rates. While these papers focus on the share of women at the police station, we explore another dimension which is access to justice for women holding constant women representation. In addition, these papers mainly focus on whether there is change in reporting gender-based violence. We complement this research by also showing that actual violence against women decline and children's outcomes improve.

This paper is also related to the literature linking economic conditions and gender-based violence. While it is often assumed that improving the economic situation of women and ensuring that women get an equal share of resources within the household will alleviate gender-based violence, recent research has shown this is not always the case ([Bobonis, González-Brenes and Castro, 2013](#); [Eswaran and Malhotra, 2011](#); [Bloch, Rao and Desai, 2004](#)). In contrast, we find that access to justice that reduces gender-based violence generates additional benefits to women and their children. We argue that these results are consistent with a bargaining model in which the threat point is determined by access to justice.

Relatedly, this paper complements the literature studying household educational investments on children. While much of the literature on educational investment has focused on the role of policies that alleviate resource constraints such as conditional cash transfers or scholarships programs (see [Glewwe and Muralidharan \(2016\)](#) for a review of this literature),

we instead examine the role of a policy that affects the cost of violence against women by providing access to justice. Our results highlight that investments outside the education sector such as GBV interventions can have first order effects on children human capital. Finally, this paper complements previous literature showing that an increase in women's income appear to benefit children more than an increases in men's income ([Attanasio and Lechene, 2002](#); [Thomas, 1990](#); [Lundberg, Pollak and Wales, 1997](#)).⁸ While in this case we do not find a change in women's income or labor force participation, when justice increases they are more likely to invest on children.

The remainder of this paper is organized as follows. Section 1.2 presents a brief background on the prevalence of domestic violence in Peru and on the WJC center intervention. Section 1.3 describes the data. Section 1.4 presents the empirical strategy. Section 1.5 presents the main results. Section 1.6 investigates the channels through which WJC center introduction affects domestic violence and schooling. Section 1.7 provides supporting evidence consistent with the identification assumptions. Section 1.8 concludes.

1.2 Background

1.2.1 Gender-based violence in Peru

Gender-based violence is currently one of the most pressing social problems in Latin America and the Caribbean. Even though the region has received much attention on conflict, crime, political and economic instability, it is easily overlooked that violence against women is among the most pervasive types of violence in the region ([Fregoso and Bejarano, 2009](#); [Heinemann and Verner, 2006](#); [Londoño et al., 2000](#)).

Among the Latin American countries, Peru has gained a considerable amount of attention in recent years, largely due to the high prevalence and severity of gender-based violence in this country. According to a study carried out in 10 countries by the World Health Organization in 2006, the prevalence of physical violence by a male partner ranges from 13% in Japan's urban regions to 61% in rural areas of Peru and 49% in urban areas of Peru ([Garcia-Moreno et al., 2006](#); [Morrison, Orlando and Pizzolitto, 2007](#)). [Flake and Forste \(2006\)](#) study the relationship between

⁸Most of this literature finds that households in which women's income share is higher spend a larger fraction of their income on children's clothing and food.

household characteristics and the likelihood of experiencing gender-based violence in Colombia, Dominican Republic, Haiti, Nicaragua and Peru. They find that although the prevalence of gender-based violence is high in all five countries, Peru had the highest percentage of instances at 38.9% followed by Nicaragua (26.1%), Dominican Republic (22.6%), Colombia (19%) and then Haiti (15.7%). Data collected by *Instituto Nacional de Estadística e Informática* (INEI) through the Demographic Health Surveys have found that although the prevalence of violence (physical and/or sexual) affecting women has declined from 41.2% to 32.6% from 2000 to 2015, it still remains quite high (INEI, 2001, 2015). All this evidence suggests that Peru is very high on the world ranking of registered cases of domestic violence and among the leaders in Latin America in terms of prevalence of violence against women.

While the majority of intimate partner violence is perpetrated within the domestic sphere, Peru's institutions also have a reputation for gender-based violence, including sexual violence. For many decades, women in Peru have been subject to abuse- even by the one entity supposed to protect them: the state. For instance, in the 1990s and early 2000s, Peru witnessed one of the most heinous violations of women's rights in recent history: under the administration of Alberto Fujimori, thousands of women were forcibly sterilized in an attempt to prevent overpopulation and poverty. The state is also complicit in institutional violence against women, ranging from insults to injury in its hospitals, health centers and schools (Boesten, 2012).

Despite legislative progress in identifying and addressing the problem, the legal system has constantly been characterized as ill-equipped to efficiently process complaints. In the early 1990s, Peru was one of the first countries in the region to develop legislation and policy to address violence against women. The Law for Protection from Family Violence was first adopted in 1993 and strengthened in 1997, attempting to codify intimate partner violence as a criminal offence while producing a distinct and expedited procedure for victims to lodge complaints. However, these legal reforms in the area of violence against women lacked a clear legal framework and have done very little to curb its persistence. In short, "many women do not bother to file complaints because the legal system is too slow to act" (UNHCR, 2010). In addition, in rural Peruvian communities, women are often assumed to have a certain level of blame in a conflict (Revilla, 1999). Traditional methods of justice based on local customs also are often discriminatory towards women and rarely impartial (Franco and González, 2009).

A major contributor to the persistence of gender-based violence in Peru is a deeply embed-

ded inequality of gender roles and status, which is usually represented through the notion of *machismo*. Machismo defines gendered behaviors, which makes Peruvian families more susceptible to domestic violence, since women are expected to accomplish familial obligations unconditionally within a patriarchal family system. In addition, the high rates of intimate partner violence might also be explained through dynamics of historical and cultural factors which are based on the subordination and discrimination of women (Mitchell, 2013). Within this context, the persistence of violence against women is a clear sign of women disempowerment, which impinges on women's autonomy within the household.

1.2.2 Women's Justice Centers Program

The 1994 Inter-American "Belem do Pará" Convention on "Prevention, Punishment, and Eradication of Violence against Women" significantly expanded Latin America's definition of domestic and sexual violence. As a consequence, many countries in the region modified or enacted new legislation incorporating those issues into their political agenda. In particular, Peru altered its Police and Justice System's jurisdiction to encompass domestic and sexual violence complaints. This new legal framework paired with the government's awareness of the country's high levels of domestic violence led in 1999 to the creation of the women justice centers (WJCs) – "*Centros de Emergencia para Mujeres*" – by the *Peruvian Ministry for Women and Vulnerable Populations* (MIMP) as part of the National Program against Sexual and Family Violence.⁹

The *women's justice centers* (WJC) are free of charge public centers that aim to strengthen the justice system's capacity to detect, process and assist victims of domestic and sexual violence from an inter-disciplinary approach that includes legal, social and psychological dimensions. Basically, incoming victims receive a service designed to integrate all steps of the complaint process (e.g. police station, attorney's office and medical doctor) in a single office in order to reduce as much as possible the time dedicated to issue the complaint and to follow the legal procedure in the corresponding court of justice. Hence, WJCs are frequently located within a short distance from partner establishments such as police stations, prosecutors' offices and

⁹The Peruvian Ministry for Women and Vulnerable Populations, known as *Ministerio de la Mujer y Poblaciones Vulnerables* - (MIMP) used to be called as Ministry for Women and Social Development (*Ministerio de la Mujer y Desarrollo Social* - MIMDES) when the WJC center program was rollout in 1999. <http://www.mimp.gob.pe/contigo/contenidos/pncontigo-articulos.php?codigo=14>

health facilities.¹⁰

In addition to assisting incoming victims, WJC center's aim is also to undertake a local violence prevention program. The prevention component intends to identify, control and reduce the risk factors. In this regard, the WJC centers have put in practice courses for training justice promoters –“*facilitadoras en accion*” and “*promotores juveniles*”–, which are volunteer women that advocate and execute campaigns, talks, workshops and seminars to raise awareness about the problem of domestic violence in their region. Lastly, WJCs keep a record of cases that allow for monitoring and evaluation of the persistence of domestic and sexual violence (MIMDES, 2007).¹¹

The first women's justice center opened in the District of Lima in 1999. During 1999-2014, the number of centers has grown from 13 in the first year to 226 by the end of 2014, covering 100% of the 24 regions of Peru and 96% of the provinces (188 of 196 provinces). Figure 1.1 shows the distribution and growth of the opening of the WJC centers over time. Whereas WJCs centers opened gradually throughout the first years of implementation, the program expanded exponentially after 2006. Up to that year, the average opening rate was about 6 WJCs/year; from 2006 to 2014 it augmented to 22 WJCs/year. Such escalation was provoked by a 2006 decentralization decree that granted local governments the right to open their own WJCs at the district level.

From a geographical coverage point of view, by 2014 most of the WJCs are concentrated in Metropolitan Lima and Lima Provinces (31 WJCs); in the Callao region there are 4 WJCs; the rest of the coastal region have 46 WJCs; in the sierra region there are 117 WJCs and in the jungle region there are 28 WJCs (Figure 1.2). Given the before-mentioned strong ties to local justice and health institutions, WJCs are highly located within urban areas.

According to MIMP's statistics, the number of domestic violence cases registered in the WJC centers has increased substantially: from 29,759 in 2002 to more than 60,000 in 2016 (See Figure 1.3). Whereas 40% of reported cases are from women between 25 and 45 years old, children and teenagers (0-17 years old) constitute the second largest group – 30%. Additionally, a 2006-2008 survey administered by MIMP on 51 WJCs revealed that for the majority of the

¹⁰The service provided in these centers is staffed by representatives of various government institutions such as police officers, prosecutors, counsellors, psychologists and public welfare agents whose objective is to help the victims of domestic abuse (MIMDES, 2007).

¹¹Ministerio de la Mujer y Desarrollo Social. 2007. ¿Que son los Centros de Emergencia Mujer?. Available at http://www.mimp.gob.pe/files/programas_nacionales/pncvfs/Centros_Emergencia_Mujer_MIMDES1.pdf

women (75%) who attended a WJC, domestic violence stopped during and after the program's intervention. However, the remaining 25% indicated a persistence of violence even after having attended a WJC center (MIMDES, 2009).¹²

Albeit these evaluations, the program's effectiveness remains unclear. MIMP's statistics lack a rigorous study of the causality of factors and its mechanisms. Although globally, WJCs seem to have a positive effect on curbing the incidence of domestic violence, it remains uncertain if such effect is due to the centers' ability to act as catalysts for women empowerment—which indirectly could enhance women's capacity to care for their children through improved social protection and access to justice. Additionally, given our detailed micro-level of analysis, our results could become instrumental to understand the effectiveness of similar programs implemented in other Latin American countries.

1.3 The Data

This paper makes use of three different types of datasets which provide variation across geographical regions and time at different levels of aggregation: individual and household-level data, school level data and administrative data on WJC centers, femicides, female deaths due to aggression and female hospitalizations for mental health problems at the district level.

1.3.1 Individual and Household Level Data

To study the impact of WJC centers on women's and their children's outcomes, we rely on microdata from the *Peruvian Demographic and Health Survey* (DHS), which is collected over the period 2000-2014.¹³ These surveys are cross-sections designed to be representative at the national and regional (second administrative) levels. The DHS employs a stratified random cluster sampling procedure in which the country is divided into several primary sampling units (in this case, districts) and clusters of households are randomly selected.

The *Peruvian DHS* collects primarily demographic and health information from women

¹²Ministerio de la Mujer y Desarrollo Social. 2009. Investigación operativa: "Eficacia de la intervención de los Centros Emergencia Mujer". Available at http://www.mimp.gob.pe/files/programas_nacionales/pncvfs/estadistica/eficacia_intervencion_cem.pdf

¹³The *Encuesta Demografica y de Salud Familiar* (ENDES) is the Peruvian version of the Demographic and Health Surveys (DHS). These surveys are available for the following years: 2000, 2004-2008 and 2009-2014. The *Peruvian DHS* is exceptionally a continuous survey, which means that the data is collected quarterly instead of every five years since 2004.

aged 15 to 49 years that include their fertility, weight, marital status, employment status, household decision making and socio-economic characteristics among others. In addition to this, it also includes some demographic and socioeconomic characteristics for each of the women's household members (e.g. husband and other children), which we exploit in our analysis.

In addition to the standard survey, the *Peruvian DHS* also includes a domestic violence module which asks eligible women if they have ever experienced physical, sexual or emotional abuse from their current or previous partner in the last 12 months.¹⁴ While all women between the ages of 15 to 49 are asked to participate in the standard survey, only one woman in each household, who has ever been married or partnered, is randomly selected to complete the domestic violence module. Women who are never married or never cohabited are excluded from the sample. This selection process is taken by the DHS program in order to minimize underreporting of domestic violence events.¹⁵ The DHS captures four different types of domestic violence: moderate physical violence, severe physical violence, sexual violence and emotional violence. These domestic violence categories are defined by the DHS as ex-post classified questions.¹⁶ Since the last measure is less visible and more difficult to measure, in this study we define exposure to a domestic violence event if the woman has ever experienced

¹⁴It should be noted that though this is an important measure of domestic violence, it does not report the different forms of gender-based violence that affect women beyond spouses and inter-family relationships.

¹⁵The domestic violence module of questions is implemented only to a subsample of the women selected for the *Peruvian DHS* sample. There are three security and ethical precautions increasingly mandated by the DHS program for the collection of data on domestic violence. The first requires that the interviewer does not continue with the questions on domestic violence if privacy cannot be ensured. In general, the interviewers are women trained to elicit trust from the respondents. The second requires that only one eligible woman in each selected household is to be administered the module questions. In sample households where more than one woman is eligible for the DHS survey, the domestic violence module is administered to only one randomly selected woman. By interviewing only one woman in each household, possible security breaches, due to other persons in the household knowing that information on domestic violence was given, are minimized. The third requires that the domestic violence questions should be only administered to ever-married or cohabiting women, even though the DHS sample includes all women age 15-49. Only 1% of the eligible women was not interviewed because privacy was not made possible in the household. Despite the selection measures taken by the DHS program, this empirical analysis may still suffer from measurement issues due to underreporting. In order to account for this, we employ several different outcomes to measure violence against women: femicides and female deaths due to aggression

¹⁶More specifically, the DHS defines *moderate physical violence* if the woman experiences at least one of these acts from their spouse or partner: (a) spouse ever pushed, shook or threw something, (b) spouse ever slapped respondent, (c) spouse ever punched respondent with fist or something harmful, (d) spouse ever kicked or dragged respondent. *Severe physical violence* is defined if the woman experiences at least one of the following acts: (e) spouse ever tried to strangle or burn, (f) spouse ever threatened with knife/gun or other weapon, (g) spouse ever attacked with knife/gun or other weapon. *Sexual violence* is defined if the woman experiences at least one of the following acts: (h) spouse ever physically forced sex when not wanted, (i) spouse ever forced other sexual acts when not wanted (j) spouse ever twisted arm or pulled hair.

any type of moderate, severe physical or sexual violence during the last 12 months.

The main advantage of using this household survey is that we can link children's outcomes (e.g. school attendance status, child labor) with their mother's self-reported domestic violence. This information is crucial in order to be able to understand the mechanisms behind the results. Since we do not observe enrollment rates with the School Census, we use the *Peruvian DHS* to estimate the share of children in primary and secondary level who are enrolled and attending school.¹⁷ This survey also allows us to measure children's schooling performance (e.g. passed grade, repeated, dropped out) and whether he/she is performing any child labor.

Panel B of Tables 1.1 and 1.2 provides summary statistics on women's characteristics and on children's school attendance status during 2006-2014, respectively.¹⁸ According to the *Peruvian DHS*, the data indicate that 39% of ever-partnered Peruvian women declared to have experienced abuse from their spouse during the last 12 months, which is remarkably high. As for children's education outcomes, the school attendance rate in primary level is 97% for both boys and girls, which is almost universal. The school attendance rate in secondary level is also quite high (89%) and very similar between genders. Given that secondary school is not compulsory, the drop-out rate reaches 9% of the students in this educational level.

In addition, the *Peruvian DHS* also records GPS coordinates for every cluster of households in a certain district, which allows us to measure not only presence of WJC center in the district of residence but also proximity to the WJC center. Although this data was collected yearly, in this study we were able to obtain the GPS cluster locations only for the 2000, 2004-2008, 2009-2011 and 2014 *Peruvian DHS* Surveys. Since the DHS does not disclose the name of the villages (*centros poblados*) where the clusters are located, the final sample is a repeated cross-section of individuals (women and children), where the lowest geographical unit we can condition on is the district.

Our concern with this database is linked to the fact that GPS locations of the sampled DHS clusters of households are displaced before public release to preserve confidentiality of respondents. The GPS displacement is randomly carried out so that: urban clusters are uniformly displaced up to 2 kilometers and rural clusters are displaced up to 5 kilometers,

¹⁷For the children's school attendance analysis, we also use the 1996 *Peruvian DHS* in order to assess the validity of the identification strategy.

¹⁸We focus our analysis in the middle of the rollout period, 2006-2014, for which identifying assumptions are likely to hold. We discuss this choice in more detail in Section 1.7.

with 1% of the rural clusters displaced up to 10 kilometers. In addition, the displacement is restricted so that the points stay within the second administrative level, which is the province. Therefore, the GPS displacement procedure introduces a random error, which can substantively affect the results of the analysis (Burgert et al., 2013).

Nevertheless, Perez-Heydrich et al. (2013) propose several recommendations in order to reduce any distance measurement error. Firstly, they suggest that the amount of measurement error depends on the spatial density of the resource facilities. As the density of the resource facilities decreases, the probability that a DHS cluster is linked to the correct closest WJC center increases for all types of locations (urban and rural). In Peru, there are a total of 226 WJC centers by 2014, which means that the spatial density of the WJC centers is quite low and, thus, the measurement error is quite reduced. Secondly, the authors recommend to study the effect of the service within a reasonable buffer distance, rather than using the closest-distance to the resource facility. For this reason, we are going to measure exposure to the WJC center through different groups of Euclidean distance buffers. Lastly, we are also going to limit the analysis to urban areas because in these locations the range of displacement is less than in rural areas.

1.3.2 School Level Data

We use two school level datasets: the Peruvian School Census (*Censo Escolar*, CE) and the Census Evaluation of Students (*Evaluacion Censal de Estudiantes*, ECE). The Peruvian School Census is a large panel dataset on primary and secondary school enrollment, which covers the universe of schools in Peru during the period 1998 - 2014. This dataset is collected on a yearly basis by the Peruvian Ministry of Education, with exception of the year 2003 and it contains a rich set of information at the school level.

More specifically, the School Census collects comprehensive data on the total number of enrolled students by age, grade and gender. These data are designed to reflect enrollment (not attendance) statistics corresponding to the months of May-July. The School Census also collects data on school characteristics, such as language of instruction, public or private, urban or rural area and other physical plant characteristics (i.e. electricity, piped water etc). We complement these data with the Census Evaluation of Students, which contains the standardized test scores of a national exam administered every year to all primary school students in second grade during the period 2007-2014. This exam has two portions: math and language (Spanish) skills.

Each school in these datasets is given a unique ID number, which allows us to follow schools over time. In addition, one of the main advantages of these school datasets is that they are geo-coded, which means that we can observe the exact location of the school. The geographic coordinates of the schools allow us to combine these data with the WJC center's locations, in order to see whether the area/district of the school is located near a WJC center and thus affected by the opening of these centers that provide specialized attention to victims of domestic and sexual violence.

Panel A of Table 1.3 shows the years of data coverage and the number of schools by rural/urban region. In order to be consistent with the individual level data, for this analysis we also use data which cover the period 2006-2014. In the later years, the dataset covers a larger share of schools. It is important to note that this dataset is not a balanced panel because during the period of study some schools have closed, while others have opened. In addition, as mentioned above, there is no data available for the year 2003, since data for this year was not collected. Although this means we do not have a balanced panel, by including school fixed effects we ensure that we compare the same schools over time. The main analysis, then, draws on a nine-year unbalanced panel dataset of 36,994 primary schools (grades one through six) and 12,811 secondary schools (grades one through five).¹⁹

Panel C of Table 1.3 provides some summary statistics on school enrollment and school characteristics. The average primary school in our sample has 95.9 students, while the average secondary school has 175 students. The proportion of primary schools is higher in rural areas, while secondary schools are more likely to be found in urban areas. The majority of primary schools are public and teach in Spanish language, but there is also a small proportion that teach in Quechua and other native languages. In contrast, a large proportion of secondary schools (40%) are private and in almost all of them the language of instruction is Spanish.

A final important issue of the School Census data is that it measures total number of children enrolled, not enrollment/attendance rates. This may lead to the concern that our results reflect changes in population. However, we discuss this issue in greater detail in Section 1.4. In addition, we also use the *Peruvian DHS* to estimate the share of children who are attending school as a robustness check.

¹⁹The primary-school sample covers between 4.1 and 3.5 million students each year, whereas the secondary school sample covers between 2.3 and 2.7 million students.

1.3.3 District Level Data

Information on the rollout of the WJC centers was provided by the *Peruvian Ministry for Women and Vulnerable Populations* (MIMP) and consists of a directory of WJC centers across all over Peru. This directory contains the name of the WJC centers, their founding dates (date-month-year), their administrative locations (district-province-department) and their addresses during the period 1999 to 2014. By using the administrative locations and addresses provided in the directory of the MIMP, we were able to geo-code all the WJC centers, which allows us to have not only the district where they are located but also their exact GPS location.

This data collection project resulted in a dataset of 226 WJC centers from 1999 till 2014. Figure 1.1 shows a histogram of WJC center founding dates and it also illustrates the evolution of the opening of WJCs since 1999 till 2016, while Figure 1.2 maps the rollout of the WJC centers at the national level, which allows to visualize the extensiveness and national scope of the program. From both graphs, we can clearly see a substantial growth in the number of centers over time, where 81% of them are founded after the year 2005.

Data on the number of femicides at the district level was obtained from the *Peruvian Crime Observatory at the Ministry of Public Affairs* and it covers the period 2009-2015. In Peru, femicides are classified in two categories: (1) *intimate femicides*, when the homicide is committed by the woman's partner, ex-partner or other family member (2) *non-intimate femicide*, when the homicide is committed by a stranger, neighbor, friend or a client who murders a sex worker (INEI, 2017). This data is recorded by each district attorney office in the country. Unfortunately, this data is only available at the district level and it is not geo-coded. In this analysis, we only consider the cases of intimate femicides in order to be consistent with the DHS data. From 2009 till 2015, 852 women have been murdered in Peru of which 762 (90%) are intimate femicides and 90 (10%) are non-intimate femicides (Figure 1.6).

We complement this information with data on female deaths due to aggression and female hospitalizations for mental health problems, which were obtained from the *Peruvian Ministry of Health - National Institute of Statistics and Informatics (INEI)*. This database contains the number of registered cases of hospitalizations by type of illness, age and gender. For the purpose of this analysis, we use female hospitalizations for mental health problems. It also records the number of hospitalizations that resulted in deaths according to different types of causes. The main female cause of mortality that is relevant to this analysis is death due to aggression. This

information is recorded by health facilities such as hospitals and is only available at the district level. The number of registered cases in health facilities includes women between the ages of 18 and 59 and covers the period 2006/7-2015. Figure 1.6 shows the number of female deaths due to aggression and female hospitalizations for mental health problems over time in Peru.

1.3.4 Measuring Exposure to the WJC Centers

In order to be able to match the data on WJC centers with the the outcomes of interest, we construct two measures of exposure to the program: (i) WJC center within a 1km Euclidean buffer of the DHS cluster/school and (ii) WJC center in the district of the DHS cluster/school.

The first measure uses the GPS coordinates of the DHS clusters/schools in order to measure a 1 kilometer Euclidean distance buffer from every DHS cluster/school location. For this method, the Euclidean buffer of 1km is first centered on each DHS cluster/school and then each DHS cluster/school is linked to a WJC center if the WJC center falls within the buffer, without consideration of district administrative borders. For instance, a DHS cluster/school located within 1km of a WJC center founded in 2008 is coded as having a WJC center within 1km of the DHS cluster/school since 2008. Figure 1.5 shows a visual representation of the Euclidean buffers for two specific regions in Peru: Lima and Tumbes.

The second measure matches the presence of a WJC center in the district, based on its date of opening and location, with the DHS cluster's/school's district. For instance, a DHS cluster/school in the district of Lima (150101) with a WJC center introduced in 2006 is coded as having a WJC center in the district of Lima since the year 2006.

Our preferred measure is the one that uses the Euclidean buffer since we want to estimate the impact of having a WJC center in the neighborhood of the school/household. The second measure is used as a robustness check because it might not always capture accurately the impact of the WJC centers due to the fact that districts in Peru have very different sizes. Therefore, rather than aggregating WJC center exposure in the district, we measure exposure based on how far the centers are from respective households, such that individuals residing at different points in the same district may have different levels of exposure to the WJC centers. Panel A of Tables 1.1 and 1.2 and Panel B of Table 1.3 show descriptive statistics of exposure to the WJC centers at the individual (women and children) and school level. The main reason for our choice of a 1km distance buffer instead of a larger buffer is not only because we believe

that these centers have a very localized effect, but also because the measure of exposure using a 5km Euclidean buffer seems to be very similar to the one that uses presence of WJC center in the district. We present results using both measures of exposure to a WJC center principally for our main outcomes of interest.

1.4 Empirical Strategy

1.4.1 Placement of WJC centers

A central methodological issue in our analysis is the fact that WJC centers are not placed randomly across the country. Even though our analysis will take advantage of variation over time, which will account for any fixed differences across districts and schools, it still remains important to understand what drives placement since placement decisions may not be orthogonal to other factors that could affect women's and children's outcomes of interest.

We address this concern in a number of ways which lead us to believe that the link between the opening of the WJC centers and the outcomes of interest is casual. First, we had several discussions with the Peruvian policymakers and WJC center managers about the location choices. Since the foundation of the first WJC center in 1999 till the end of 2005, the primary criteria they cited when deciding where to locate were population density and level of infrastructure at the regional level. In this stage, capitals and large cities were prioritized locations to open a WJC center. Starting from 2006, after the decentralization process which transferred the responsibility of the WJC centers to the local governments (districts), the Peruvian policymakers decided to open new WJC centers at the district level and they incorporated additional criteria such as proximity to police stations, district attorney offices (known as *fiscalias*) and health establishments.

Even though program guidelines suggested that priority should be given to poorer districts with sufficient judicial and medical infrastructures, in several occasions, political representatives had certain autonomy in deciding the order in which districts received the program. There is also anecdotal evidence from the authorities that WJC center's placement has been primarily developed taking into account the population density but failed to take into account the rate of incidence of violence against women. This is likely due to the lack of reliable data on domestic violence or femicides for all the districts in Peru prior to the opening of the centers.

For instance, official data on femicides in Peru started to be recorded since 2009.²⁰ Moreover, our conversations with the Peruvian policymakers suggest that educational considerations, and in particular enrollment rates or schooling performance, were never factored into program placement decisions.

Second, we are able to evaluate this endogenous placement statistically using our data. To do this we estimate at the district level: (a) the determinants of having a WJC center by the end of the sample in 2014 and (b) the determinants of adding a WJC center during 2006-2014, which is the period when the program grew substantially. We focus on several variables at the district level cited by the Peruvian policymakers such as: number of justice courts, number of district attorney offices, number of police stations and number of health establishments. We also control for district population at baseline and department fixed-effects. Moreover, in order to verify that education patterns before the program do not predict where the WJC centers are introduced, we also control for pre-program changes in primary and secondary school enrollment at the district level. Unfortunately, we are unable to perform the same test for self-reported domestic violence or femicides due to lack of pre-program data on these variables for all the districts in Peru. However, we control for baseline (self-reported) domestic violence at the district level by using the *2000 Peruvian DHS* which contains a representative sample of 700 districts in Peru.

The results from these regressions are shown in Table 1.4. In general, the results corroborate the evidence we collected from our conversations with the Peruvian policymakers and WJC center managers. Districts with more police stations, more district attorney offices, more health establishments and more densely populated are more likely to have WJC centers by 2014 and more likely to add them during 2006-2014. Clearly, urban areas with more infrastructure development are more likely to have these specialized centers for women. In addition, pre-program changes in primary and secondary district school enrollment do not seem to have any impact. Neither coefficient is statistically significant and both are very small. Similarly, domestic violence does not have any impact on WJC placement. These findings suggest that WJC center placement between 2006-2014 does not seem to have been based neither on

²⁰Several ministerial reports have documented the fact that WJC centers failed to consider the rate of incidence of violence against women in program placement. See, for instance, Ombudsman Office, *Informe Defensorial N 144. Centros de Emergencia Mujer: Supervisión de los servicios especializados en la atención de víctimas de violencia familiar y sexual*, July 2009. Ministerio de la Mujer y Desarrollo Social, *Investigación operativa: "Eficacia de la intervención de los Centros Emergencia Mujer"*, August 2009

pre-program changes in schooling nor on baseline domestic violence.

Finally, we note two additional concerns that might threaten the validity of our research design. First, one might be worried that another shift (e.g. a government program or policy change) might be rolled out during the same period and in the same places as the WJC centers, which might also have an impact on education outcomes. An obvious candidate is the CCT program *Juntos*, which was launched in September of 2005, right at the time when the WJC centers started to be implemented more intensively.²¹ In addition to this, *Juntos* integrates two broad objectives: in the short run, it aims to reduce poverty by providing households with cash transfers; and in the long run, it aims to break the inter-generational transmission of poverty by promoting human capital through improving access to education and health services.

In spite of this, several reasons lead us to believe that *Juntos* is not a confounding factor in our empirical strategy. Districts were selected for program participation based on an index that includes poverty and percentage of villages affected by violence during civil conflict. The aim of *Juntos* was to reach some of the most vulnerable and marginalized segments of the population and focused particularly on rural areas with high poverty rates and limited access to State services.²² By 2014, about 1142 districts have CCTs and 225 districts have WJC centers. However, more than half of the districts with WJC centers (123 districts) are not covered by the CCT *Juntos* program. This evidence clearly suggests that while WJC centers were more likely to be implemented in urban areas, the CCT program was more likely to cover dispersed populations in the poorest rural areas. We test this assumption more directly by analyzing whether the WJC placement at the district level was correlated with the CCT *Juntos* implementation. Columns 2 and 4 in Table 1.4 indicate that the WJC center placement was not determined by the rollout of the CCT *Juntos* program.²³

²¹See Figure 1.7 on the presence of both programs at the district level and Figure 1.8 on the timing of CCT *Juntos* and WJC centers programme implementation. There are two large expansions of the CCT *Juntos* implementation, first in 2007 and then in 2012.

²²*Juntos* is targeted to the population living in poverty and extreme poverty: households with children under 14, pregnant women, widowed parents and/or older adults. It is particularly focused on getting children out of poverty, improving their education, health and nutrition. This programme is also explicitly seen as a way to tackle the special vulnerability of populations who were most affected by the political violence that was prevalent in Peru between 1980-2000. Most of the victims of this conflict were poor populations living in rural areas and Quechua speakers.

²³We also construct a panel database at the district level on WJC center and CCT *Juntos* placement from 2005 till 2014, which allows us to better analyze whether program implementations were correlated over space and time. By using a fixed-effects model, we can control for any time-invariant locality factors at the district level and also year dummies. The results in Table 1.A.1 corroborate the idea that the CCT *Juntos* is not a confounding factor in our research design.

The second concern related to WJC center placement is that if we estimate the impact of the WJC centers on all areas, our results might be identified off of rural areas which are not at risk of having a WJC center and these may not be an accurate comparison for those areas which get a WJC center. Given this, we will focus our analysis on a specification in which we limit the sample to urban areas (urban school and households), which are the ones more “at risk” of opening a WJC center. As a further robustness check, we will also limit our samples to districts which ever have a WJC center during the sample period.

1.4.2 Individual Level Specification

We use a difference-in-difference empirical strategy to estimate the impact of WJC centers on women’s and children’s outcomes. We exploit the variation created by the differential timing in the opening of the WJC centers and also the spatial variation in the exposure of a woman/child to a WJC center. In order to estimate the impact of WJC centers on women’s and children’s outcomes, the following specification is used:

$$y_{idt} = \gamma_0 + \gamma_1 WJC_{idt} + \alpha_d + \lambda_{pt} + \delta X'_{idt} + \varepsilon_{it} \quad (1.1)$$

where (y_{idt}) represents the outcome of interest of woman i (or the child of woman i) at year t who resides in district d , (WJC_{idt}) is an indicator variable that takes the value of one if there is a WJC center within 1km of the women’s/child’s household or in the district of residence of women/child i in year t , (α_d) is a district fixed-effect, (λ_{pt}) is a province-by-year fixed-effect, (X'_{idt}) is a vector individual-level characteristics for women/child i depending on the sample of interest and (ε_{idt}) is a random error term. Standard errors are clustered at the district level and we also include district-specific time trends. The inclusion of districts fixed-effects account for possible time-invariant unobserved characteristics at the district level, such as cultural differences or attitudes towards the role of women/children. However, this does not account for any differential trends in women’s/children’s outcomes associated with WJC center placement. To address this, we allow the year fixed-effects to differ by province. Province-by-year fixed effects rule out the concern that our results are driven by changes that vary by province and year such as an increase in political corruption or a decrease in provincial resources.

There are two main measures of domestic violence to be used as dependent variables for

women's specification. The first is a measure of physical domestic violence which is defined as a binary indicator that takes value 1 if the woman reports any moderate, severe or sexual abuse from the intimate partner during the previous year. The second measure is a binary indicator for emotional violence, which is based on three questions referred to behaviors or situations that are considered as strong indicators of mistreatment by experts. We also use a set of outcomes for women's health/nutritional status such as anemia status, weight, body mass index etc. The vector X'_{idt} includes a set of control variables for woman's age, age at first marriage, number of children, years of education, number of household members, number of families in the dwelling, marital status and rural-urban residence.

Since our school level data contain number of students enrolled, but not enrollment rates, we use the *Peruvian DHS* to estimate the impact of WJC centers on children's school attendance status. The main child outcome variable is a dummy variable indicating whether the child is attending school during the year of the survey. We also use additional schooling performance outcomes, which are defined as a changes in school attendance status between one year and the next, conditional on the child being enrolled at school. Therefore, the dependent variable can be classified as: (a) currently attending school, (b) passed grade (c) repeated grade (d) dropped out and (e) left school more than 2 years ago. For the children's specification, we also include a set of control variables such as age, gender, household's head years of education, number of children in the household aged 0-18, number of children in the household aged 0-5, number of female adults, number of male adults and rural-urban residence.

The coefficient of interest is (γ_1) , which captures the average change in outcomes of women/children that are located near the WJC centers or in districts with WJC center, to the average change in outcomes of women/children that are not reached by a WJC center. The identification assumption is that in the absence of the WJC centers, treatment households (women and children) would otherwise have changed similarly, on average, to control households within their same province. Note that in this specification we cannot control for individual fixed-effects because the *Peruvian DHS* databases of women and children are repeated cross-sections.

1.4.3 District Level Specification

We then estimate the following equation to capture the impact of WJC centers on femicides, female deaths due to aggression and female hospitalizations for mental health problems at the

district level:

$$y_{dt} = \gamma_0 + \gamma_1 WJC_{dt} + \alpha_d + \lambda_{pt} + \delta X'_{dt} + \varepsilon_{dt} \quad (1.2)$$

where (y_{dt}) refers to alternative domestic violence metrics (e.g. femicides by intimate partner, female deaths due to aggression) and hospitalizations for mental health problems aggregated at the district level in year t , (WJC_{dt}) is an indicator variable that takes the value of one starting in the first year in which district d offers a WJC center, (α_d) is a district fixed-effect, (λ_{pt}) is a province-by-year fixed-effect, (X'_{dt}) represents time-varying district level covariates (e.g. district population), and (ε_{dt}) is a random error term. In this case, we are unable to use exposure to a WJC center within a 1km Euclidean buffer as treatment since the outcome variables are only available at the district level and they are not geo-coded. For this specification, the dependent variables are defined using the logarithm (instead of the level).

This is a standard fixed-effects model, where identification is derived from changes in domestic violence/mental health outcomes correlated to changes in the introduction of WJC centers in the district. This empirical strategy allows us to account for both time-invariant characteristics of districts, and time-varying characteristics common between treatment and control districts. Therefore, the identification assumption is that any unobserved time-varying covariates that affect domestic violence/mental health outcomes are uncorrelated with the rollout of the WJC centers within their same province.

1.4.4 School Level Specification

Lastly, using the same identification strategy, we study the overall effect of WJC centers on education outcomes at the school level by using the following regression equation:

$$Y_{st} = \beta_0 + \beta_1 WJC_{st} + \alpha_s + \lambda_{pt} + \gamma_t X'_s + \varepsilon_{st} \quad (1.3)$$

where (Y_{st}) is the education outcome (i.e. total number of children enrolled, standardized test scores) in school s at year t , (WJC_{st}) is an indicator variable that takes the value of one if the school has a WJC center within 1km/in the district of the school, (α_s) is a school fixed-effect, (λ_{pt}) is a province-by-year fixed-effect, $(\gamma_t X'_s)$ is a year-interacted vector of school's initial characteristics (including initial school enrollment, presence of electricity, presence of piped water, school language (Spanish), urbanization and public school dummy) and (ε_{st}) is a random error term. The inclusion of school fixed-effects accounts for any time-invariant characteristics

at the school level. We also allow the year fixed-effects to differ by province and by measures of school's baseline enrollment and baseline infrastructure. Since initially-different schools might be more likely to change differently, this empirical specification focuses on comparing changes in treatment and control schools with similar initial characteristics that might drive WJC center placement.

The coefficient of interest is (β_1) , which captures the average change in enrollment in schools that are located near the WJC centers or in districts with WJC center, to the average change in enrollment in schools that did not have a WJC center. The identification assumption is that treatment schools located in the proximity of a WJC center/in districts with WJC center would otherwise have changed similarly, on average, to those controls schools that are not exposed to the services of a WJC center. In practice, by controlling for province-by-year fixed-effects (λ_{pt}) and by variables that drive WJC center placement, the identification assumption is that treatment schools would otherwise have changed similarly, on average, to control schools within their same province and with similar initial characteristics. Throughout this analysis, we cluster our standard errors at the school level. We also estimate this regression including district-specific time trends.

Nevertheless, we are concerned about the possibility that the results are driven by time-varying variables which might influence both the opening of the WJC centers and school enrollment. A related issue is the possibility that WJC center managers consciously decide to introduce centers where school enrollment is increasing. To address both of these issues, we use the panel nature of the school data in order to construct a placebo treatment based on the timing of the WJC centers introduction. We estimate whether *future* WJC centers predict current enrollment using equation 1.4 below:

$$Y_{st} = \beta_0 + \beta_1 WJC_{st} + \beta_2 WJC_{st+1} + \beta_3 WJC_{st+2} + \beta_4 WJC_{st+3} + \alpha_s + \lambda_{pt} + \gamma_t X'_s + \varepsilon_{st} \quad (1.4)$$

where (WJC_{st+1}) , (WJC_{st+2}) and (WJC_{st+3}) are indicator variables that takes the value of one if the school has a WJC center within 1km/in the district of the school starting from the year $t + 1$, $t + 2$ and $t + 3$. If $\beta_2 > 0$, $\beta_3 > 0$ and $\beta_4 > 0$ are positive and significant, this would indicate that WJC centers are being introduced in areas where schooling is increasing more rapidly. While, if $\beta_2 = \beta_3 = \beta_4 = 0$ this would indicate that WJC centers are introduced in

areas in which school enrollment is growing for other reasons.²⁴ Therefore, the coefficients β_2 , β_3 and β_4 effectively capture the effect of future openings for areas that are not covered by the WJC centers in t . Our hypothesis for the placebo regression is that total enrollment in schools that do not have a WJC center within 1km/in the district should *not* be affected by the fact that a WJC center may open in the future in the proximity of these schools.

1.5 Results

1.5.1 Impact of WJC Centers on Gender-Based Violence

We begin by investigation the impact of WJC center's introduction on the incidence of gender-based violence against women. From estimating equation 1.1 for the sample of women, Table 1.5 presents the results of regressing the likelihood of experiencing domestic violence (by the intimate partner) in the last 12 months against the presence of a WJC center within 1km/in the district after controlling for several covariates, district fixed-effects, district-specific time trends and province-by-year fixed effects.

Panel A of Table 1.5 shows our domestic violence estimates when exposure to the program is measured through the presence of a WJC center within a 1km Euclidean buffer. Column 1 presents our results using the entire sample of women.²⁵ Introducing a WJC center within 1km of the women's residence decreases domestic violence by 2.2 percentage points, which represents a 5.6% decrease in domestic violence. Column 2 shows this regression after including district-specific trends to address the concern that districts that have a WJC center are trending differently than those that do not. This coefficient is slightly smaller (1.8 percentage points) but still significant. Our preferred specification is shown in Columns 3, in which we limit the sample to just urban clusters, which means that control areas are most comparable to those which are affected by the introduction of a WJC center. Even though this reduces the sample significantly, the coefficient is a bit higher in magnitude to the overall sample (2.9 percentage points) and highly significant. Lastly, Column 4 limits further to areas that ever have a WJC

²⁴This technique has already been used to address this concern by [La Ferrara, Chong and Duryea \(2012\)](#) and [Oster and Steinberg \(2013\)](#).

²⁵The full sample of women in the *Peruvian DHS* surveys consists of 210,847 respondents aged 15 to 49 over the period 2000-2014. However, this sample is reduced to 121,404 eligible women since we only include partnered women who are eligible for the domestic violence module. When we run estimations using the geo-coded cluster locations during the period 2006-2014, this sample is reduced even further to 64,366 observations of women.

center (including those that change and those that always have a center). The coefficient in this case is still negative and similar in magnitude but not statistically significant, which may be due to the sample size restriction.

In Panel B of Table 1.5 we explore the impact of WJC centers on domestic violence by using the alternative measure of exposure, presence of a WJC center in the district. We use this alternative explanatory variable as a robustness check and also to explore whether the opening of a WJC center matters in broader surroundings. These findings also show that women living in a district with a WJC center are significantly less likely to suffer from physical violence by their spouse compared to those living in districts that do not have this type of institution. The magnitude of the coefficients is relatively similar to the ones of Panel A. These results are robust to including district specific trends and to limiting the sample to urban clusters and districts which ever have a WJC center. In Table 1.A.4 of the Appendix, we also show that these results are driven by older and more educated women, which are the ones that are more likely to have better outside options.

Moreover, in Table 1.6 we present the impact of the WJC centers on different types of emotional violence. In general, we find a negative but not statistically significant effect except for one mistreatment emotional behaviour outcome. For instance, we find that proximity to a WJC center can be associated to less likely threats of taking the children from the spouse.

One limitation of the *Peruvian DHS* data collected on domestic violence is that it is self-reported by women respondents and, therefore, subject to recall bias, cultural values and willingness to report domestic violence. Since empirical work on gender-based violence generally suffers from measurement issues, in order to corroborate our results, we also use administrative district level data on femicides and female deaths due to aggression as alternative outcomes of violence against women. Tables 1.7 and 1.8 present the results of regressing the logarithm of femicides and female deaths due to aggression against the number of WJC centers in the district, respectively (equation 1.2). These findings provide suggestive evidence of a reduction in femicides and female mortality due to aggression. More precisely, the coefficients indicate that the opening of a WJC center in the district can be associated with a statistically significant reduction in femicides and female hospitalizations for assault. The largest effect is found for women aged 20 to 39 years old, which is reassuring in terms of the results found with the self-reported domestic violence data.

We also explore whether an improvement in access to justice for women has an impact on their health. Table 1.9 shows the effect of introducing a WJC center in the proximity of the residence on a set of women's health outcomes. In particular, women living within 1km of the WJC center experience an increase in their weight compared to those living further away. Finally, Table 1.10 shows the effects on female hospitalizations due to mental health problems using district and year variation in the openings. We find that after the opening of a WJC in the district, women mental health problems decline by 10% over the period of analysis. Moreover, we do not find these effects for men. These results show some suggestive evidence of an improvement in women's health.

1.5.2 Impact of WJC Centers on Children's School Attendance

Given the reduction on gender-based violence, in this section we analyze whether there are positive spillover effects on children's outcomes. We start by analyzing the impact of WJC centers on children's school attendance rate and their attendance status since a downside of our school-level data is that we observe number of students enrolled, not enrollment rates. Tables 1.11 and 1.12 summarize the estimated impacts of WJC centers on children's school attendance in primary and secondary level, respectively, from estimating equation 1.1 for the sample of children. While, Table 1.13 presents the results for children's attendance status (e.g. passed grade, repeated, dropped-out).

First, Panel A of Table 1.11 indicates that children in primary school living in household's located near a WJC center are significantly more likely to attend school. More specifically, living in the proximity of a WJC center increases children's school attendance by approximately 2 percentage points. Focusing on our preferred specifications in Columns 3 and 4, we find a positive and statistically significant effect on children's primary school attendance after the opening of a WJC center in the proximity of the household and also in the district of residence. These results are robust to using the different measures of exposure to the program. The magnitude of the findings in Table 1.11 could be considered very large given the primary school attendance rate of 97%. In order to better interpret these results, in Table 1.A.5 of the Appendix we analyze domestic violence in the household by children's primary level school attendance status. Effectively, we find that domestic abuse is higher among the 3% of households who do not send their children to primary school and this difference is driven by

urban areas. In addition, we also analyze the impact of WJC centers by the distribution of the primary school attendance. Information on primary school attendance is used to assign children into four distinct school attendance quintiles. Results in Table 1.A.6 of the Appendix indicate that the effect of opening a WJC center within 1km of a child's residence on primary school attendance is only statistically significant for those children located in areas with the lowest school attendance rates.

Second, in Table 1.12 we also find a positive and statistically significant impact of WJC centers on secondary school attendance for those children living within 1km of the center. These estimates range between 2 to 3 percentage points. However, this effect is no longer significant when we use presence of a WJC center in the district as a measure of exposure. Due to the GPS displacement issue in the Peruvian DHS data, we also estimate the impact of WJC centers using two additional Euclidean buffers: 3km and 5km. Results in Tables 1.A.7 and 1.A.8 show that when we analyze the effect of the WJC in broader surroundings we do not find a significant impact for both primary and secondary school attendance rates.

Lastly, the impact of WJC centers on school attendance status - grade advancement conditional on staying in school, repeating grade, recent drop-out and old drop-out is also estimated using the same method as reported for school attendance. Results in Table 1.13 show that children located near a WJC center are significantly more likely to pass a grade and they are also significantly less likely to drop out of school. However, we do not find an effect on grade repetition nor on having left school more than two years before the opening of the centers. These results are robust to using different samples of children (i.e. children of the women selected for the domestic violence module).

What we find, overall, is that investments in children's human capital, especially those in primary level, are affected positively by the introduction of the WJC centers.

1.5.3 Impact of WJC Centers on School Enrollment

The evidence above suggests that overall primary school attendance increases in response to WJC center introduction. This section analyzes our estimates of the impact of the WJC centers on education outcomes at the school level as an additional robustness check. From estimating equation 1.3, Tables 1.14 and 1.15 present estimated impacts of WJC centers on average enrollment in primary schools and secondary schools, respectively. While, Table 1.16

presents the impact of WJC centers on standardized test scores for second grade students in primary level.

Panel A of Table 1.14 shows our primary school enrollment estimates when exposure to the program is measured through the presence of a WJC center within a 1km Euclidean buffer. The coefficient on WJC center within 1km in Column 1 is positive and statistically significant. This result indicates that the introduction of a WJC center within 1km of a school is associated with an increase of 2.8% in the number of children enrolled in primary school in the year after the center was opened. The coefficient in Column 2, after controlling for district-specific trends, is almost unchanged (2.7%) and still highly significant. In Column 3, we include district population as a time-varying control in order to rule out the concern that our results might be driven by mechanical changes in population, especially due to the fact that our school data measure number of students enrolled, not enrollment rates. After controlling for district population, the impact of WJC centers on primary school enrollment is even larger (3.3%) and statistically significant. Our preferred specifications are shown in Columns 4 and 5, in which we limit the sample to just urban schools and districts that ever have a WJC center. Although this restricts the sample significantly, the coefficient for urban schools in Column 4 is also larger in magnitude to the overall sample (3.2%) and highly significant. Lastly, the impact for districts that ever have a WJC center is a bit smaller in magnitude (2.4%) and significant, despite the fact that we restrict the sample size even further.

In Panel B of Table 1.14 we explore the impact of WJC centers on primary school enrollment by using the alternative measure of exposure, presence of a WJC center in the district. Panel B shows that introducing a WJC center in the district also has a positive and significant effect, but the coefficient is a bit lower (1%), indicating that the effect probably decreases with distance. Focusing on our preferred specifications in Columns 4 and 5, we find that adding a WJC center in the district increases the total number of children in primary school between 1.2% and 1.9%. These results are also similar in magnitude to the results found with the individual-level data, which is reassuring.

Table 1.15 shows the impact of WJC centers on secondary school enrollment, using the different measures of exposure to the program. We also find a positive impact on the number of children enrolled in secondary school (2.9%) when we use the entire sample, but the effect is not robust to controlling for district specific trends and to limiting the sample to districts

which ever have WJC center. The specification with urban schools is the only one that yields a positive and significant coefficient of 3.4% for secondary school enrollment.

Lastly, consistent with these results, we also find some suggestive evidence of a positive effect on standardized test scores for primary school children located in schools near a WJC center. Table 1.16 shows that test scores of children in schools located in the proximity of a WJC center are 0.02 - 0.05 standard deviations higher. Even though these results are not robust to all the different specifications, they are positive and highly significant for urban schools.

All these findings suggest a strong connection between the presence of WJC centers and total number of children in primary school. They also indicate that these findings are localized to within few kilometers and they are mostly driven by urban areas. In Table 1.A.3 of the Appendix, we also show these effects broken down by gender and grade. We find that these effects are similar for boys and girls, even though they seem to be driven mostly by girls. We also find that the impact is equally distributed among the different grades.

Placebo regression: *Future* WJC centers

As mentioned earlier, one of the main threats to this identification strategy is the possibility that WJC centers were rolled out in response to changes in enrollment, rather than causing them. This is strongly linked to the issue of endogenous WJC center placement. Even though, we account for characteristics which are constant over time through school fixed-effects, one concern that remains is the possibility that WJC centers are placed in areas where enrollment is increasing more rapidly since center managers or policymakers are targeting more densely populated areas. Another concern is posed by time-varying unobservables correlated to both the timing of the WJC centers and school enrollment. For example, it could be that areas reached by the WJC centers are also hit by a positive economic shock or there are improvements in public welfare programs at the time they are opening the WJC centers. We already account for this by controlling for province-by-year fixed effects.

However, another way to address the concern that WJC centers are located in areas that are changing in other ways that we do not observe is by constructing a placebo treatment based on the timing of the WJC center openings. We estimate analogous regressions to the ones in Tables 1.14 and 1.15 (our baseline school-level specification), but instead of only looking at the effects of opening a WJC center on current enrollment, we also look at the effects of *future* openings.

The idea is that if *future* WJC center openings predict current enrollment, this would suggest that WJC center placement anticipates changes in schooling, rather than causing them. Table 1.A.10 and 1.A.11 show the results for this falsification exercise for primary and secondary school enrollment, respectively. We find that the effect of future WJC centers is virtually zero and not statistically precise, suggesting no strong evidence of pre-trends. In addition, the inclusion of future WJC centers does not affect our estimate of the impact of current WJC center on school enrollment.

1.6 Discussion: Mechanisms

In this section we provide some evidence on the mechanisms that might potentially drive the findings in this study. In the context of Peru, the presence of the WJC centers can reduce domestic violence either by improving women's intra-household bargaining power, or by increasing women's trust in the institutional system.

Firstly, we argue that the presence of a WJC center in proximity of the household may allow women to send a signal to their husbands regarding their outside options. The availability of easier access to justice thus may generate a more credible threat to the potential offenders through greater chances of demanding police involvement and criminal penalties. Several economic theories of household bargaining power suggest that policies aimed at affecting spouse's outside option from an abusive relationship may also affect within-household distribution through changes in their relative bargaining positions (Farmer and Tiefenthaler, 1996; McElroy and Horney, 1981; Manser and Brown, 1980). In other words, women's threat point increases when they have alternatives and when support services are more helpful.

Central to this analysis is the relationship between household decision-making or bargaining power in the household and the WJC center introduction. In order to test this, we use the *Peruvian DHS* which records who has the final say on a variety of household decisions. For example, a woman is asked "*who makes the final decision on large household purchases?*" or "*who makes the final decision on money husband earns?*". Responses include: respondent only, jointly with partner and partner only. For these categories, we construct three measure of equal decision-making. The first one is a score that ranges from 0 to 6 and counts the number of times the respondent makes decision jointly with partner. The second one is also a score that ranges from 0 to 1 and counts the share of decisions made jointly with partner. The third

one is a dummy that takes the value of 1 when at least one decision is made jointly with the partner. In addition to decision-making, we also estimate the effect of WJC centers on women's earnings relative to their husbands.

Table 1.17 provides the estimates of the impact of WJC centers on decision-making and bargaining power. We find suggestive evidence of an improvement in the bargaining power of women in the household. In particular, we find that women living near a WJC center are more likely to make joint decision-making with their husband. They are also less likely to earn less than their husband and more likely to earn as much as their husband. We also analyze whether the WJC centers have an effect on women's labor force participation. Results in Table 1.18 indicate that women's labor supply does not seem to be affected by the opening of these centers.²⁶ This result may seem a bit puzzling but one possible interpretation is that WJC centers might allow women to be more productive in their jobs by supplying more hours and therefore increasing their income.

Alternatively, the introduction of WJC centers may have contributed to break the silence regarding violence against women and to turn it into a public issue. The different awareness campaigns by the "justice promoters" may have contributed to convert these centers into a trustworthy public service that ensures protection and support to women. By increasing women's trust in the institutional system, these specialized institutions may be changing the discriminatory social values and power structures that underlie violence against women in Peru.

For this purpose, we analyze patterns of institutional trust in case of a domestic violence event. We focus on relating proximity to a WJC center to the type of institutional help sought by women in case of having suffered from domestic violence abuse. More specifically, we analyze nine institutional trust variables: whether women sought help in a police station, justice court, district attorney office, DEMUNA (Defence center for children and adolescents), WJC center, ombudsman office, health facility, NGO and other institution. Results in Table 1.19 indicate that married or cohabiting women living within 1km of a WJC center are less likely to seek help in a regular police station, but instead they are significantly more likely to put their trust in the women's justice centers. In particular, when we analyze the impact of WJC centers on the overall help sought for the entire sample of women, we find that women in the proximity

²⁶In addition, we find that WJC center have no effect on civil status and fertility outcomes, suggesting that there is no selection into the domestic violence module (results upon request).

of a WJC center are significantly less likely to seek help. However, when we analyze the same effect conditional on the occurrence of a domestic violence event, the effect is still negative but not significant. This negative effect could reflect the drop in domestic violence in Peru during the period studied and thus the drop in institutional help sought by women. Despite this, we also find that proximity to a WJC center has a positive and statistically significant effect on women's trust to seek help in a WJC center for the conditional sample. This result suggests that institutional trust is reinforced in the area close to the WJC center, which might trigger a potential female empowerment. This finding has important implications for state capacity, since it may enhance government's ability to fight the problem of domestic violence.

Both mechanisms lead to the conclusion that WJC center's intervention in households with abuse may change the behavior of offenders and victims. In other words, the opening of WJC centers might be a powerful tool to reduce the incentives of the spouse to choose domestic violence through an improvement of the bargaining power of women in the household or/and an increase in institutional trust which, in turn, might also improve women's health and their ability to take care of their children.

Furthermore, our paper shows that the impact of WJC centers is not limited to the direct recipients of domestic abuse. We find strong evidence of positive externalities in terms of their children's human capital investment. In the context of Peru, empowering and supporting women through the presence of WJC centers can contribute to children's education outcomes by basically improving women's welfare. Previous research has shown that women affected by domestic violence may be limited in their ability to take care of their children. Children growing up in households where there is violence among intimate partners can suffer from behavioural and emotional problems, which may lead them to drop out of school and even engage in child labor.

In the context of Peru, if a woman suffers from domestic violence and thus probably from health problems, the burden of household chores falls automatically on their daughters since sons are not expected to get involved in such activities. To better understand why empowering women would help promote school attendance in the context of Peru, we analyze the impact of WJC center on child labor. Table 1.20 reports regression results of the impact of WJC centers on child labor for children aged 6 to 14 years old. These findings show that proximity to a WJC center is associated with a statistically significant reduction in child labor, especially for young

girls.

To shed light on the mechanisms behind the schooling effects on children we divide the sample between households where the grandmother was subject to domestic violence by the grandfather. Previous literature suggests that having a mother subject to violence make women more prone to be subject to violence in their own household. We find that most of the effects are driven by these type of households showing that the effect is having an effect on children by affecting most vulnerable households that would be subject to violence (see Table 1.A.9 in the Appendix).

Finally, having shown the inter-generational benefits of improving access to justice through the WJC centers, we also analyze how other government programs in Peru could interact with WJC centers by changing parental incentives in the affected areas. For this purpose, we exploit the gradual rollout of the conditional cash transfer program (CCT *Juntos*), which provides monetary transfers to parents with the condition that children attend school on a daily basis. In particular, we examine how the presence of CCTs in the district interacts with the presence of a WJC center within 1km of the school/child's residence.

Tables 1.21 and 1.22 present analogous versions of Equations 1.1 and 1.3 for children and schools of primary level including the interaction with CCT_{dt} and a dummy indicating whether district d has CCT in year t , respectively. Using the *Peruvian DHS*, we do not find any synergies between better access to justice and CCTs. This result is probably due to the lack of variation in CCT rollout in the Peruvian DHS since we are only analyzing the 2006-2014 period. We even find a negative correlation between CCT in the district and primary school attendance when we further limit the sample to districts which ever have a WJC center. In this case, the CCT rollout is most likely representing the most marginalized districts of Peru in terms of access to education and not necessarily the CCT program. However, using the school census data which covers the universe of schools in Peru, we find that conditional cash transfers can enhance the positive effect of WJC centers on primary school enrollment. We find that the effects of CCTs are larger once a WJC opens in the district.

1.7 Robustness Checks

1.7.1 Assessing the Internal Validity of the Research Design

In this section we present several robustness checks that support the validity of the identification assumption of the paper. Identification using the difference-in-difference approach relies on the assumption that the path of the outcome variables for the treatment and control households/schools should not be systematically different in the absence of the WJC centers. More precisely, this means that the introduction of WJC centers should be the only factor that drives the treatment group to experience a change in an outcome variable, such as a relative reduction in domestic violence.

However, the main threat to this identification strategy is the correlation between the order of the opening of the WJC centers and the trends in domestic violence and education patterns before the rollout of the program. Basically, the average effect of the WJC centers would be biased if the timing of WJC centers creation was correlated with pre-program changes in domestic violence and education outcomes. To address this concern, we first argue that pre-program changes in domestic violence and education patterns are uncorrelated to the timing of future WJC centers introduction. Second, we conduct an event study to show that pre-program trends are not driving our results. We also use this analysis to provide a sense of the dynamic effect of WJC centers.

In order to test this, we begin by estimating a regression of pre-program changes in school enrollment on indicators for the year the WJC center was introduced within a 1km radius of the school:

$$\Delta \text{Log}(Y_{st}) = \text{Log}(Y_{st-1}) - \text{Log}(Y_{st}) = \gamma + \alpha_t + \sum_{k \geq t} \delta_k I(\text{WJCyear}_{<1km,s} = k) + \varepsilon_{st} \quad (1.5)$$

The dependent variable is ΔY_{st} is the change in education outcomes at the school level from year $t - 1$ to year t (e.g change in the log of primary/secondary total school enrollment, change in school test scores). The set of dummy variables ($\text{WJCyear}_{<1km} = k$) take the value of 1 in the year in which a WJC center was opened within 1km of the school. Year fixed-effects are denoted as α_t . The data for this test is derived exclusively from the School Census (CE) panel database and the sample is restricted to those schools that were reached by the program between 2006 and 2014. The reference group is the opening of a WJC center in 2006. If (WJCyear) effects are

jointly significant it would indicate that year of WJC center creation within 1km of the school was correlated with pre-program changes in total school enrollment.

Unfortunately, we cannot perform exactly the same test with the *Peruvian DHS* since we do not observe the same clusters of households over time. This means that we cannot exploit the variation generated by proximity to the WJC center through Euclidean buffers. However, we can still check whether the timing of WJC center introduction in the district is correlated with changes in women's domestic violence and children's school attendance rates in the district. For this case, we regress pre-program changes in the outcomes of interest for women and children at the district level (e.g. domestic violence, primary school attendance rate, secondary school attendance rate) on yearly indicators of WJC center introduction in the district:

$$\Delta y_{dt} = y_{dt-1} - y_{dt} = \gamma + \alpha_t + \sum_{k \geq t} \delta_k I(WJCyear_d = k) + \varepsilon_{dt} \quad (1.6)$$

In Tables 1.A.12 and 1.A.13 of the Appendix, we report the results of estimating Equation (1.5) and (1.6) on three different windows of pre-program changes in education outcomes at the school and district level, respectively. These findings show that pre-program changes in education at the beginning of the rollout might be correlated with the timing of the WJC center introduction. While, the other two windows of pre-program education results indicate that the rollout year is not correlated with pre-program changes in education outcomes. For this reason, we decide to focus our analysis in the middle of the rollout, that is, from 2006 till 2014, for which identifying assumptions are likely to hold.

We do not find evidence that pre-program trends in education patterns are correlated with the order of the WJC center implementation during the period 2006-2014. In particular, results in Table 1.A.12 indicate that opening a WJC center within 1km of the school does not significantly explain pre-program changes in primary and secondary school enrollment between 1998-2005. Similarly, results in Table 1.A.13 show that the opening of a WJC center in the district is not correlated to pre-program changes in district school attendance rates between 1996-2005. Results in Table 1.A.14 also indicates that pre-program changes in standardized test scores at the school level are not correlated with WJC center introduction. In all cases, we are unable to reject the null hypothesis of the joint test. These findings strongly suggest that pre-program time trends for the education outcomes of interest are not correlated with the introduction of the WJC centers between 2006-2014.

Moreover, Table 1.A.15 reports the results of estimating Equation 1.6 using women's self-reported domestic violence as an outcome variable. Column 1 shows that the timing of WJC centers in the district is not significantly correlated with pre-program changes in district level domestic violence and the p-value for the joint test is 0.416. The lack of a significant correlation between the year of WJC center introduction in the district and changes in district level domestic violence for different windows provides evidence that pre-program time trends in domestic violence were not correlated with introduction of WJC centers in the district.²⁷

In conclusion, we have presented evidence that pre-program changes in domestic violence and education patterns are uncorrelated to the timing of future WJC centers introduction in the district/within 1km. The pre-program patterns for each relevant outcome of interest are also depicted by Figure 1.9.

1.7.2 Accounting for the dynamic impact of WJC centers

Next, we exploit the fact that we have information prior to the introduction of the WJC, since the rollout was done gradually each year, in order to conduct an additional formal test of whether pre-trends in the outcomes of interest are correlated with the opening of the WJCs in Peru. This test also allows us to better understand the dynamics of the WJC center introduction and to disentangle the effect over time. For instance, how quickly school enrollment or attendance rates increase after the opening of a WJC and whether this impact accelerates, stabilizes or mean reverts. To explore these dynamics, we conduct an event study analysis, where we analyze the impact of leads and lags of the WJC introduction. Formally, we estimate the following regressions at the individual and school level, respectively:

$$y_{idt} = \gamma_0 + \sum_{i=-3}^4 WJC_d * \beta_i I(\tau_{dt} = i) + \alpha_d + \lambda_{pt} + \delta X'_{it} + \varepsilon_{idt} \quad (1.7)$$

$$Y_{st} = \beta_0 + \sum_{i=-3}^4 WJC_s * \beta_i I(\tau_{st} = i) + \alpha_s + \lambda_{pt} + \gamma X'_{st} + \varepsilon_{st} \quad (1.8)$$

where τ_t denotes the event year, defined so that $\tau = 0$ for the year the WJC was introduced within 1km/in the district of the household i /school s , $\tau = 1$ for one year after the WJC centers

²⁷Unfortunately, we are unable to perform this test for other women's outcomes due to data unavailability for the pre-program period. For instance, official data on femicides in Peru started to be recorded since 2009 and female hospitalizations since 2006/7.

began to operate, and so on. For $\tau \leq -1$, school and households were untreated by the WJC introduction. The coefficients are measured relative to the omitted coefficient $\tau = -1$. In other words, we add indicator variables for up to 3 years before implementation and 0 - 4 years after implementation.²⁸ For each outcome, we expect that coefficients on dummies for years -3 and -2 (the years prior to the WJC centers) should not be significant, because if this was the case, the validity of the parallel trends assumption would be violated.

Figures 1.10 and 1.11 plot the coefficient of the interaction for the years leading up to the opening of the WJC centers and the years after the introduction of the WJC centers from estimating Equation 1.7 for each of the women's and children's outcomes, respectively. Similarly, Figure 1.12 plots the coefficient of the interaction for the leads and lags of the WJC center introduction from estimating Equation 1.8 for the education outcomes at the school level.

For women's outcomes (e.g. domestic violence, emotional violence, female deaths due to aggression, joint decision-making and labor supply), the coefficients on the years leading up to the opening of the WJC centers are close to zero and not significant, showing no evidence of an anticipatory response within district about to introduce the WJC centers. In particular, we find that women located in districts with a WJC center present lower propensity to experience domestic violence since the year of the opening of the WJC center. This declining pattern reaches its largest impact 2 years after the opening of the center. A similar pattern is found for mental health, one year after hospital entries related to mental health problems decline by 20%.

For primary school enrollment and attendance, we find that the treated schools and households did exhibit a rising trend (relative to the control group) prior to the WJC implementation but this difference is not statistically significant. In particular, primary school attendance increases by 2 percentage points two years after the opening of a WJC in the district. This increase reaches its peak in the third year of WJC introduction, which is also accompanied by a significant improvement in schooling performance and by a decrease in drop-out rates in the same year.

Similarly, in the year of the opening, primary school enrollment increases substantially by 1.8% for schools located within 1km of the WJC, after which this increment fluctuates around 2% over the subsequent 3 years. For standardized test scores, there is also no difference in

²⁸Of these seven indicator variables, note that $\tau = -3$ is a dummy that takes the value one for more than 3 years before the WJC center was introduced. The next five dummies are equal to one only in the relevant year of WJC center opening, while the final variable $\tau = 4$ is equal to one in each year starting with the fourth year of adoption.

pre-program trends between school located near a WJC center and those further away. Indeed, the graphs show an absence of a strong pre-trend and evidence of a trend break after the WJC opened within 1km and in the district. For secondary school enrollment, we find that schools exposed to a WJC center have a lower enrollment and attendance prior to the opening of the WJC. However, this decline is opposite to the direction we observe after the WJC introduction and it is not statistically significant.

Overall, for school and households located in districts with a WJC presence, the greatest impact on primary school enrollment, primary school attendance and standardized test scores is found 2 years after the opening of the centers, which coincides with the negative impact on women's self-reported domestic violence. The similar timing of the effects on education and domestic violence provides further evidence that improving access to justice for women might be an important mechanism for allowing women to take better care of their children (i.e. investing in their human capital) by increasing their threat point in the intra-household bargaining power, their trust in the institutional system and also by improving their health.

1.8 Conclusion

In this paper we argue that the opening of WJC centers in Peru has a positive impact on women's status and their children's human capital investment, and that these impacts are concentrated in the very local areas around the WJC center. To the best of our knowledge, this is the first quantitative analysis that attempts to explore the impact of an unexamined dimension of institutional intervention that provides better access to justice for women, namely the WJC centers, on the prevalence of gender-based violence and education outcomes.

We deal with the potential endogeneity in the WJC center placement by exploiting the variation generated by the rollout of the women's justice centers in Peru. Basically, in order to ensure that our results are not driven by selection or time-varying unobservables, we use a difference in-differences strategy, which exploits variation created by the differential timing in the opening of the WJC centers and also the spatial variation in the exposure of a school/household to a WJC center, together with province-by-year fixed effects. We provide evidence in support of the identifying assumptions, and account for two key time-varying confounders: the fact that WJC center introduction might anticipate changes in schooling and unobservable changes in variables that might affect both the timing of the WJC centers and the

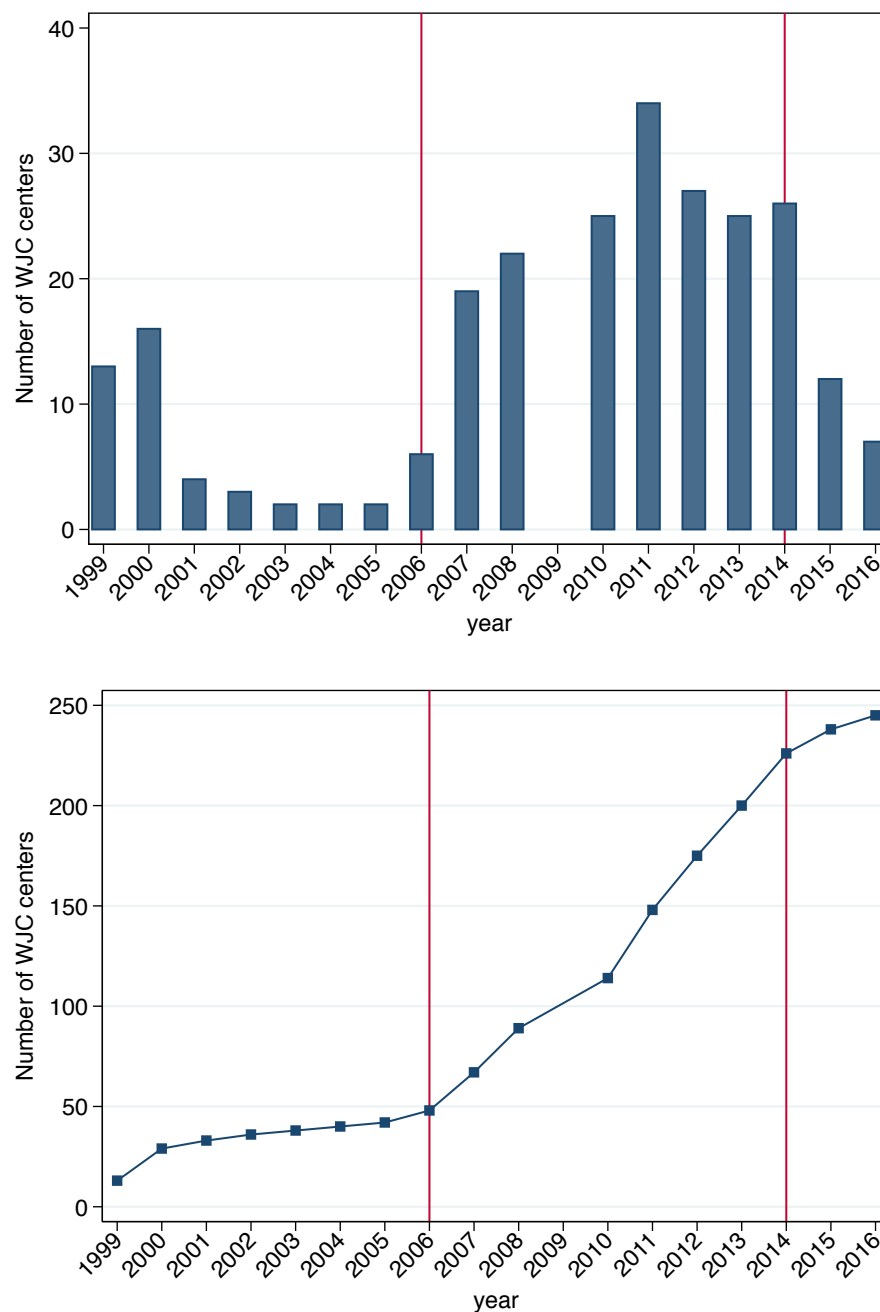
education outcomes.

Our main finding is that women's status and investments in children's human capital are affected positively by the introduction of the WJC centers. In particular, our results first reveal that providing better access to justice for women can reduce domestic violence, femicides and female deaths due to aggression and consequently improve their mental health. Then we find that children in primary school living in household's located near a WJC center are significantly more likely to attend school, to pass a grade and they are also significantly less likely to drop out of school. We also show that introducing a WJC center within 1km of a school causes an increase of 3% in the total number of children enrolled in primary schools, which reinforces our previous results. Moreover, primary school second graders have better test scores in reading and mathematics. Consistent with the results on education, we also find that young girls are less likely to be working after the opening of the WJC centers. Most of these effects are localized within a few kilometers and they are mostly driven by urban areas.

Lastly, we analyze which are the potential mechanisms that might be driving this pattern of results. First, we provide evidence that these improvements might be driven by an increase in the bargaining power of women inside the household. Next, we also find some evidence that after these centers opened women resort more to formal institutions in case of violence suggesting an increase in trust on state institutions which might also lead to an improvement in women's intra-household bargaining power. From a public policy standpoint, our analysis implies that providing better access to justice for women can be a powerful tool to reduce gender-based violence and increase human capital investment of children, suggesting a positive inter-generational benefit of the program.

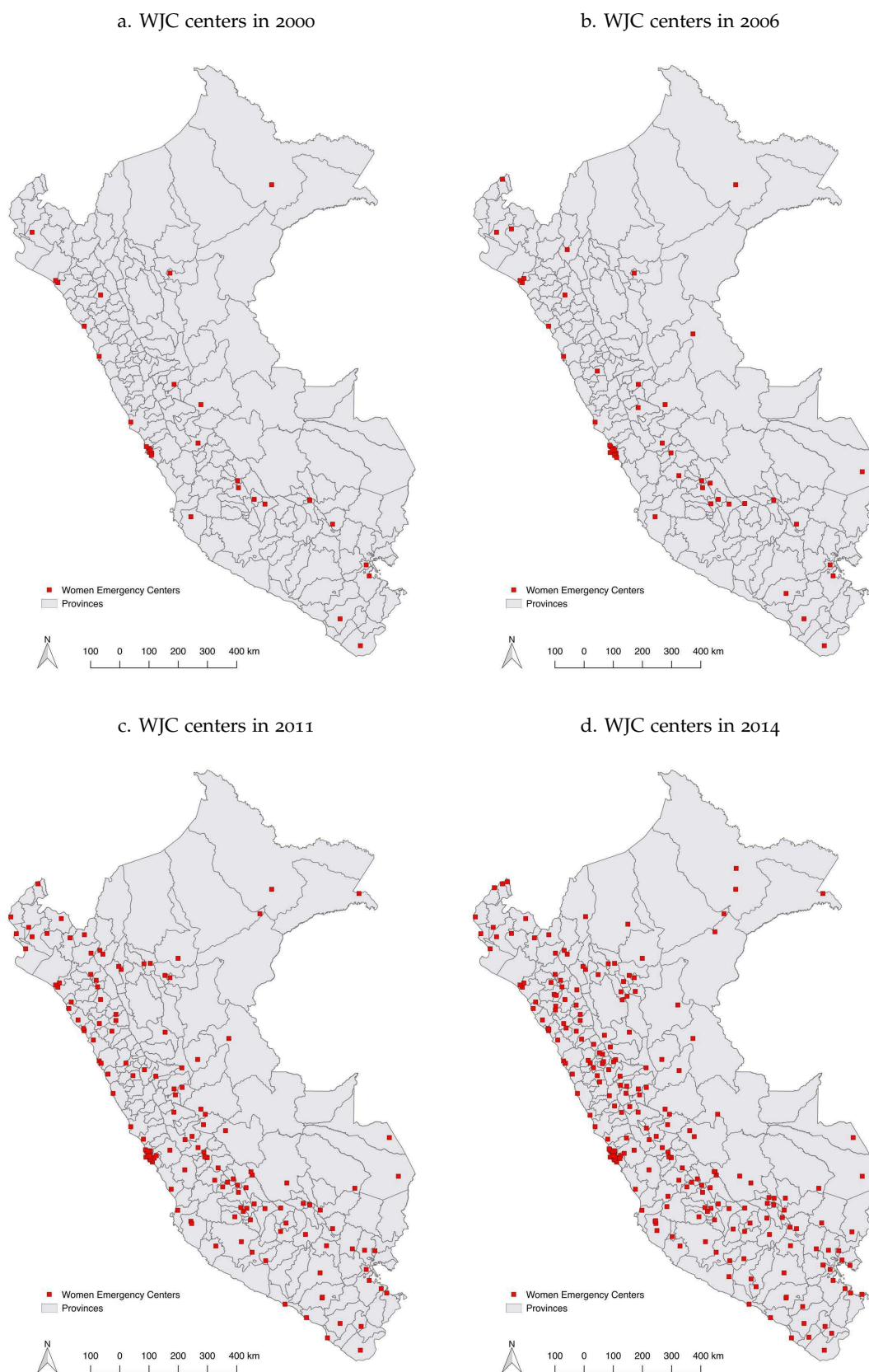
Chapter 1 - Figures and Tables

Figure 1.1: *Distribution and Growth of the Opening of the women's justice centers (WJCs) by Year - Peru (1999-2016)*



Notes: Author's estimates based on WJC centers data from the Peruvian Ministry for Women and Vulnerable Populations (MIMP).

Figure 1.2: Rollout of the WJCs across Time and Space (1999-2014)



Notes: Author's estimates based on WJC centers data from the Peruvian Ministry for Women and Vulnerable Populations (MIMP).

Figure 1.3: Total Number of Persons Attended in WJC Centers by Year (2002-2016)

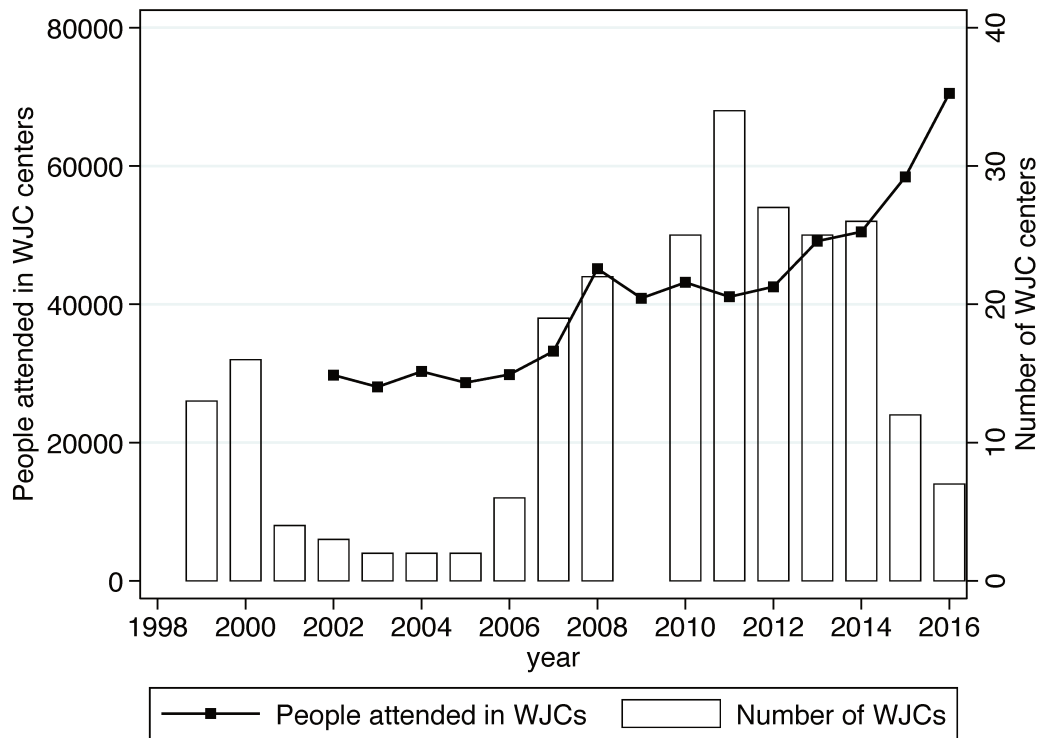
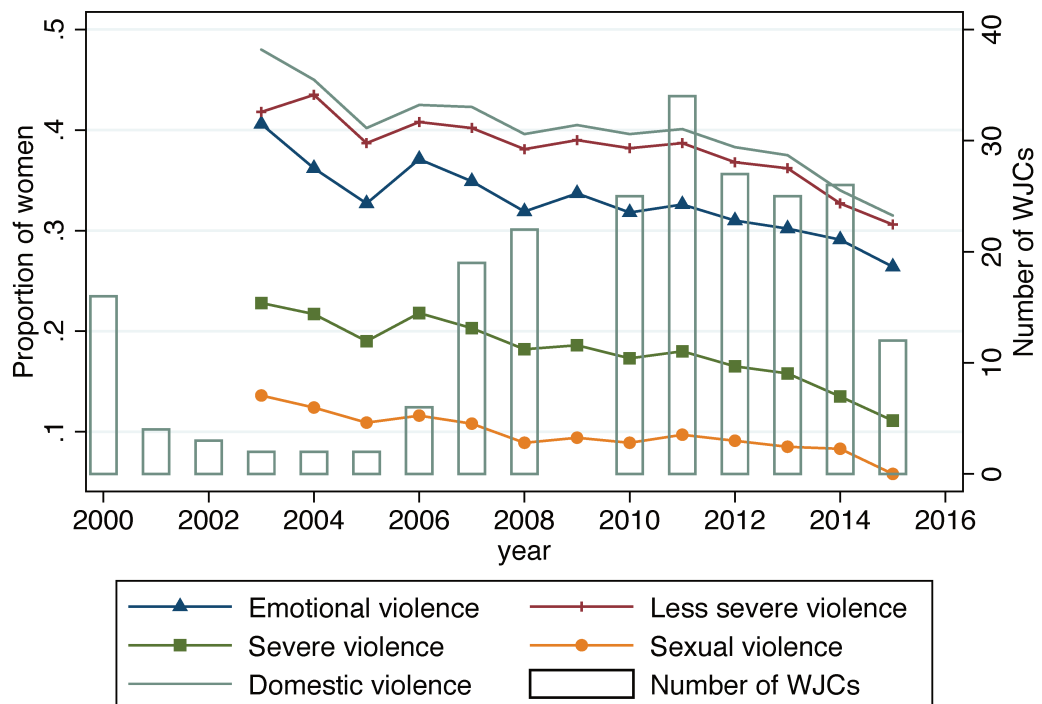


Figure 1.4: Domestic Violence in Peru (2003-2015)



Source: 2003-2015 Peru DHS

Figure 1.5: *Euclidean Distance Buffers and WJC centers (Schools and DHS Clusters of Households) - Lima and Tumbes*

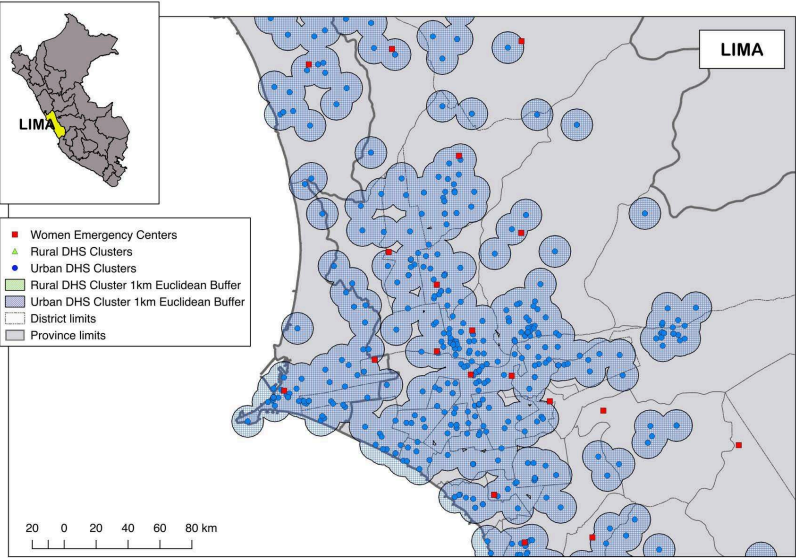
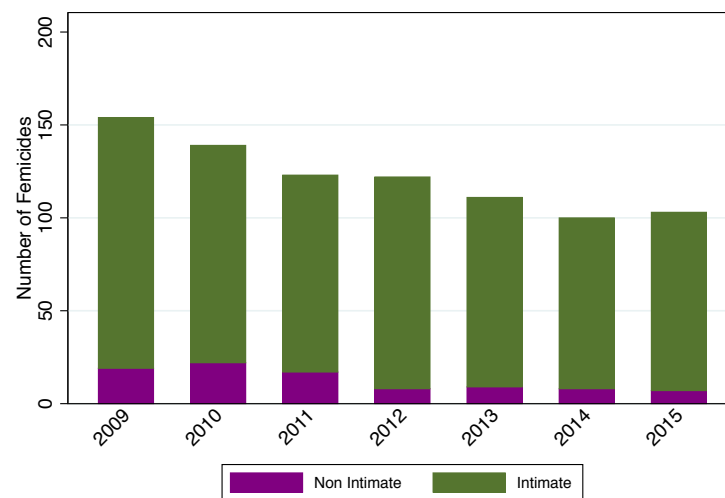
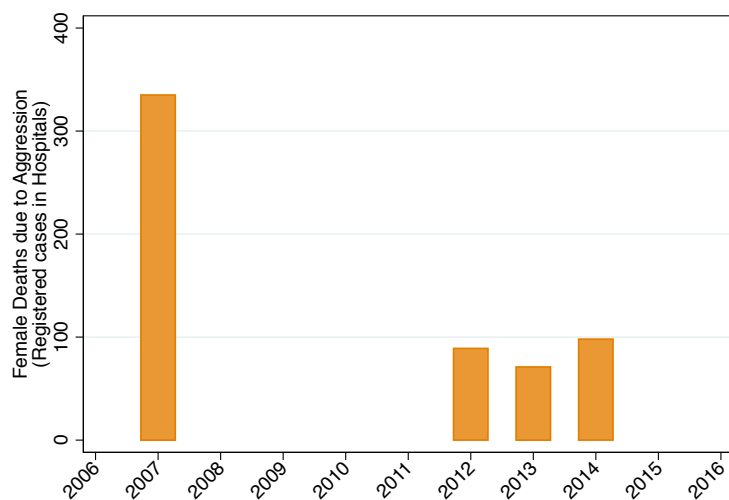


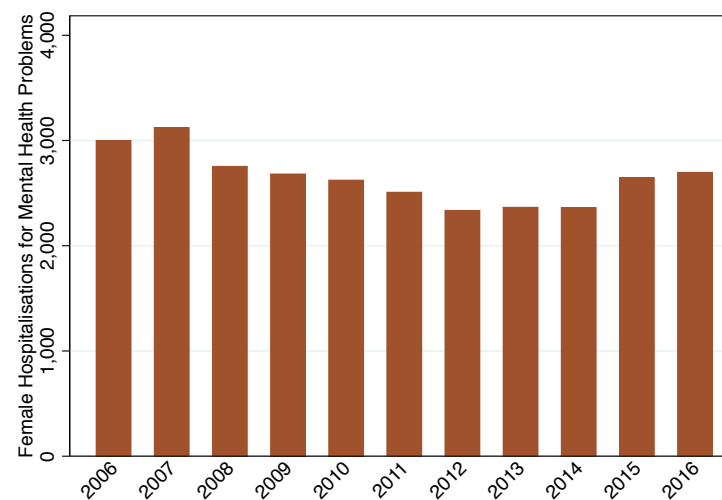
Figure 1.6: *Femicides, Female Deaths due to Aggression and Female Hospitalizations for Mental Health Problems - Peru*



Source: Peruvian Crime Observatory at the Ministry of Public Affairs



Source: Ministry of Health - General Office of Statistics and Informatics



Source: Ministry of Health - General Office of Statistics and Informatics

Notes: Author's estimates based on femicides data from the Peruvian Crime Observatory at the Ministry of Public Affairs, on female deaths data due to aggressions (cases registered in hospitals) and on female hospitalizations for mental health problems from the Peruvian Ministry of Health - General Office of Statistics and Informatics.

Figure 1.7: *WJC center and CCT Juntos presence in the district*

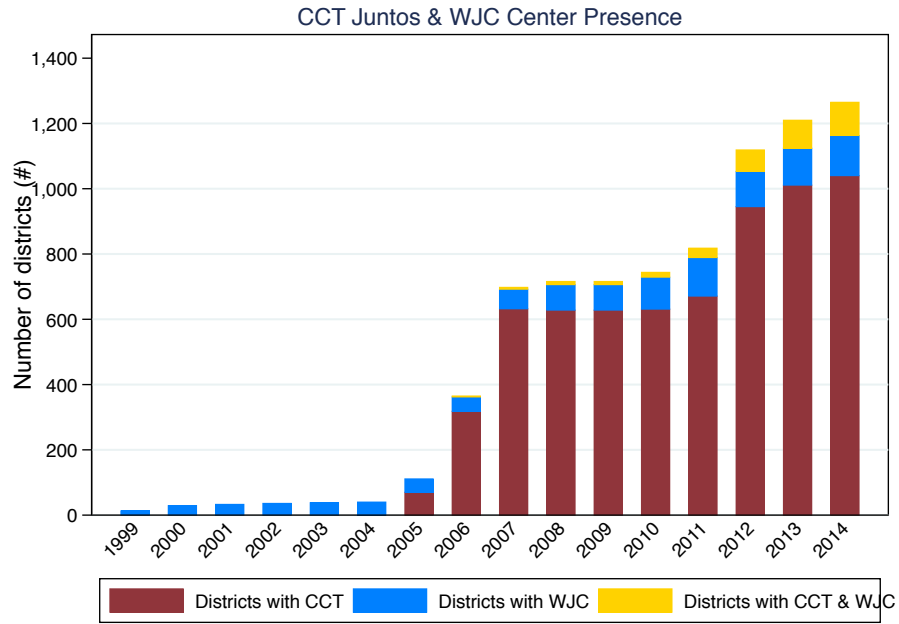


Figure 1.8: *WJC center and CCT Juntos entry in the district*

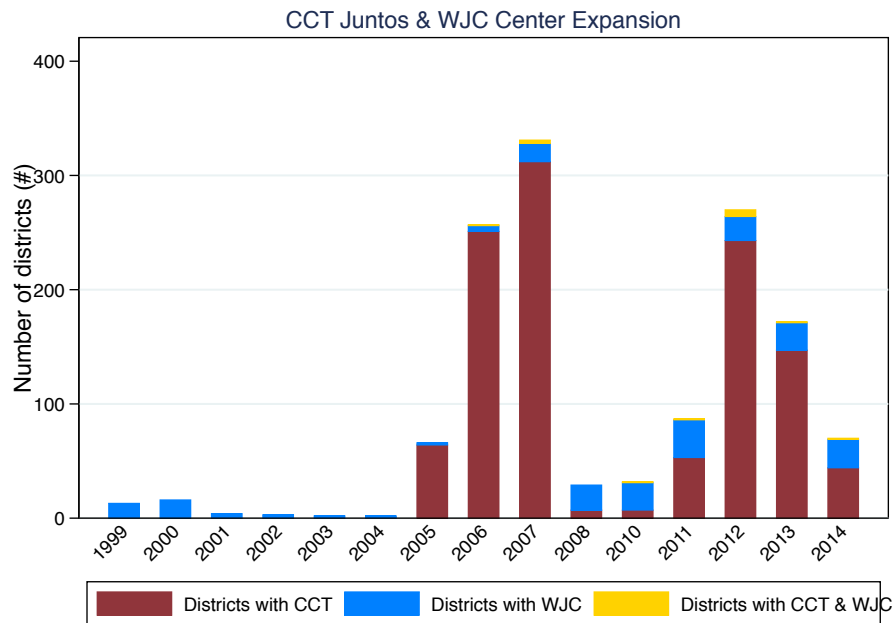


Table 1.1: Women's Descriptive Statistics and WJC Center Exposure - DHS (2006-2014)

	Women: 15-49 years old		
	All	Urban	Rural
<i>Panel A.1: Number of women by exposure to a WJC center</i>			
No WJC within 1km	55,323	29,432	25,891
WJC within 1km	9,040	8,965	75
No WJC within 5km	38,603	13,841	24,762
WJC within 5km	25,760	24,556	1,204
Total of women	64,363	38,397	25,966
<i>Panel A.2: Number of women by exposure to a WJC center</i>			
No WJC in the district	61,946	28,540	33,406
WJC in the district	34,614	30,041	4,573
Total of women	96,560	58,581	37,979
	Women: 15-49 years old		
	Obs	Mean	Std. Dev.
<i>Panel B: Women's Summary Statistics</i>			
Domestic violence last 12 months	64,363	0.390	0.487
Less severe violence	64,363	0.376	0.484
Severe violence	64,363	0.174	0.379
Sexual violence	64,363	0.093	0.291
Emotional violence	64,363	0.323	0.467
Anemic	57,540	0.220	0.414
Weight (kg)	59,460	61.57	11.10
BMI	59,460	26.80	4.416
Underweight	59,460	0.006	0.079
Overweight	59,460	0.511	0.499
Obese	59,460	0.208	0.406
Smokes	64,363	0.035	0.184
Age	64,363	33.93	8.336
Age at first marriage	64,363	20.14	4.739
# Total children ever born	64,363	2.811	1.993
# Years of education	64,363	8.577	4.481
# Household Members	64,363	4.626	1.818
Married	64,363	0.356	0.478
Living together	64,363	0.517	0.499
Widowed	64,363	0.007	0.089
Divorced/Not living together	64,363	0.118	0.319
Urban cluster	64,363	0.596	0.490
Currently working	64,363	0.684	0.464

Notes: The GPS data was not available for the years 2012 and 2013 in the Peru DHS. Source: Peru DHS (2006-2014)

Table 1.2: Children's Descriptive Statistics and WJC Center Exposure - DHS (2006-2014)

	Primary Level (Children: 6-11 years old)			Secondary Level (Children: 12-16 years old)		
	All	Urban	Rural	All	Urban	Rural
<i>Panel A.1: Number of children by exposure to a WJC center - (GPS data)</i>						
No WJC within 1km	42,914	19,654	23,260	29,494	14,282	15,212
WJC within 1km	5,789	5,740	49	4,025	3,991	34
No WJC within 5km	32,066	9,706	22,360	21,691	7,087	14,604
WJC within 5km	16,637	15,688	949	11,828	11,186	642
Total of children	48,703	25,394	23,309	33,519	18,273	15,246
<i>Panel A.2: Number of children by exposure to a WJC center - (All data)</i>						
No WJC in the district	48,895	19,250	29,645	33,392	13,999	19,393
WJC in the district	22,971	19,084	3,887	16,069	13,490	2,579
Total of children	71,866	38,334	33,532	49,461	27,489	21,972
	Primary Level (Children: 6-11 years old)			Secondary Level (Children: 12-16 years old)		
	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
<i>Panel B: Children's Summary Statistics</i>						
Currently Attending	48,703	0.970	0.169	33,519	0.895	0.305
Female Attendance	24,689	0.970	0.169	18,549	0.899	0.300
Male Attendance	24,014	0.970	0.169	14,970	0.891	0.311
Passed Grade	48,213	0.919	0.271	30,380	0.782	0.412
Repeated Grade	48,213	0.048	0.215	30,380	0.038	0.191
Dropped Out	48,213	0.022	0.146	30,380	0.090	0.287
Left School +2 years ago	48,213	0.002	0.047	30,380	0.084	0.278
Age	48,703	8.467	1.700	33,519	13.786	1.384
Head's Years of Education	48,703	8.602	7.159	33,519	8.348	7.025
Urban Cluster	48,703	0.521	0.499	33,519	0.545	0.497
# Female Adults in HH	48,703	1.219	0.532	33,519	1.218	0.541
# Male Adults in HH	48,703	1.101	0.611	33,519	1.120	0.669
# HH Members 0-18 years old	48,703	3.166	1.522	33,519	3.248	1.551

Notes: The GPS data was not available for the years 2012 and 2013 in the Peru DHS. Source: Peru DHS (2006-2014)

Table 1.3: School Descriptive Statistics and WJC Center Exposure - School Census (2006-2014)

	Primary Schools (1st - 6th Grade)			Secondary Schools (1st - 5th Grade)		
	All	Urban	Rural	All	Urban	Rural
<i>Panel A: Years of coverage and number of schools</i>						
Number of schools in						
First year of coverage (2006)	32,817	12,007	20,810	9,693	6,822	2,871
Last year of coverage (2014)	36,859	14,325	22,534	12,773	8,488	4,285
<i>Panel B: Number of schools by exposure to a WJC center</i>						
Never had WJC within 1km	34,372	11,883	22,489	11,287	7,018	4,269
WJC within 1km	2,575	2,524	51	1,522	1,504	18
Never had WJC within 5km	26,418	5,095	21,323	7,282	3,164	4,118
WJC within 5km	10,529	9,312	1,217	5,527	5,358	169
Total of schools	36,947	14,407	22,540	12,809	8,522	4,287
Never had WJC in the district	24,439	6,530	17,909	7,481	4,040	3,441
WJC in the district	12,555	7,884	4,671	5,330	4,484	846
Total of schools	36,994	14,414	22,580	12,811	8,524	4,287
	Primary Schools (1st - 6th Grade)			Secondary Schools (1st - 5th Grade)		
	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
<i>Panel C: School Summary Statistics</i>						
Total Enrollment	315,221	95.9	142.5	102,685	174.8	206.8
Female Enrollment	315,221	46.9	73.6	102,685	84.4	114.9
Male Enrollment	315,221	49.0	75.4	102,685	90.4	113.1
Public School	315,221	0.797	0.402	102,685	0.636	0.481
Urban School	315,221	0.378	0.485	102,685	0.679	0.466
School Language (Spanish)	315,221	0.815	0.387	102,685	0.905	0.292
School Language (Quechua)	315,221	0.124	0.330	102,685	0.000	0.242
School with electricity	315,221	0.671	0.469	102,685	0.872	0.334
Schools with piped water	315,221	0.729	0.444	102,685	0.845	0.361
Reading test-scores (2nd grade)	181,240	510.18	73.08			
Math test-scores (2nd grade)	181,240	507.74	81.68			
Both test-scores (2nd grade)	181,240	508.9	73.44			

Notes: The GPS data was not available for 49 schools (47 primary schools and 2 secondary schools) in the Peruvian School Census. Source: Peru School Census (2006-2014)

Table 1.4: *Placement of WJC Centers in the District*

Dependent variables	WJC in district, by 2014		Added WJC in district during 2006-2014			
	(1)	(2)	(3)	(4)	(5)	(6)
# Criminal Attorney Offices	-0.022* (0.013)	-0.022* (0.013)	-0.050*** (0.015)	-0.050*** (0.015)	-0.048*** (0.016)	-0.050*** (0.015)
# Family Attorney Offices	0.090** (0.036)	0.089** (0.036)	0.110*** (0.040)	0.111*** (0.040)	0.089** (0.040)	0.109*** (0.040)
# Mixed Attorney Offices	0.106*** (0.033)	0.107*** (0.033)	0.069 (0.043)	0.071* (0.043)	0.037 (0.049)	0.070 (0.043)
# Criminal Courts	0.005 (0.018)	0.005 (0.018)	-0.001 (0.024)	-0.001 (0.024)	-0.001 (0.023)	-0.001 (0.024)
# Family Courts	-0.093** (0.040)	-0.092** (0.040)	-0.126** (0.058)	-0.127** (0.058)	-0.107* (0.058)	-0.125** (0.058)
# Mixed Courts	0.183*** (0.035)	0.181*** (0.035)	0.233*** (0.041)	0.233*** (0.042)	0.213*** (0.040)	0.232*** (0.041)
# Police Stations	0.082*** (0.012)	0.083*** (0.012)	0.049*** (0.015)	0.048*** (0.015)	0.048*** (0.016)	0.049*** (0.015)
# of Health Establishments	0.246*** (0.043)	0.247*** (0.042)	0.194*** (0.050)	0.167*** (0.049)	0.167*** (0.059)	0.196*** (0.049)
Log. Population, 2000	0.017*** (0.005)	0.014** (0.006)	0.012** (0.005)	0.012** (0.005)	0.014 (0.012)	0.011* (0.006)
△ Primary Enrollment, (1998-2005)			0.0002 (0.0002)		0.0001 (0.0003)	0.0001 (0.0002)
△ Secondary Enrollment, (1998-2005)				0.00008 (0.00008)	0.00002 (0.0002)	0.00006 (0.00009)
Domestic Violence, 2000					0.065 (0.063)	
CCT <i>Juntos</i> in the district		-0.010 (0.017)				-0.0001 (0.019)
# Households with CCT <i>Juntos</i> , 2014		0.00001 (0.00001)				0.00001 (0.00001)
Observations (Districts)	1,843	1,838	1,843	1,843	700	1,838
R-squared	0.703	0.702	0.535	0.534	0.487	0.535
Department FE	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses

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

Notes: This table shows the effects of district characteristics on WJC center's placement. The left hand side variable in Columns 1 and 2 is the number of WJC centers in the district by 2014; in Columns 3 to 6 it is whether any centers were added during the sample period 2006-2014. Standard errors are in parentheses, clustered at the district level. Source: MIMP (*Ministerio de la Mujer y Poblaciones Vulnerables*)

Table 1.5: *The Effect of WJC centers on Domestic Violence - (2006-2014)*

Dep. variable	Domestic Violence in last 12 months			
	All women	All women	Only urban clusters	Ever WJC in district
Sample	Standard	District trends	Standard	Standard
Controls	(1)	(2)	(3)	(4)
<i>Panel A: WJC center within a distance buffer from the cluster of residence</i>				
WJC within 1km	-0.022** (0.010)	-0.018* (0.011)	-0.029*** (0.010)	-0.017 (0.012)
Observations	64,363	64,363	38,395	27,996
Number of districts	1167	1167	485	215
Mean dep. var	0.390	0.390	0.399	0.397
<i>Panel B: WJC center in the district of residence</i>				
WJC in district	-0.024** (0.011)	-0.060*** (0.020)	-0.023* (0.014)	-0.032* (0.018)
Observations	96,560	96,560	58,579	42,393
Number of districts	1293	1293	531	225
Mean dep. var	0.387	0.387	0.397	0.394
District FE	YES	YES	YES	YES
Province*Year FE	YES	YES	YES	YES
Covariates	YES	YES	YES	YES
Robust standard errors in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				

Notes: The dependent variable is a dummy indicating whether the women suffered any type of domestic violence (less severe, severe or sexual violence) during the last 12 months. The independent variables measures the presence of a WJC within a 1km Euclidean buffer of the women's cluster of residence and presence of WJC center in the women's district. Robust standard errors (in parentheses) are clustered at the district level. The sample includes women between the ages of 15 and 49. Women who were never married or never cohabited are excluded from the sample. Covariates include age, age at first marriage, number of children, years of education, number of household members, number of households in the dwelling, marital status (married=1), rural residence dummy, district fixed-effects and province-by-year fixed effects. Source: Peru DHS 2006-2014.

Table 1.6: *Impact of WJC centers on Emotional Violence - (2006-2014)*

Dep. variables	Emotional violence (1)	Spouse ever humiliated (2)	Spouse ever threatened with harm (3)	Spouse ever threatened to take children (4)
<i>Sample A: All women 15-49 years old</i>				
WJC within 1km	-0.010 (0.010)	-0.002 (0.009)	-0.003 (0.006)	-0.017* (0.010)
Observations	64,364	64,364	64,364	64,364
Number of districts	1167	1167	1167	1167
Mean dep.var.	0.323	0.229	0.119	0.206
<i>Sample B: Only women in urban clusters</i>				
WJC within 1km	-0.018 (0.011)	-0.009 (0.010)	-0.007 (0.007)	-0.024** (0.011)
Observations	38,396	38,396	38,396	38,396
Number of districts	485	485	485	485
Mean dep.var.	0.337	0.239	0.114	0.219
District FE	YES	YES	YES	YES
Province*Year FE	YES	YES	YES	YES
Covariates	YES	YES	YES	YES
Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1				

Notes: The dependent variable is a dummy indicating whether the women suffered any type of emotional violence during the last 12 months. The independent variables measures the presence of a WJC within a 1km Euclidean buffer of the women's cluster of residence. Robust standard errors (in parentheses) are clustered at the district level. The sample includes women between the ages of 15 and 49. Women who were never married or never cohabited are excluded from the sample. Covariates include age, age at first marriage, number of children, years of education, number of household members, number of households in the dwelling, marital status (married=1), rural residence dummy, district fixed-effects and province-by-year fixed effects. Source: Peru DHS 2006-2014.

Table 1.7: *WJC centers and Femicides at the District Level - (2009-2015)*

Dep. var Sample	Log(# Femicides)					
	All (1)	All (2)	Age 0-19 (3)	Age 20-39 (4)	Age 40-59 (5)	Age 60+ (6)
WJC centers in the district	-0.008 (0.015)	-0.008 (0.015)	0.012 (0.008)	-0.021* (0.012)	0.003 (0.008)	0.002 (0.001)
Log (District population)		0.023 (0.036)	-0.015 (0.022)	0.017 (0.026)	0.015 (0.018)	0.012* (0.007)
Observations	12,915	12,894	12,894	12,894	12,894	12,894
Number of districts	1845	1842	1842	1842	1842	1842
Mean dep. var	0.058	0.058	0.010	0.035	0.010	0.001
District FE	YES	YES	YES	YES	YES	YES
Province-Year FE	YES	YES	YES	YES	YES	YES

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

Notes: The dependent variable number of femicides at the district level was obtained from Peru's Crime Observatory at the Ministry of Public Affairs.

Table 1.8: *WJC centers and Female Deaths due to Aggression (registered cases in hospitals) at the District Level - (2007-2014)*

Dep. var.	Log(# Female Deaths due to Aggression)	
	2007-2014	
	(1)	(2)
WJC centers in the district	-0.074** (0.031)	-0.075** (0.031)
Log (Population)		-0.057 (0.051)
Observations	7,384	7,372
Number of clusters	1846	1843
Mean dep. var.	0.080	0.080
District FE	YES	YES
Province-Year FE	YES	YES

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

Notes: Female deaths due to aggression at the district level was obtained from the Peruvian Ministry of Health. The sample of female deaths due to aggression includes women between the ages of 18 and 59 and covers the rounds 2007, 2012, 2013 and 2014.

Table 1.9: *The Effect of WJC centers on Women's Health Outcomes - (2006-2014)*

Dep. variables	Anemic (1)	Weight (kg) (2)	BMI (3)	Underweight (4)	Overweight (5)	Obese (6)	Smokes (7)
<i>Sample A: All women 15-49 years old</i>							
WJC within 1km	-0.012 (0.009)	0.508** (0.212)	0.105 (0.089)	0.001 (0.002)	-0.003 (0.009)	0.004 (0.009)	0.005 (0.003)
Observations	57,540	59,460	59,460	59,460	59,460	59,460	64,363
Number of districts	1134	1134	1134	1134	1134	1134	1167
Mean dep. var	0.261	61.72	26.86	0.006	0.511	0.208	0.037
<i>Sample B: Only women in urban clusters</i>							
WJC within 1km	-0.008 (0.011)	0.576** (0.259)	0.127 (0.107)	0.001 (0.002)	-0.005 (0.011)	0.002 (0.011)	0.007 (0.005)
Observations	34,387	35,534	35,495	35,495	35,495	35,495	38,395
Number of districts	477	477	477	477	477	477	485
Mean dep. var	0.207	63.6	27.4	0.006	0.517	0.252	0.051
District FE	YES	YES	YES	YES	YES	YES	YES
Province-Year FE	YES	YES	YES	YES	YES	YES	YES
Covariates	YES	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses

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

Notes: Source: Peru DHS 2006-2014.

Table 1.10: *WJC centers and Female Hospitalizations for Mental Health Problems at the District Level - (2006-2016)*

Dep. var.	Log(# Female Mental Health Problems) 2006-2016	
	(1)	(2)
WJC centers in the district	-0.0781* (0.043)	-0.0875** (0.043)
Log (Population)		0.685*** (0.189)
Observations	4,529	4,529
Number of clusters	1846	1843
Mean dep. var.	0.608	0.608
District FE	YES	YES
Province-Year FE	YES	YES

Robust standard errors in parentheses

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

Notes: The dependent variable number of female hospitalizations for mental health problems was obtained from the Peruvian Ministry of Health. The sample of female hospitalizations for mental health problems includes women between the ages of 18 and 59.

Table 1.11: *The Effect of WJC Centers on Children's Primary School Attendance - (2006-2014)*

Dep. variable Sample	Currently Attending Primary Level			
	All children 6-11 y.o	All children 6-11 y.o	Only urban clusters	Ever WJC in district
Controls	Standard (1)	District trends (2)	Standard (3)	Standard (4)
<i>Panel A: WJC center within a distance buffer from the cluster of residence</i>				
WJC within 1km	0.019** (0.008)	0.018* (0.009)	0.027*** (0.009)	0.023*** (0.008)
Observations	48,703	48,703	25,391	19,563
Number of districts	1159	1159	485	215
Mean dep. var	0.970	0.970	0.971	0.969
<i>Panel B: WJC center in the district of residence</i>				
WJC in the district	0.005 (0.007)	-0.005 (0.011)	0.016** (0.008)	0.022** (0.009)
Observations	71,866	71,866	38,330	29,051
Number of districts	1286	1286	531	225
Mean dep. var	0.970	0.970	0.970	0.967
District FE	YES	YES	YES	YES
Province*Year FE	YES	YES	YES	YES
Covariates	YES	YES	YES	YES

Robust standard errors in parentheses

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

Notes: The dependent variable is a dummy indicating whether the child is currently attending primary school. The independent variables measures the presence of a WJC within a 1km Euclidean buffer of the child's cluster of residence and presence of a WJC center in the child's district. Robust standard errors (in parentheses) are clustered at the district level. The sample includes children between the ages of 6 and 11. Covariates include age, gender, household's head years of education, number of children in the household aged 0-18, number of children in the household aged 0-5, number of female adults, number of male adults, rural residence dummy, district fixed effect and province-by-year fixed effect. Source: Peru DHS 2006-2014.

Table 1.12: *The Effect of WJC Centers on Children's Secondary School Attendance - (2006-2014)*

Dep. variable	Currently Attending Secondary Level			
	All children 12-16 y.o	All children 12-16 y.o	Only urban clusters	Ever WJC in district
Sample	Standard	District trends	Standard	Standard
Controls	(1)	(2)	(3)	(4)
<i>Panel A: WJC center within a distance buffer from the cluster of residence</i>				
WJC within 1km	0.022* (0.012)	0.027* (0.014)	0.029** (0.012)	0.027** (0.013)
Observations	33,519	33,519	18,266	13,570
Number of clusters	1140	1140	480	215
Mean dep. var	0.895	0.895	0.916	0.908
<i>Panel B: WJC center in the district of residence</i>				
WJC in the district	0.012 (0.016)	0.039** (0.018)	0.027 (0.020)	0.036 (0.024)
Observations	49,461	49,461	27,482	20,275
Number of districts	1270	1270	528	224
Mean dep. var	0.896	0.896	0.913	0.904
District FE	YES	YES	YES	YES
Province*Year FE	YES	YES	YES	YES
Covariates	YES	YES	YES	YES
Robust standard errors in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				

Notes: The dependent variable is a dummy indicating whether the child is currently attending secondary school. The independent variables measures the presence of a WJC within a 1km Euclidean buffer of the child's cluster of residence and presence of a WJC center in the child's district. Robust standard errors (in parentheses) are clustered at the district level. The sample includes children between the ages of 12 and 16. Covariates include age, gender, household's head years of education, number of children in the household aged 0-18, number of children in the household aged 0-5, number of female adults, number of male adults, rural residence dummy, district fixed effect and province-by-year fixed effect. Source: Peru DHS 2006-2014.

Table 1.13: School Attendance Status and Proximity to a WJC center - (2006-2014)

Sample Dep. variables	Primary School Attendance Status Children: 6-11 years old				Secondary School Attendance Status Children: 12-16 years old			
	Passed grade (1)	Repeated grade (2)	Dropped out (3)	Left school +2 years ago (4)	Passed grade (5)	Repeated grade (6)	Dropped out (7)	Left school +2 years ago (8)
<i>Sample A: All Children</i>								
WJC within 1km	0.020** (0.010)	-0.004 (0.005)	-0.018** (0.009)	0.001 (0.001)	0.020* (0.013)	-0.000 (0.005)	-0.017* (0.012)	-0.002 (0.009)
Observations	64,921	64,921	64,921	64,921	53,378	53,378	53,378	53,378
Number of districts	1165	1165	1165	1165	1161	1161	1161	1161
Mean dep. var.	0.917	0.048	0.023	0.002	0.778	0.036	0.094	0.085
<i>Sample B: Children of the women selected for the DV Module</i>								
WJC within 1km	0.023*** (0.008)	-0.006 (0.005)	-0.019*** (0.007)	0.001 (0.001)	0.030** (0.013)	-0.007 (0.005)	-0.018 (0.012)	-0.003 (0.009)
Observations	48,213	48,213	48,213	48,213	30,380	30,380	30,380	30,380
Number of districts	1155	1155	1155	1155	1135	1135	1135	1135
Mean dep. var.	0.919	0.048	0.022	0.002	0.782	0.038	0.090	0.084
District FE	YES	YES	YES	YES	YES	YES	YES	YES
Province-Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Covariates	YES	YES	YES	YES	YES	YES	YES	YES

Robust standard errors in parentheses

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

Notes: The dependent variable is a dummy indicating the school attendance status of the child. The independent variables measures the presence of a WJC within a 1km Euclidean buffer of the child's cluster of residence. Robust standard errors (in parentheses) are clustered at the district level. The sample for primary level includes children between the ages of 6 and 11 and the sample for secondary level includes children between the ages of 12 and 16. Covariates include age, gender, household's head years of education, number of children in the household aged 0-18, number of children in the household aged 0-5, number of female adults, number of male adults, rural residence dummy, district fixed effect and province-by-year fixed effect. Source: Peru DHS 2006-2014.

Table 1.14: *The Effect of WJC Centers on Primary School Enrollment (2006-2014)*

Dep. variable Sample	Log (Primary School Enrollment)				
	All schools	All schools	All schools	Only urban schools	Ever WJC in district
Controls	Standard (1)	District trends (2)	Standard (3)	Standard (4)	Standard (5)
<i>Panel A: WJC center within a distance buffer from the school</i>					
WJC within 1km	0.028*** (0.008)	0.027*** (0.008)	0.033*** (0.008)	0.032*** (0.008)	0.024** (0.010)
Log (District Population)			0.443*** (0.023)	0.424*** (0.031)	0.415*** (0.055)
Observations	315,221	315,221	315,221	119,232	103,662
Number of schools	36947	36947	36947	14405	12413
Mean dep. var	95.9	95.9	95.9	177.8	127.7
<i>Panel B: WJC center in the district of the school</i>					
WJC in the district	0.009* (0.005)	0.002 (0.004)	0.005 (0.005)	0.012** (0.006)	0.019** (0.009)
Log (District Population)			0.439*** (0.023)	0.417*** (0.031)	0.398*** (0.056)
Observations	315,407	315,407	315,407	119,270	103,730
Number of schools	36994	36994	36994	14412	12427
Mean dep. var	95.9	95.9	95.9	177.8	127.7
School FE	YES	YES	YES	YES	YES
Province*Year FE	YES	YES	YES	YES	YES
Covariates	YES	YES	YES	YES	YES
Robust standard errors in parentheses					
*** p<0.01, ** p<0.05, * p<0.1					

Notes: The dependent variable is the logarithm of enrollment plus one. The independent variables measures the number of WJC centers within a 1km Euclidean buffer from the school and presence of WJC center in school's district. Standard errors (in parentheses) are clustered at the school level. All regressions are weighted by initial school enrollment level. Covariates include school fixed effects, year fixed effects, year-by-province fixed effects, and a vector of controls of baseline school characteristics interacted with academic year (including initial school enrollment, presence of electricity, presence of piped water, school language (Spanish), urban and public school dummy).Source: Peruvian School Census 2006-2014.

Table 1.15: *The Effect of WJC Centers on Secondary School Enrollment (2006-2014)*

Dep. variable Sample	Log (Secondary School Enrollment)				
	All schools	All schools	All schools	Only urban schools	Ever WJC in district
Controls	Standard (1)	District trends (2)	Standard (3)	Standard (4)	Standard (5)
<i>Panel A: WJC center within a distance buffer from the school</i>					
WJC within 1km	0.029** (0.012)	0.017 (0.014)	0.030** (0.012)	0.034*** (0.013)	-0.005 (0.019)
Log (District Population)			0.427*** (0.038)	0.426*** (0.043)	0.442*** (0.082)
Observations	102,685	102,685	102,685	69,686	41,324
Number of schools	12809	12809	12809	8516	5175
Mean dep. var	174.8	174.8	174.8	215.3	195.3
<i>Panel B: WJC center in the district of the school</i>					
WJC in the district	0.023*** (0.008)	-0.004 (0.008)	0.014* (0.008)	0.019** (0.008)	-0.005 (0.013)
Log (District Population)			0.420*** (0.038)	0.417*** (0.043)	0.448*** (0.083)
Observations	102,691	102,691	102,691	69,692	41,324
Number of schools	12811	12811	12811	8518	5175
Mean dep. var	174.8	174.8	174.8	215.3	195.3
School FE	YES	YES	YES	YES	YES
Province*Year FE	YES	YES	YES	YES	YES
Covariates	YES	YES	YES	YES	YES
Robust standard errors in parentheses					
*** p<0.01, ** p<0.05, * p<0.1					

Notes: The dependent variable is the logarithm of enrollment plus one. The independent variables measures the number of WJC centers within a 1km Euclidean buffer from the school and presence of WJC center in school's district. Standard errors (in parentheses) are clustered at the school level. All regressions are weighted by initial school enrollment level. Covariates include school fixed effects, year fixed effects, year-by-province fixed effects, and a vector of controls of baseline school characteristics interacted with academic year (including initial school enrollment, presence of electricity, presence of piped water, school language (Spanish), urban and public school dummy).Source: Peruvian School Census 2006-2014.

Table 1.16: *The Effect of WJC Centers on Primary Level 2nd Grade Test Scores - (2006-2014)*

Dep. variable Sample	Standardized Test Scores (2nd Grade)			
	All schools	All schools	Only urban schools	Ever WJC in district
Controls	Standard (1)	District trends (2)	Standard (3)	Standard (4)
<i>Panel A: WJC center within a distance buffer from the school</i>				
WJC within 1km	0.028* (0.017)	0.018 (0.019)	0.040** (0.018)	0.027 (0.021)
Observations	181,240	181,240	92,666	69,822
Number of schools	29737	29737	13507	10858
Mean dep. var	508.9	508.9	536.9	526.9
<i>Panel B: WJC center in the district of the school</i>				
WJC in the district	0.026** (0.011)	-0.020 (0.016)	0.050*** (0.013)	0.050*** (0.016)
Observations	181,279	181,279	92,681	69,838
Number of schools	29747	29747	13510	10862
Mean dep. var	508.9	508.9	537.0	527.0
School FE	YES	YES	YES	YES
Province*Year FE	YES	YES	YES	YES
Covariates	YES	YES	YES	YES
Robust standard errors in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				

Notes: The dependent variable is the average of the standardized reading and math test scores for 2nd grade of primary school. The independent variables measures the number of WJC centers within a 1km Euclidean buffer from the school and presence of WJC center in school's district. Standard errors (in parentheses) are clustered at the school level. All regressions are weighted by initial school enrollment level. Covariates include school fixed effects, year fixed effects, year-by-province fixed effects, and a vector of controls of baseline school characteristics interacted with academic year (including initial school enrollment, presence of electricity, presence of piped water, school language (Spanish), urban and public school dummy). Source: Peru ECE 2007-2014.

Table 1.17: *Impact of WJC centers on Decision Making and Bargaining Power in the Household - (2006-2014)*

Dep. variable	Joint decision-making		
	score (0-6) (1)	score (0-1) (2)	dummy(0/1) (3)
<i>Sample: Married or cohabiting women 15-49 years old</i>			
WJC within 1km	0.040 (0.047)	0.007 (0.008)	0.017* (0.009)
Observations	72,009	72,009	72,009
Number of clusters	1168	1168	1168
Mean dep.var.	2.238	0.373	0.798
Dep. variable	Earnings compared to husband		
	Earns more than husband	Earns Less than husband	Earns the same as husband
<i>Sample: Married or cohabiting women 15-49 years old</i>			
WJC within 1km	0.008 (0.011)	-0.034* (0.018)	0.029** (0.014)
Observations	33,767	33,767	33,767
Number of districts	1094	1094	1094
Mean dep.var.	0.125	0.676	0.189
District FE	YES	YES	YES
Province*Year FE	YES	YES	YES
Covariates	YES	YES	YES
Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1			

Notes: In the DHS women are asked who makes decisions on a variety of household issues. For instance, a women is asked “*who makes the final decision on your own health care?*” “*who makes the final decision on large household purchases?*” etc. Responses include: respondent only, jointly with partner, and partner only. From these replies, we construct three measures of equal decision-making, that is, when the women makes decisions jointly with the partner. Robust standard errors (in parentheses) are clustered at the district level. The sample includes women between the ages of 15 and 49. Covariates include age, age at first marriage, number of children, years of education, number of household members, number of households in the dwelling, marital status (married=1), rural residence dummy, district fixed-effects and province-by-year fixed effects. Source: Peru DHS 2006-2014.

Table 1.18: *Impact of WJC centers on Women's Labor Force Participation - (2006-2014)*

Dep. variables	Currently working (1)	Works for family (2)	Works for someone else (3)	Self- employed (4)
<i>Sample A: All women 15-49 years old</i>				
WJC within 1km	-0.010 (0.010)	-0.004 (0.005)	-0.010 (0.008)	0.005 (0.007)
Observations	113,785	113,786	113,786	113,786
Number of clusters	1168	1168	1168	1168
Mean dep.var.	0.646	0.211	0.305	0.236
<i>Sample B: Married or cohabiting women selected for the DV module</i>				
WJC within 1km	-0.009 (0.014)	-0.004 (0.009)	-0.024 (0.017)	0.017 (0.011)
Observations	64,354	64,354	64,354	64,354
Number of districts	1167	1167	1167	1167
Mean dep.var.	0.684	0.209	0.269	0.300
District FE	YES	YES	YES	YES
Province*Year FE	YES	YES	YES	YES
Covariates	YES	YES	YES	YES
Robust standard errors in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				

Notes: The dependent variable is a dummy indicating women's labor force participation during the last 12 months. The independent variables measures the presence of a WJC within a 1km Euclidean buffer of the women's cluster of residence. Robust standard errors (in parentheses) are clustered at the district level. The sample includes women between the ages of 15 and 49. Covariates include age, age at first marriage, number of children, years of education, number of household members, number of households in the dwelling, marital status (married=1), rural residence dummy, district fixed-effects and province-by-year fixed effects. Source: Peru DHS 2006-2014.

Table 1.19: *The Effect of WJC centers on Type of Institution Women Choose to Seek Help from Domestic Violence Events- (2006-2014)*

	Sought institutional help after domestic violence event									
	Any inst. Help (1)	Police Station (2)	Justice Court (3)	District Attorney Office (4)	DEMUNA (5)	WJC Center (6)	Ombudsman Office (7)	Health Facility (8)	NGO (9)	Other (10)
<i>Sample A: All Married or cohabiting women 15-49 years old</i>										
WJC within 1km	-0.014* (0.008)	-0.019*** (0.007)	-0.001 (0.002)	-0.002 (0.002)	0.002 (0.003)	0.004* (0.002)	-0.00009 (0.0005)	-0.001 (0.002)	0.00008 (0.0001)	0.002 (0.002)
Observations	64,363	64,363	64,363	64,363	64,363	64,363	64,363	64,363	64,363	64,363
Number of districts	1167	1167	1167	1167	1167	1167	1167	1167	1167	1167
Mean dep. var	0.135	0.089	0.019	0.011	0.014	0.004	0.0009	0.008	0.0001	0.014
<i>Sample B: Only married or cohabiting women 15-49 years old who have suffered domestic violence</i>										
WJC within 1km	-0.022 (0.023)	-0.036* (0.020)	-0.0004 (0.004)	-0.008 (0.006)	0.004 (0.009)	0.009* (0.005)	-0.0002 (0.002)	0.00008 (0.004)	0.00008 (0.0001)	-0.003 (0.005)
Observations	25,090	25,090	25,090	25,090	25,090	25,090	25,090	25,090	25,090	25,090
Number of districts	1110	1110	1110	1110	1110	1110	1110	1110	1110	1110
Mean dep. var	0.293	0.193	0.043	0.026	0.033	0.011	0.002	0.019	0.0003	0.029
District FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Province-Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Covariates	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Robust standard errors in parentheses										
*** p<0.01, ** p<0.05, * p<0.1										

Notes: Source: Peru DHS 2006-2014. DEMUNA is the acronym for "Defensoria Municipal del Niño y el Adolescente" or "Municipal Defense Centers for Children and Adolescents"

Table 1.20: *The Effect of WJC Centers on Child Labor - (2006-2014)*

Dep. variable Sample	Child Labor			
	All children 6-14 y.o (1)	All children 6-14 y.o (2)	Female (3)	Male (4)
<i>Panel A: All Children</i>				
WJC within 1km	-0.021*** (0.005)	-0.008* (0.004)	-0.014** (0.006)	-0.003 (0.006)
Observations	97,933	97,933	48,108	49,816
Number of districts	1169	1169	1162	1164
Mean dep. var	0.070	0.070	0.064	0.075
<i>Panel B: Children of the women selected for the DV module</i>				
WJC within 1km	-0.024*** (0.006)	-0.012** (0.005)	-0.018** (0.008)	-0.006 (0.007)
Observations	71,410	71,410	35,162	36,215
Number of districts	1163	1163	1145	1147
Mean dep. var	0.065	0.065	0.059	0.070
District FE	YES	YES	YES	YES
Province*Year FE	YES	YES	YES	YES
Covariates	NO	YES	YES	YES

Robust standard errors in parentheses

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

Notes: The dependent variable is a dummy indicating whether the child is currently working. The independent variables measures the presence of a WJC within a 1km Euclidean buffer of the child's cluster of residence. Robust standard errors (in parentheses) are clustered at the district level. The sample includes children between the ages of 6 and 14 years old. Covariates include age, gender, household's head years of education, number of children in the household aged 0-18, number of children in the household aged 0-5, number of female adults, number of male adults, rural residence dummy, district fixed effect and province-by-year fixed effect. Source: Peru DHS 2006-2014.

Table 1.21: CCT Juntos and WJC Centers on Children's Primary School Attendance - (2006-2014)

Dep. variable	Currently Attending Primary Level			
	All children 6-11 y.o	All children 6-11 y.o	Only urban clusters	Ever WJC in district
Sample	Standard	District trends	Standard	Standard
Controls	(1)	(2)	(3)	(4)
<i>Panel A: WJC center within a distance buffer from the cluster of residence</i>				
WJC within 1km * CCT in district	-0.027* (0.015)	-0.010 (0.017)	0.014 (0.046)	-0.013 (0.017)
WJC within 1km	0.021** (0.008)	0.019** (0.009)	0.026*** (0.009)	0.024*** (0.008)
CCT in the district	0.012 (0.029)	-0.055 (0.057)	0.005 (0.042)	-0.016** (0.007)
Observations	48,632	48,632	25,377	19,563
Number of districts	1155	1155	483	215
Mean dep. var	0.970	0.970	0.971	0.969
<i>Panel B: WJC center in the district of residence</i>				
WJC in district * CCT in district	-0.019 (0.012)	0.010 (0.022)	0.000 (0.024)	-0.030 (0.033)
WJC in the district	0.009 (0.007)	-0.004 (0.012)	0.016* (0.009)	0.024** (0.010)
CCT in the district	0.008 (0.017)	-0.038 (0.031)	-0.008 (0.027)	0.011 (0.028)
Observations	71,772	71,772	38,309	29,051
Number of districts	1281	1281	529	225
Mean dep. var	0.970	0.970	0.970	0.967
District FE	YES	YES	YES	YES
Province*Year FE	YES	YES	YES	YES
Covariates	YES	YES	YES	YES

Robust standard errors in parentheses

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

Notes: The dependent variable is a dummy indicating whether the child is currently attending primary school. The independent variables measures the presence of a WJC within a 1km Euclidean buffer of the child's cluster of residence and presence of a WJC center in the child's district. Robust standard errors (in parentheses) are clustered at the district level. The sample includes children between the ages of 6 and 11. Covariates include age, gender, household's head years of education, number of children in the household aged 0-18, number of children in the household aged 0-5, number of female adults, number of male adults, rural residence dummy, district fixed effect and province-by-year fixed effect. Source: Peru DHS 2006-2014.

Table 1.22: CCT Juntos and WJC Centers on Primary School Enrollment - (2006-2014)

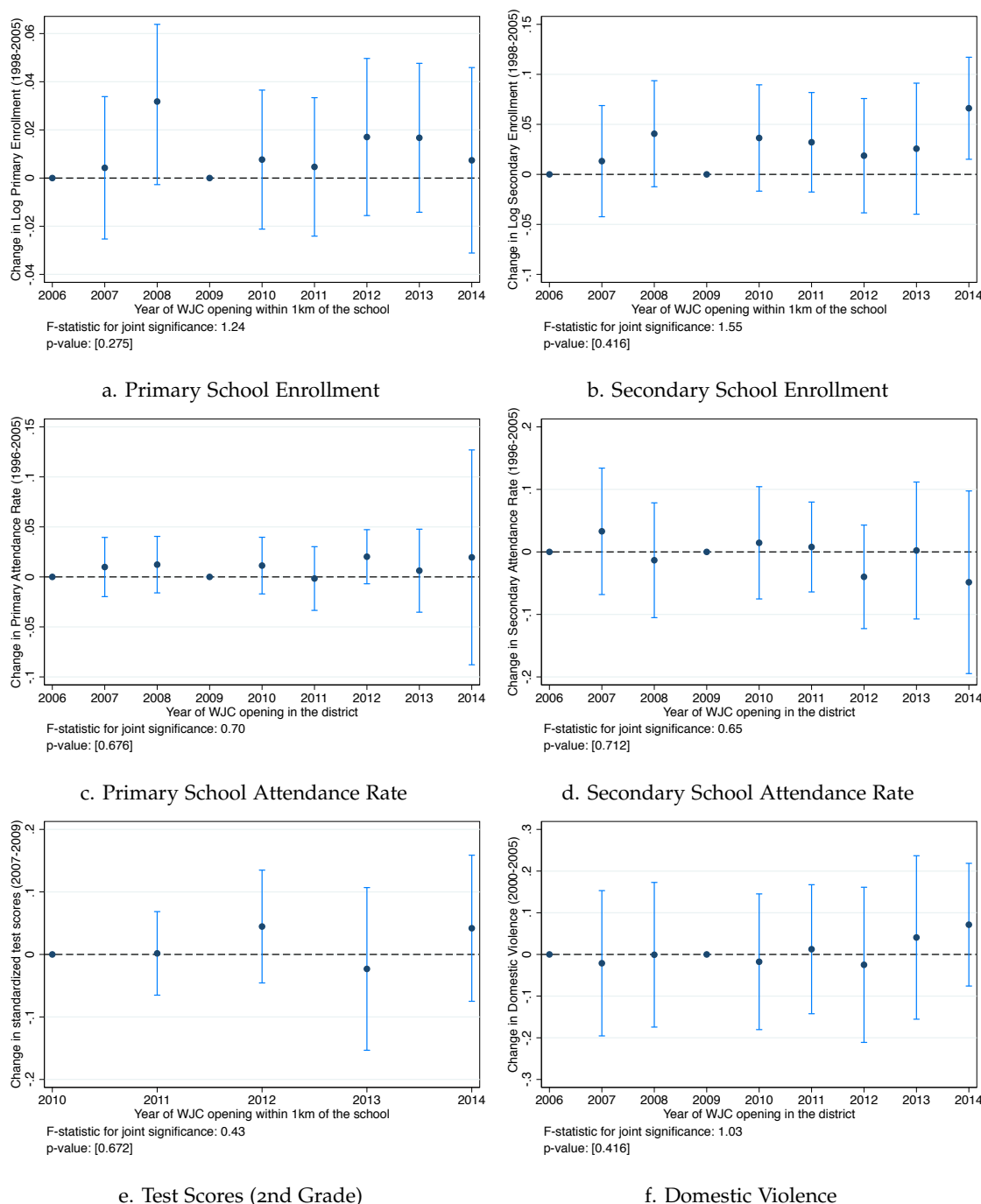
Dep. variable Sample	Log (Primary School Enrollment)				
	All schools	All schools	All schools	Only urban schools	Ever WJC in district
Controls	Standard (1)	District trends (2)	Standard (3)	Standard (4)	Standard (5)
<i>Panel A: WJC center within a distance buffer from the school</i>					
WJC within 1km * CCT in district	0.050*** (0.013)	0.035*** (0.013)	0.046*** (0.013)	0.068*** (0.014)	0.032** (0.016)
WJC within 1km	0.013 (0.008)	0.017** (0.008)	0.019** (0.008)	0.013 (0.009)	0.016 (0.011)
CCT in district	0.012*** (0.004)	0.007** (0.003)	0.012*** (0.004)	0.022*** (0.007)	0.094*** (0.023)
Log (District Population)			0.444*** (0.023)	0.423*** (0.031)	0.396*** (0.056)
Observations	314,410	314,410	314,410	118,598	103,662
Number of schools	36844	36844	36844	14325	12413
Mean dep. var	95.7	95.7	95.7	177.5	127.7
<i>Panel B: WJC center in the district of the school</i>					
WJC in district * CCT in district	0.006 (0.008)	-0.005 (0.006)	0.011 (0.007)	0.054*** (0.012)	-0.015 (0.027)
WJC in the district	0.006 (0.006)	0.003 (0.005)	-0.000 (0.005)	0.001 (0.007)	0.016* (0.009)
CCT in the district	0.015*** (0.004)	0.010*** (0.003)	0.014*** (0.004)	0.020*** (0.007)	0.105*** (0.026)
Log (District Population)			0.443*** (0.023)	0.421*** (0.031)	0.385*** (0.056)
Observations	314,596	314,596	314,596	118,636	103,730
Number of schools	36891	36891	36891	14332	12427
Mean dep. var	95.9	95.9	95.9	177.8	127.7
School FE	YES	YES	YES	YES	YES
Province*Year FE	YES	YES	YES	YES	YES
Covariates	YES	YES	YES	YES	YES

Robust standard errors in parentheses

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

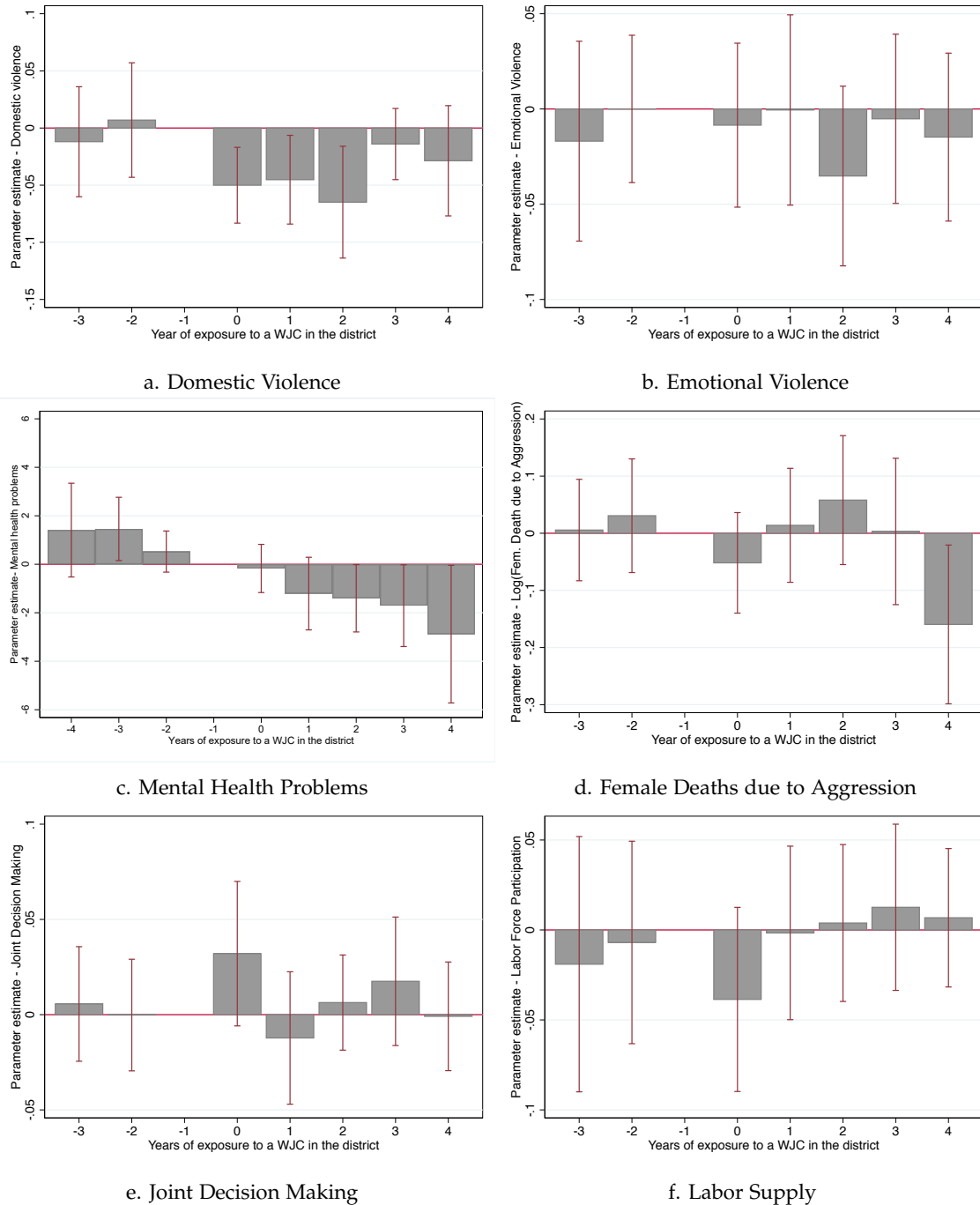
Notes: The dependent variable is the logarithm of enrollment plus one. The independent variables measures the number of WJC centers within a 1km Euclidean buffer from the school and presence of WJC center in school's district. Standard errors (in parentheses) are clustered at the school level. All regressions are weighted by initial school enrollment level. Covariates include school fixed effects, year fixed effects, year-by-province fixed effects, and a vector of controls of baseline school characteristics interacted with academic year (including initial school enrollment, presence of electricity, presence of piped water, school language (Spanish), urban and public school dummy).Source: Peruvian School Census 2006-2014.

Figure 1.9: *Effect of WJC center rollout on changes in pre-program outcomes*



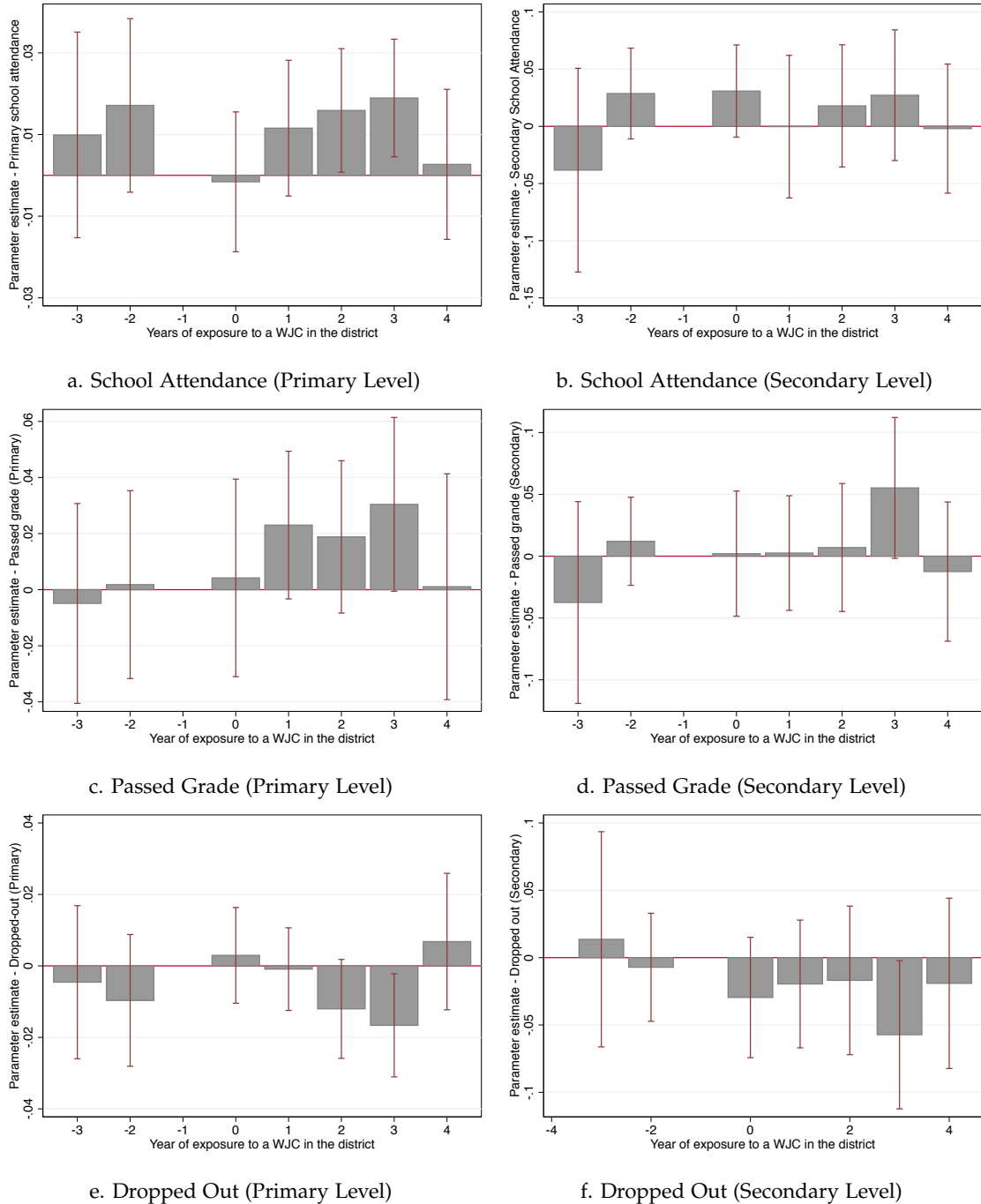
Notes: This figure shows coefficient estimates from changes in the outcomes of interest during pre-program periods (e.g. 1996-2005, 1998-2005, 2000-2005 depending on data availability) on year of WJC center introduction indicators (within 1km or in the district) and year fixed effects.

Figure 1.10: *Event Study: Pre-WJC and Post-WJC Trends in Women's Outcomes*



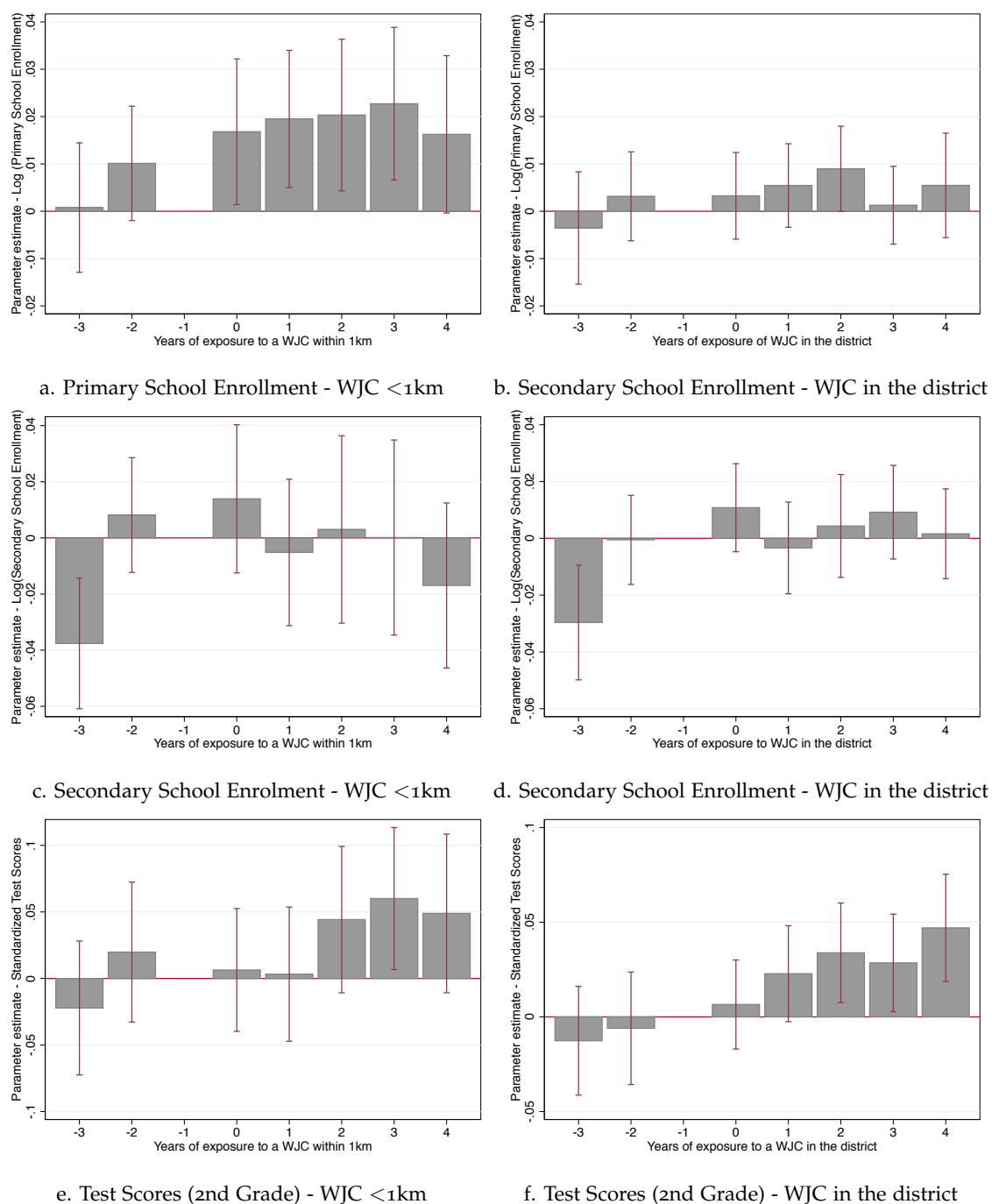
Notes: These graphs plot the coefficient obtained from a regression of the outcomes on the interaction between presence of WJC within 1km/in the district and dummies for the years leading up to the opening of the WJC centers and years after the WJC introduction. Each bar represents the estimated coefficients and the capped, vertical line shows the estimated 95% confidence interval. Covariates include district fixed effects, year fixed effects, year-by-province fixed effects, and individual controls.

Figure 1.11: *Event Study: Pre-WJC and Post-WJC Trends in Children's School Attendance*



Notes: These graphs plot the coefficient obtained from a regression of the outcomes on the interaction between presence of WJC within 1km/in the district and dummies for the years leading up to the opening of the WJC centers and years after the WJC introduction. Each bar represents the estimated coefficients and the capped, vertical line shows the estimated 95% confidence interval. Covariates include district fixed effects, year fixed effects, year-by-province fixed effects, and individual controls.

Figure 1.12: *Event Study: Pre-WJC and Post-WJC Trends in School Enrollment and Test Scores*



Notes: These graphs plot the coefficient obtained from a regression of the outcomes on the interaction between presence of WJC within 1km/in the district and dummies for the years leading up to the opening of the WJC centers and years after the WJC introduction. Each bar represents the estimated coefficients and the capped, vertical line shows the estimated 95% confidence interval. Covariates include school fixed effects, year fixed effects, year-by-province fixed effects, and a vector of controls of baseline school characteristics interacted with academic year.

Appendix

1.A Supplementary Tables

Table 1.A.1: Correlation between WJC center and CCT Juntos program implementation (2005-2014)

Dep. var.	WJC center entry _d		WJC center presence _d	
	(1)	(2)	(3)	(4)
CCT Juntos entry _d	0.002 (0.003)	0.005 (0.004)		
CCT Juntos presence _d			-0.027*** (0.008)	0.001 (0.008)
Observations	18,390	18,390	18,390	18,390
Number of districts	1839	1839	1839	1839
District FE	NO	YES	NO	YES
Year FE	NO	YES	NO	YES
Robust standard errors in parentheses				
*** p<0.01, ** p<0.05, * p<0.1				

Notes: Standard errors that allow for clustering at the district level level are reported in parentheses. Program (WJC or CCT) entry is equal to one only in the year of introduction in the district. Program presence is equal to one in every year beginning with the first year after the program entry.

Table 1.A.2: School Enrollment Effects by Gender and Grade

Dep. variable	School Enrollment					
	Primary Schools			Secondary Schools		
	Obs. (1)	Mean (2)	WJC within 1km (3)	Obs. (4)	Mean (5)	WJC within 1km (6)
<i>Panel A: Results for Schools by Gender</i>						
Log(Female enrollment)	315,221	46.9	0.033*** (0.010)	102,685	84.42	0.009 (0.017)
Log(Male enrollment)	315,221	49.9	0.021 (0.013)	102,685	90.40	0.067*** (0.014)
<i>Panel B: Results for Schools by Grade</i>						
Grade 1 enrollment	315,221	15.57	0.019* (0.010)	102,685	40.97	0.027** (0.014)
Grade 2 enrollment	315,221	17.08	0.030*** (0.009)	102,685	38.18	0.034** (0.014)
Grade 3 enrollment	315,221	16.55	0.026*** (0.009)	102,685	35.18	0.023 (0.015)
Grade 4 enrollment	315,221	16.07	0.031*** (0.009)	102,685	31.84	0.043** (0.018)
Grade 5 enrollment	315,221	15.70	0.023** (0.009)	102,685	28.64	0.044** (0.019)
Grade 6 enrollment	315,221	14.97	0.033*** (0.009)			
School FE			YES			YES
Province*Year FE			YES			YES
Covariates			YES			YES
Robust standard errors in parentheses						
*** p<0.01, ** p<0.05, * p<0.1						

Notes: The dependent variable is the logarithm of enrollment plus one. The independent variables measures the number of WJC centers within a 1km Euclidean buffer from the school and presence of WJC center in school's district. Standard errors (in parentheses) are clustered at the school level. All regressions are weighted by initial school enrollment level. Covariates include school fixed effects, year fixed effects, year-by-province fixed effects, and a vector of controls of baseline school characteristics interacted with academic year (including initial school enrollment, presence of electricity, presence of piped water, school language (Spanish), urban and public school dummy).Source: Peruvian School Census 2006-2014.

Table 1.A.3: School Enrollment and Children's School Attendance Status Effects by Gender

Dep. variable	Currently Attending School					
	Primary Level			Secondary Level		
	Children 6-11 y.o.			Children: 12-16 y.o.		
	WJC within			WJC within		
	Obs. (1)	Mean (2)	1km (3)	Obs. (4)	Mean (5)	1km (6)
<i>Sample: Female</i>						
School attendance	23,973	0.970	0.020** (0.009)	14,855	0.891	0.022 (0.019)
Passed grade	23,573	0.917	0.033*** (0.010)	12,808	0.781	0.031 (0.024)
Repeated grade	23,573	0.047	-0.010* (0.005)	12,808	0.028	-0.020 (0.009)
Dropped out	23,573	0.022	-0.025** (0.010)	12,808	0.088	-0.003 (0.018)
Left school +2 year ago	23,573	0.002	-0.0009 (0.001)	12,808	0.098	-0.006 (0.014)
<i>Sample: Male</i>						
School attendance	24,646	0.970	0.015* (0.008)	18,474	0.899	0.022 (0.015)
Passed grade	24,543	0.919	0.012 (0.009)	17,358	0.784	0.023 (0.021)
Repeated grade	24,543	0.050	-0.001 (0.008)	17,358	0.045	0.00007 (0.008)
Dropped out	24,543	0.021	-0.012* (0.007)	17,358	0.091	-0.032* (0.018)
Left school +2 year ago	24,543	0.002	0.001 (0.001)	17,358	0.074	0.009 (0.011)
District FE			YES			YES
Province*Year FE			YES			YES
Covariates			YES			YES

Robust standard errors in parentheses

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

Notes: The dependent variable is a dummy indicating whether the child is currently attending primary or secondary school. The independent variables measures the presence of a WJC within a 1km Euclidean buffer of the child's cluster of residence. Robust standard errors (in parentheses) are clustered at the district level. The sample for primary level includes children between the ages of 6 and 11 and the sample for secondary level includes children between the ages of 12 and 16. Covariates include age, gender, household's head years of education, number of children in the household aged 0-18, number of children in the household aged 0-5, number of female adults, number of male adults, rural residence dummy, district fixed effect and province-by-year fixed effect. Source: Peru DHS 2006-2014.

Table 1.A.4: Domestic Violence Effects by Age, Education Level and Type of Domestic Violence - (2006-2014)

Dep. variable	Domestic violence in last 12 months					
	Obs. (1)	Mean (2)	WJC within 1km (3)	Obs. (4)	Mean (5)	WJC in the district (6)
<i>Panel A: Results for Women by Age</i>						
Women 15-33 years old	31,442	0.349	-0.004 (0.018)	47,136	0.355	-0.013 (0.016)
Women 34-49 years old	32,886	0.402	-0.038*** (0.019)	49,380	0.418	-0.038*** (0.018)
<i>Panel B: Results for Women by Education Level</i>						
No education	2,254	0.374	-0.102 (0.110)	3,380	0.374	0.134 (0.119)
Primary Level	22,198	0.402	-0.035 (0.026)	32,844	0.390	-0.025 (0.024)
Secondary Level	24,989	0.415	-0.018 (0.015)	37,834	0.394	-0.042** (0.016)
Higher Level	14,033	0.331	-0.029* (0.016)	21,435	0.316	0.013 (0.025)
<i>Panel C: Results for Women by Type of Domestic Violence</i>						
Less severe violence	64,366	0.376	-0.029*** (0.010)	96,560	0.373	-0.018 (0.012)
Severe violence	64,366	0.171	-0.014* (0.009)	96,560	0.171	-0.006 (0.009)
Sexual violence	64,366	0.092	0.001 (0.006)	96,560	0.092	-0.007 (0.007)
District FE			YES			YES
Province-Year FE			YES			YES
Covariates			YES			YES
Robust standard errors in parentheses						
*** p<0.01, ** p<0.05, * p<0.1						

Notes: The dependent variable is a dummy indicating whether the women suffered any type of domestic violence (less severe, severe or sexual violence) during the last 12 months. The independent variables measures the presence of a WJC within a 1km Euclidean buffer of the women's cluster of residence. Robust standard errors (in parentheses) are clustered at the district level. The sample includes women between the ages of 15 and 49. Women who were never married or never cohabited are excluded from the sample. Covariates include age, age at first marriage, number of children, years of education, number of household members, number of households in the dwelling, marital status (married=1), rural residence dummy, district fixed-effects and province-by-year fixed effects. Source: Peru DHS 2006-2014.

Table 1.A.5: Domestic Violence by Children's Primary Level School Attendance Status - (2006-2014)

Primary Level (6-11 y.o.)	Children's School Attendance Status		
	Not Attending (1)	Attending (2)	Diff (3)
Domestic violence (All)	0.435 (0.010)	0.408 (0.001)	0.026** (0.010)
Observations	2,131	69,735	
Domestic violence (Urban Areas)	0.469 (0.014)	0.430 (0.002)	0.038*** (0.014)
Observations	1,149	37,185	
Domestic violence (Rural Areas)	0.395 (0.015)	0.384 (0.002)	0.010 (0.015)
Observations	982	32,550	

Table 1.A.6: The Effect of WJC Centers on Primary Level Attendance Quintiles - (2006-2014)

Dep. variable	Currently Attending Primary Level			
	Quintile 1	Quintile 2	Quintile 3	Quintile 4
<i>Sample: Children 6 to 11 years old</i>				
WJC within 1km	0.067** (0.029)	0.014 (0.010)	0.021 (0.014)	0.0002 (0.003)
Observations	11,802	8,944	9,403	18,549
Number of clusters	171	139	109	740
Mean dep. var	0.917	0.969	0.985	0.998
% Rural	0.335	0.349	0.250	0.486
District FE	YES	YES	YES	YES
Province*Year FE	YES	YES	YES	YES
Covariates	NO	YES	YES	YES

Robust standard errors in parentheses

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

Notes: Source: Peru DHS 2006-2014.

Table 1.A.7: *The Effect of WJC Centers on Children's Primary School Attendance - (2006-2014) - Alternative Euclidean Buffers*

Dep. variable Sample	Currently Attending Primary Level			
	All children 6-11 y.o	All children 6-11 y.o	Only urban clusters	Ever WJC in district
Controls	Standard (1)	District trends (2)	Standard (3)	Standard (4)
<i>Panel A: WJC center within a distance buffer from the cluster of residence</i>				
WJC within 3km	0.007 (0.011)	0.004 (0.012)	0.015 (0.014)	0.010 (0.016)
Observations	48,703	48,703	25,391	19,563
Number of districts	1159	1159	485	215
Mean dep. var	0.970	0.970	0.971	0.969
<i>Panel B: WJC center in the district of residence</i>				
WJC within 5km	-0.007 (0.008)	-0.004 (0.008)	0.005 (0.011)	0.006 (0.007)
Observations	48,703	48,703	25,391	19,563
Number of clusters	1159	1159	485	215
Mean dep. var	0.970	0.970	0.970	0.967
District FE	YES	YES	YES	YES
Province*Year FE	YES	YES	YES	YES
Covariates	YES	YES	YES	YES
Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1				

Notes: The dependent variable is a dummy indicating whether the child is currently attending primary school. The independent variables measures the presence of a WJC within a 3km and 5km Euclidean buffer of the child's cluster of residence. Robust standard errors (in parentheses) are clustered at the district level. The sample includes children between the ages of 6 and 11. Covariates include age, gender, household's head years of education, number of children in the household aged 0-18, number of children in the household aged 0-5, number of female adults, number of male adults, rural residence dummy, district fixed effect and province-by-year fixed effect. Source: Peru DHS 2006-2014.

Table 1.A.8: *The Effect of WJC Centers on Children's Secondary School Attendance - (2006-2014) - Alternative Euclidean Buffers*

Dep. variable Sample	Currently Attending Secondary Level			
	All children 12-16 y.o	All children 12-16 y.o	Only urban clusters	Ever WJC in district
Controls	Standard (1)	District trends (2)	Standard (3)	Standard (4)
<i>Panel A: WJC center within a distance buffer from the cluster of residence</i>				
WJC within 3km	0.008 (0.012)	0.009 (0.014)	0.016 (0.014)	0.012 (0.017)
Observations	33,519	33,519	18,266	13,570
Number of clusters	1140	1140	480	215
Mean dep. var	0.895	0.895	0.916	0.908
<i>Panel B: WJC center in the district of residence</i>				
WJC within 5km	-0.011 (0.013)	-0.001 (0.015)	-0.001 (0.016)	-0.003 (0.019)
Observations	33,519	33,519	18,266	13,570
Number of clusters	1140	1140	480	215
Mean dep. var	0.896	0.896	0.913	0.904
District FE	YES	YES	YES	YES
Province*Year FE	YES	YES	YES	YES
Covariates	YES	YES	YES	YES
Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1				

Notes: The dependent variable is a dummy indicating whether the child is currently attending secondary school. The independent variables measures the presence of a WJC within a 3km and 5km Euclidean buffer of the child's cluster of residence. Robust standard errors (in parentheses) are clustered at the district level. The sample includes children between the ages of 12 and 16. Covariates include age, gender, household's head years of education, number of children in the household aged 0-18, number of children in the household aged 0-5, number of female adults, number of male adults, rural residence dummy, district fixed effect and province-by-year fixed effect. Source: Peru DHS 2006-2014.

Table 1.A.9: Heterogeneity by Violent Households

<i>Sample A: Children of households where the grandmother was subject to violence by grandfather</i>					
	Enrolled (1)	Passed grade (2)	Repeated grade (3)	Dropped out (4)	Left school +2 years ago (5)
WJC within 1km	0.025*** (0.009)	0.037*** (0.011)	-0.015** (0.006)	-0.026*** (0.009)	0.003 (0.002)
Observations	20,636	19,475	19,475	19,475	19,475
R-squared	0.164	0.154	0.135	0.188	0.089
<i>Sample B: Children of households where the grandmother was not subject to violence by grandfather</i>					
WJC within 1km	0.018* (0.010)	0.014 (0.011)	-0.002 (0.009)	-0.015 (0.009)	0.000 (0.001)
Observations	27,795	28,613	28,613	28,613	28,613
R-squared	0.148	0.117	0.094	0.151	0.071
District FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Province*Year FE	YES	YES	YES	YES	YES
Covariates	YES	YES	YES	YES	YES
Robust standard errors in parentheses					
*** p<0.01, ** p<0.05, * p<0.1					

Notes: The dependent variable is a dummy indicating the school attendance status of the child. The independent variables measures the presence of a WJC within a 1km Euclidean buffer of the child's cluster of residence. Robust standard errors (in parentheses) are clustered at the district level. The sample for primary level includes children between the ages of 6 and 11 and the sample for secondary level includes children between the ages of 12 and 16. Covariates include age, gender, household's head years of education, number of children in the household aged 0-18, number of children in the household aged 0-5, number of female adults, number of male adults, rural residence dummy, district fixed effect and province-by-year fixed effect. Source: Peru DHS 2006-2014.

Table 1.A.10: *Placebo regressions, Impact of Future WJC Centers on Primary School Enrollment*

Dep. variable Sample	Log (Primary School Enrollment)			
	All schools	All schools	Only urban schools	Ever WJC in district
Controls	Standard	District trends	Standard	Standard
	(1)	(2)	(3)	(4)
<i>Panel A: WJC center within a distance buffer from the school</i>				
WJC within 1km, t	0.024*** (0.008)	0.023*** (0.007)	0.027*** (0.008)	0.019** (0.009)
WJC within 1km, $t + 1$	0.004 (0.006)	0.005 (0.006)	0.003 (0.007)	0.006 (0.007)
WJC within 1km, $t + 2$	0.002 (0.006)	-0.002 (0.006)	-0.003 (0.007)	0.002 (0.008)
WJC within 1km, $t + 3$	0.004 (0.008)	0.011 (0.008)	0.004 (0.009)	-0.002 (0.010)
Observations	315,221	315,221	119,232	103,518
Number of schools	36947	36947	14405	12398
P-value joint test	0.987	0.493	0.831	0.767
Mean dep. var	95.9	95.9	177.8	127.7
<i>Panel B: WJC center in the district of the school</i>				
WJC in the district, t	0.008* (0.004)	0.000 (0.004)	0.017*** (0.006)	0.029*** (0.008)
WJC in the district, $t + 1$	0.002 (0.004)	-0.000 (0.004)	0.006 (0.005)	-0.002 (0.006)
WJC in the district, $t + 2$	0.003 (0.004)	-0.001 (0.004)	0.001 (0.005)	0.015** (0.006)
WJC in the district, $t + 3$	-0.007 (0.005)	-0.009** (0.004)	-0.004 (0.006)	-0.012 (0.008)
Observations	315,407	315,407	119,270	103,586
Number of schools	36994	36994	14412	12412
P-value joint test	0.200	0.148	0.408	0.071
Mean dep. var	95.9	95.9	177.8	127.7
School FE	YES	YES	YES	YES
Province*Year FE	YES	YES	YES	YES
Covariates	YES	YES	YES	YES

Robust standard errors in parentheses

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

Notes: The dependent variable is the logarithm of enrollment plus one. The independent variable measures the presence of a WJC center within 1km/in the district in year t and controls for openings of future WJC centers in year $t + 1$, $t + 2$ and $t + 3$. All regressions are weighted by initial school enrollment level. Covariates include school fixed effects, year fixed effects, year-by-province fixed effects, and a vector of controls of baseline school characteristics interacted with academic year (including initial school enrollment, presence of electricity, presence of piped water, school language (Spanish), urban and public school dummy). Source: Peruvian School Census: 2006-2014.

Table 1.A.11: Placebo regressions, Impact of Future WJC centers on Secondary School Enrollment

Dep. variable Sample	Log (Secondary School Enrollment)			
	All schools	All schools	Only urban schools	Ever WJC in district
Controls	Standard	District trends	Standard	Standard
	(1)	(2)	(3)	(4)
<i>Panel A: WJC center within a distance buffer from the school</i>				
WJC within 1km, t	0.033*** (0.012)	0.023* (0.013)	0.039*** (0.013)	0.006 (0.019)
WJC within 1km, $t + 1$	-0.017 (0.013)	-0.017 (0.013)	-0.020 (0.014)	-0.032* (0.018)
WJC within 1km, $t + 2$	0.010 (0.013)	0.008 (0.014)	0.004 (0.014)	0.008 (0.020)
WJC within 1km, $t + 3$	0.014 (0.014)	0.011 (0.014)	0.023 (0.015)	0.013 (0.020)
Observations	102,685	102,685	69,686	41,277
Number of schools	12809	12809	8516	5170
P-value joint test	0.162	0.215	0.073	0.163
Mean dep. var	174.8	174.8	215.3	195.3
<i>Panel B: WJC center in the district of the school</i>				
WJC in the district, t	0.026*** (0.007)	0.002 (0.008)	0.032*** (0.008)	0.015 (0.012)
WJC in the district, $t + 1$	-0.013* (0.007)	-0.018** (0.007)	-0.008 (0.008)	-0.014 (0.011)
WJC in the district, $t + 2$	0.008 (0.008)	0.002 (0.008)	0.003 (0.009)	0.009 (0.013)
WJC in the district, $t + 3$	0.010 (0.009)	-0.002 (0.008)	0.010 (0.011)	0.009 (0.015)
Observations	102,691	102,691	69,692	41,277
Number of schools	12811	12811	8518	5170
P-value joint test	0.047	0.119	0.314	0.288
Mean dep. var	174.8	174.8	215.3	195.3
School FE	YES	YES	YES	YES
Province*Year FE	YES	YES	YES	YES
Covariates	YES	YES	YES	YES

Robust standard errors in parentheses

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

Notes: The dependent variable is the logarithm of enrollment plus one. The independent variable measures the presence of a WJC center within 1km/in the district in year t and controls for openings of future WJC centers in year $t + 1$, $t + 2$ and $t + 3$. All regressions are weighted by initial school enrollment level. Covariates include school fixed effects, year fixed effects, year-by-province fixed effects, and a vector of controls of baseline school characteristics interacted with academic year (including initial school enrollment, presence of electricity, presence of piped water, school language (Spanish), urban and public school dummy). Source: Peruvian School Census: 2006-2014.

Table 1.A.12: Relationship between WJCs within 1km rollout and pre-program school enrollment

	Schools matched to WJC within 1km, Pre-WJC period					
	$\Delta \text{Log(Primary School Enrollment)}$			$\Delta \text{Log(Secondary School Enrollment)}$		
	(1) Δ_{98-00}	(2) Δ_{98-05}	(3) Δ_{98-10}	(4) Δ_{98-00}	(5) Δ_{98-05}	(6) Δ_{98-10}
WJC within 1km in 2002	0.028 (0.032)			0.060 (0.047)		
WJC within 1km in 2003	-0.016 (0.036)			0.042 (0.050)		
WJC within 1km in 2004	-0.021 (0.035)			-0.070 (0.054)		
WJC within 1km in 2005	-0.054 (0.156)			-0.207*** (0.066)		
WJC within 1km in 2006	-0.014 (0.031)			-0.048 (0.056)		
WJC within 1km in 2007	-0.011 (0.029)	0.004 (0.015)		-0.020 (0.046)	0.013 (0.028)	
WJC within 1km in 2008	-0.006 (0.029)	0.032 (0.035)		-0.032 (0.045)	0.041 (0.027)	
WJC within 1km in 2009	-	-		-	-	
WJC within 1km in 2010	-0.034 (0.028)	0.008 (0.015)		0.003 (0.045)	0.036 (0.027)	
WJC within 1km in 2011	-0.022 (0.027)	0.005 (0.015)		-0.052 (0.045)	0.032 (0.025)	
WJC within 1km in 2012	0.002 (0.035)	0.017 (0.017)	0.005 (0.009)	-0.016 (0.043)	0.019 (0.029)	0.000 (0.013)
WJC within 1km in 2013	-0.029 (0.029)	0.017 (0.016)	0.009 (0.011)	-0.007 (0.045)	0.026 (0.033)	0.004 (0.018)
WJC within 1km in 2014	-0.021 (0.031)	0.007 (0.020)	0.004 (0.011)	-0.003 (0.043)	0.066** (0.026)	0.031* (0.016)
Observations	2,190	6,372	6,157	1,115	3,400	3,540
Number of schools	1179	1247	678	607	710	404
Year FE	YES	YES	YES	YES	YES	YES
P-value joint test	0.536	0.275	0.925	0.001	0.148	0.197

Robust standard errors in parentheses

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

Notes: Standard errors that allow for clustering at the school level are reported in parentheses. The dependent variable in columns 1-6 is the change in the logarithm of school enrollment plus one. The observations correspond to three windows of pre-WJC center period for each school. All regressions include year fixed-effects.

Table 1.A.13: Relationship between WJCs in the district rollout and pre-program school attendance

	Districts matched to WJC locations, Pre-WJC Δ 1996-2005					
	Δ Primary School Attendance			Δ Secondary School Attendance		
	(1)	(2)	(3)	(4)	(5)	(6)
	Δ 96-00	Δ 96-05	Δ 96-10	Δ 96-00	Δ 96-05	Δ 96-10
WJC in the district in 2002	0.002 (0.036)			-0.071 (0.060)		
WJC in the district in 2003	-0.056 (0.060)			0.032 (0.062)		
WJC in the district in 2004	-0.005 (0.036)			0.041 (0.082)		
WJC in the district in 2005	0.016 (0.036)			-0.051 (0.060)		
WJC in the district in 2006	-0.057 (0.052)			-0.078 (0.087)		
WJC in the district in 2007	-0.031 (0.040)	0.010 (0.015)		-0.065 (0.109)	0.033 (0.051)	
WJC in the district in 2008	-0.011 (0.039)	0.012 (0.014)		-0.008 (0.098)	-0.013 (0.046)	
WJC in the district in 2009	-	-	-	-	-	-
WJC in the district in 2010	-0.026 (0.040)	0.011 (0.014)	-0.009 (0.008)	-0.062 (0.071)	0.015 (0.045)	-0.013 (0.028)
WJC in the district in 2011	-0.034 (0.041)	-0.002 (0.016)	-0.016 (0.009)	0.030 (0.067)	0.008 (0.036)	-0.029 (0.024)
WJC in the district in 2012	0.012 (0.039)	0.020 (0.014)	0.006 (0.008)	0.022 (0.076)	-0.040 (0.042)	-0.052 (0.041)
WJC in the district in 2013	-0.008 (0.049)	0.006 (0.021)	-0.012 (0.011)	0.055 (0.101)	0.002 (0.055)	-0.015 (0.030)
WJC in the district in 2014	-0.073 (0.076)	0.020 (0.054)	-0.007 (0.038)	-0.152 (0.125)	-0.049 (0.074)	-0.030 (0.054)
Observations	90	186	228	90	184	226
Number of districts	90	106	102	90	106	102
Year FE	YES	YES	YES	YES	YES	YES
P-value joint test	0.000	0.676	0.222	0.000	0.712	0.778
Robust standard errors in parentheses						
*** p<0.01, ** p<0.05, * p<0.1						

Notes: Standard errors that allow for clustering at the district level are reported in parentheses. The dependent variable in columns 1-6 is the change in school attendance rate at the district level. The observations correspond to three windows of pre-WJC center period for each district. All regressions include year fixed-effects.

Table 1.A.14: Relationship between WJCs within 1km rollout and four windows of pre-program standardized test scores (2nd grade - Primary School)

	Schools matched to WJC within 1km Pre-WJC period			
	Δ Standardized Test Scores			
	(1)	(2)	(3)	(4)
	$\Delta 07-09$	$\Delta 07-10$	$\Delta 07-11$	$\Delta 07-12$
WJC within 1km in 2011	0.002 (0.034)			
WJC within 1km in 2012	0.045 (0.046)	-0.009 (0.029)		
WJC within 1km in 2013	-0.023 (0.066)	-0.029 (0.038)	-0.001 (0.034)	
WJC within 1km in 2014	0.042 (0.060)	-0.019 (0.039)	-0.009 (0.033)	-0.025 (0.034)
Observations	1,565	1,675	1,068	734
Number of schools	821	600	292	168
Year FE	YES	YES	YES	YES
P-value joint test	0.670	0.895	0.828	

Robust standard errors in parentheses

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

Notes: Standard errors that allow for clustering at the school level are reported in parentheses. The dependent variable in columns 1-4 is the change in standardized reading and math z-scores at the school level. The observations correspond to the pre-WJC center period for each school, it includes all schools which are located within 1km of a WJC center which opened between 2010-2014, 2011-2014, 2012-2014 and 2013-2014. All regressions include year fixed-effects.

Table 1.A.15: Relationship between WJCs in the district and four windows of pre-program domestic violence

	Districts matched to WJC locations, Pre-WJC period			
	Δ Domestic violence in last 12 months			
	(1)	(2)	(3)	(4)
	Δ 2000-2005	Δ 2000-2008	Δ 2000-2010	Δ 2000-2013
WJC in the district in 2007	-0.021 (0.088)			
WJC in the district in 2008	-0.001 (0.087)			
WJC in the district in 2009	-	-		
WJC in the district in 2010	-0.018 (0.082)	-0.006 (0.035)		
WJC in the district in 2011	0.013 (0.078)	0.007 (0.034)	-0.026 (0.042)	
WJC in the district in 2012	-0.025 (0.093)	0.060 (0.041)	-0.011 (0.041)	
WJC in the district in 2013	0.041 (0.098)	0.013 (0.061)	0.005 (0.050)	
WJC in the district in 2014	0.071 (0.074)	0.119** (0.078)	-0.036 (0.042)	-0.016 (0.020)
Observations	105	161	239	128
Number of districts	78	99	83	38
Year FE	YES	YES	YES	YES
P-value joint test	0.416	0.103	0.433	-

Robust standard errors in parentheses

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

Notes: Standard errors that allow for clustering at the district level are reported in parentheses. The dependent variable in columns 1-4 is the change domestic violence at the district level. The observations correspond to the pre-program period of the WJC center rollout for each district, it includes all districts that ever had a WJC center which opened between 2006-2014, 2009-2014, 2010-2014 and 2013-2014. All regressions include year fixed-effects.

Chapter 2

Fertility and Parental Labor-Force Participation: New Evidence from a Developing Country in the Balkans¹

¹I would like to thank and recognize my advisor Karen Macours for her special support and guidance throughout the elaboration of this paper. Helpful comments and advice from Ekaterina Zhuravskaya, Oliver Vanden Eynde, Gustavo Bobonis, John Giles, Marc Gurgand, Andrew Clark, Iván Torre and Martin Rossi are also gratefully acknowledged. I also benefited from comments by two anonymous referees and seminar participants at PSE, Royal Economic Society Conference 2016, GRETHA 2016 and the Spring Meeting of Young Economists 2016. The views expressed in this paper are those of the author and do not necessarily represent those of Paris School of Economics. I am responsible for all remaining errors.

Abstract

This paper examines the effect of fertility on parental labor-force participation in a developing country in the Balkans, with particular attention to the intervening role of childcare provided by grandparents in extended families. In order to address the potential endogeneity in the fertility decision, I exploit Albanian parental preference for having sons combined with the siblings sex-composition instrument as an exogenous source of variation. Using a repeated cross-section of parents with at least two children, I find a positive and statistically significant effect of fertility on parental labor supply for those parents who are more likely to be younger, less educated or live in extended families. In particular, IV estimates for mothers show that they increase labor supply, especially in terms of hours worked per week and the likelihood of working off-farm. Similarly, father's likelihood of working off-farm and having a second occupation increase as a consequence of further childbearing. The heterogeneity analysis suggests that this positive effect might be the result of two plausible mechanisms: childcare provided by non-parental adults in extended families and greater child costs which face households to a tighter budget constraint.

JEL Classification: J13, J22, C26

Keywords: fertility, parental labor-force participation, instrumental variables

2.1 Introduction

The relationship between fertility and labor supply has been of longstanding interest in economics. The vast majority of the empirical studies focus on female labor supply and they report a negative relationship between the two variables, particularly among countries that grew rapidly in the second half of the 20th century (e.g. [Angrist and Evans, 1998](#); [Boca, Pasqua and Pronzato, 2005](#); [Cristia, 2008](#); [Bruijns, 2014](#); [Hupkau and Leturcq, 2016](#)). There is also a scattering of studies in developing countries (low- and middle-income countries) which has found not only negative but also mixed results (e.g. [Porter and King, 2012](#); [Agüero and Marks, 2011](#); [Bloom et al., 2009](#); [Cruces and Galiani, 2007](#)).

A more recent literature pointing in the direction of a negative effect is suggestive that this result could be extrapolated to other contexts. This is something examined extensively by [Dehejia, Pop-Eleches and Samii \(2015\)](#) and [Bisbee et al. \(2015\)](#), who find that quasi-experimental evidence generalizes more readily to countries which share closer geographical, education, time and labor force participation characteristics. In addition, a parallel literature that uses a wide array of countries and over 200 years of history has emphasized a small and insignificant effect of fertility on labor supply at low levels of development and an economically large and negative effect at higher levels of development (e.g. [Aaronson et al., 2017](#)).

The contribution of this paper is to build upon this previous literature by arguing that, contrary to most of the existing literature, the effect of fertility on parental labor supply can be positive and statistically significant in the context of a developing country in the Balkans (Albania). Furthermore, this paper presents a comparative analysis on the relationship between fertility and labor supply based on the type of instrument. By decomposing the same-sex instrument, the results found in the paper suggest that in a context characterized by a strong son preference, all the relevant effect is coming from the *two girls* siblings sex composition.

To clearly identify the relationship between fertility and labor force participation an exogenous source of variation in family size is needed.² In contrast to American parents which

²Two methods based on natural experiments have been mainly used in the literature to tackle this endogeneity problem in the fertility decision. The first strategy was first introduced by [Rosenzweig and Wolpin \(1980\)](#) and uses the natural occurrence of multiple births or twinning at first birth to identify the effect of fertility on labor supply. They find that, although women who have had twins withdraw temporarily from the labor market, their labor-force participation is not affected overall. The second strategy was first proposed by [Angrist and Evans \(1998\)](#) and uses the preference of American parents for balancing the sex composition of their children. They find that fertility has a negative effect on the labor supply of women, but no effect on the labor supply of men.

prefer to balance the sex composition of their children, Albanian parents have another fertility preference which is generally known as *son preference*.³ Albania is characterized by a patriarchal family system where parents prefer sons to daughters, especially until they have at least one son. Therefore, in the same spirit as Angrist and Evans (1998), I exploit Albanian parental preference for having sons combined with the siblings sex-composition instrument as an exogenous source of variation in the fertility choice. Thus, the analysis focuses on a specific margin: I estimate the impact of having a third child or higher order on parental labor supply outcomes in comparison to those with only two children.

The data used in this paper comes from four different surveys conducted in Albania between 2002 and 2012. More specifically, I use the 2002, 2005 and 2012 Albanian Living Standard Measurement Surveys (LSMS) and the 2008-09 Albanian Demographic Health Survey (DHS). The advantage of these surveys is that they contain not only rich information on individual's socioeconomic and demographic characteristics, but they also contain detailed information on the family structure of each one of the households, including information about the children no longer living in the household, which makes possible the construction of the entire parental fertility history. For the purpose of this analysis, I construct a repeated-cross section of married mothers and fathers with two or more children whose oldest child is younger than 18 years old.

The main results of this paper argue that an increase in family size has a *positive* impact on parental labor-force participation in the context of Albania. More precisely, for the sample of mothers I find that having an additional child (third or higher order) increases on average a mother's likelihood of *working off-farm* by 5 percentage points and it also increases the working time by around 4.4 *hours per week* on average. Similarly relevant is the finding that fathers also change positively their labor-market behaviour in response to a change in family size. In particular, having an additional child increases father's likelihood of *working off-farm* by 6.7 percentage points, and also their likelihood of having a *second occupation* by 4 percentage points.

The second focus of the paper is to pinpoint the channels through which these results operate. I propose two plausible mechanisms behind the positive effect of fertility on parental

³There are several previous paper that use parental preference for sons as an instrument for fertility, especially in developing countries. See Ebenstein (2009) for Taiwan, Chun and Oh (2002) for South Korea, Lee (2002) and, Arnold and Zhaoxiang (1986) for China.

labor supply for the case of Albania. First, family structure in developing countries, especially in rural areas, is characterized by the presence of extended families where several generations co-reside in or near the same household ([Rosenzweig and Wolpin, 1985](#)). In response to an increase in fertility, grandparents may adjust their time allocation by supplying more time in childcare and household chores. Being freed from such tasks reduces the reservation wage for the parents, leading to an increase in labor supply. Thus, for extended families the effect of fertility on parental labor supply depends on the substitutability between grandparent's time and parent's time in caring for children.⁴ Second, children do not only require costs in terms of time dedicated to them but as well in terms of direct costs such as food, clothing, education or health. Given that a large part of the household budget in developing countries is used to cover basic necessities, such as food and shelter, an increase in financial costs might induce a negative income effect that is sufficiently large to drive some parents into the labor force.

To this end, I perform an heterogeneity analysis where I analyze whether the results vary by parent's education, parent's birth cohort and family structure. In line with these hypothesis, the analysis suggests that the labor-market consequences of fertility are more likely to be driven by poorer, less educated, and younger parents. In particular, mothers living in extended families experience the largest positive effects of fertility on labor supply suggesting that childcare provided by grandparents in extended families reduces the parental time cost of looking after children. However, the positive effect of fertility on father's labor supply is mostly driven by those living in non-extended families, which suggests that in this type of families parents do not count on childcare provided by grandparents and therefore the mother has to stay in the household while the father probably supplies more labor in the market.

In this regard, this paper not only contributes to the traditional literature on fertility and labor supply but it also adds new important insights into the functioning of labor markets in developing countries as a response to an increase in fertility. First, this paper contributes to a more recent literature on the role of grandparental co-residence in household behavior. In particular, it sheds light on the intervening role of childcare provided by grandparents in extended families in the relationship between fertility and parental labor supply. Second,

⁴Several studies have examined the effect of family structure on women's labor supply. They find that co-residence seems to increase the labor supply of women who have children (e.g. [Wong and Levine, 1992](#); [Sasaki, 2002](#); [Gong and Van Soest, 2002](#); [Maurer-Fazio et al., 2011](#); [Compton and Pollak, 2014](#); [Garcia-Moran and Kuehn, 2017](#); [Posadas and Vidal-Fernández, 2013](#); [Shen, Yan and Zeng, 2016](#); [Landman, Seitz and Steiner, 2017](#)).

I analyze parental response to an increase in fertility on the type of work they do and the varying degrees of flexibility in the labor market, such as work off-farm, work on-farm and self-employed. This is essential in a developing country since labor markets in this context are characterized by higher levels of informality and a considerable share of the employment in rural regions.⁵ Third, since there are still few studies that consider the effect of fertility on men's labor-force outcomes, I examine how the mentioned effects differ across gender by including not only women but also men in the analysis. Lastly, this study is, to the best of my knowledge, the first that exploits Albanian parental preference for having sons combined with the siblings sex-composition instrument to examine the causal link between fertility and parental labor supply.⁶

The results of this paper also have important policy implications. The main finding indicates that with lower childcare costs provided by the presence of grandparents in extended families but also greater child costs, parental labor supply is not reduced as a consequence of an increase in family size, suggesting that as an informal childcare provision mechanism, extended families play an important mitigating role for the time cost of children. However, as family structure in developing countries becomes more nuclear with economic development, the time cost of children may increase for parents which may lead to further fertility decline and fewer opportunities for non-parental household members to help with childcare or household chores. In view of this decline in intergenerational co-residence over time, governments in developing countries should establish a subsidized and universally accessible public childcare system in order to promote labor supply and sustainable economic development.

The remainder of the paper proceeds as follows. Section 2.2 provides a brief summary on fertility and labor-force participation for the case of Albania. Section 2.3 describes the datasets and some descriptive statistics. Section 2.4 presents the identification strategy and empirical estimation. The econometric results of the paper are presented in Section 2.5 and discussed in

⁵In developing countries, women are under-represented in higher proportions in the labor market and, therefore, primarily engaged in family activities. In addition, the degree of informality in labor market reveals an additional dimension of inequality in labor outcomes. Evidence suggests that women are more likely to be engaged in the agricultural sector and informal labor market (e.g. self-employed or unpaid family worker) (Blunch et al., 2001).

⁶Albania has experienced a drastic fertility decline during the last fifty years, which has fallen from almost 7 children per women in 1960 to only 1.65 in 2012 (Gjonça, Aassve and Mencarini, 2008). In addition to low fertility levels, Albania is also characterized by a declining employment rate for both men and women during the last years. According to the Albanian Institute of Statistics (INSTAT), labor-force participation has continuously fallen during the last twenty-five years from 75 % in 1989 to 62 % in 2013, being women the most damaged from this decline. Therefore, the setting of a declining fertility rate accompanied by declining employment rates and combined with a patricarchal society make Albania an interesting case for this analysis.

Section 2.6. Section 2.7 briefly concludes.

2.2 Background

2.2.1 The Evolution of Total Fertility Rate in Albania

In the mid-to-late 20th century, when modern contraceptive methods were developed and popularized in other countries, Albania was a Popular Socialist Republic, which promoted an aggressive pro-natalist policy that banned abortion. Modern family planning methods were virtually unknown and, although some were available, access was strictly controlled. The communist government believed that a larger population was necessary in order to protect the country from foreign influences that could exploit its natural resources. When communists took over the government of Albania at the aftermath of the II World War, the country had a population of just over 1.1 million people. In fact they had already found a pro-natalist environment which was reinforced by traditional patriarchal norms. After communism collapsed in 1990, the population had tripled in a period of less than 45 years to 3.3 million. With a total fertility rate of 6 children per woman in 1950, Albania had the highest fertility in Europe, which reached a peak of almost 7 children per woman by 1960. This rise in fertility during the 1950s and early 1960s mirrors the experience of many European countries during the same period, with the famous *baby boom* period of the mid-1960s.

Even though during several decades Albania has had one of the highest levels of fertility in Europe, the country has experienced a substantial fertility decline during the last fifty years, which has fallen from almost 7 children per women in 1960 to 1.65 in 2012. Figure 2.1 shows the evolution of the total fertility rate in Albania. The reduction in the fertility rate during the communist regime occurred in the absence of modern contraception and abortion, which suggests that other social and economic policies, in particular the ones that improved the social agenda in the country, might have had an indirect effect on fertility in Albania. Most significant of these were policies focused on the improvement of education in the country. For instance, the investment in education, with particular focus on the improvement of female education was unprecedented in Albania. Female illiteracy improved from 92% in 1945 to less than 8% in 1989, and by 2002 it was less than 5%, similar to most developed European countries ([Gjonça, Aassve and Mencarini, 2008](#)).

After the collapse of the communist regime in 1990, the decline in the total fertility rate has been even more substantial due to the introduction of many new laws and social policies which have aimed to regulate the transition to a more market-oriented and modern society. One of the first changes in legislation that the Albanian Government introduced in 1995 has been the legalization of abortion. Also, family planning methods have become widely available, even though the level of information and education for its use still remains low. In addition, the political openness in the early 1990s has permitted people to move freely and also to emigrate. As a consequence, an enormous emigration mass took place in a very short period of time which was concentrated in the reproductive age groups. While the move towards market economy has been associated with the emancipation of the society in general and women in particular, it has also generated negative consequences in terms of employment. The collapse of industry brought back a large unemployment, especially for women (UNFPA, 2012; Gjonça, Aassve and Mencarini, 2008).

During the transition period studied in this paper (2002 - 2012), it is clear that Albania has a new setting with regards to fertility compared to the communist period. The existence of means of birth control which were not present before 1990, the high female unemployment rates, the economic crisis and the continuing emigration are several factors that might have contributed to the declining fertility rate during this decade.

2.2.2 Labor-Force Participation in Albania

During the communist regime, the labour market in Albania was characterized by a state controlled individual decision-making and a high degree of centralization. In this system, there were social and economic policies that promoted continuously the equality between women and men in different social spheres, including the labor market.⁷ Since the communist regime was interested in full female employment, the government provided support to the families through social and economic policies including child care benefits and maternity leave schemes, complemented by a state-sponsored system of day-care nursing and kindergartens across the whole country. In addition, the prices for a number of essential products for children were subsidized by the state. As a result, women were freed from childcare responsibilities and were

⁷For instance, the constitution of the People's Socialist Republic of Albania (Article 41) stated *"The woman enjoys equal rights with a man in the work place, payment, holidays, social security, education and in all social-political activities as well as in the family"* (Kuvendi Popullor, 1976).

thus able to work. In 1989, female labor-force participation had reached 72.8%, an increase of more than 50% compared to 1960 (35.9%) and by that time there were only slight differences in labor-force participation between men (77.2%) and women (72.8%) (Gjonça, Aassve and Mencarini, 2008; INSTAT, 2004).

After the fall of the communist regime, Albania went through many radical demographic, economic and political changes. During this transitional period, women in particular have faced a substantial withdrawal from the labor market due to several reasons. First, state enterprises, where women made up the greatest percentage of the workforce before 1990 collapsed due to their inefficiency and so did social protection associated with these jobs (Tarifa, 1994). Second, the state stopped providing child care or long paid maternity leaves and at the same time many day-care nurseries and kindergartens were closed. For instance, compared to 1990, the number of kindergartens had fallen by 60% in urban areas and by 49% in rural areas in 2004 (IMF, 2006).⁸ Third, during this period the country also witnessed a massive male emigration, which has left the wives at home taking care of the children, engaging in housework or informal market and thereby decreasing their working hours out of the house. Also, the internal migration to urban areas has damaged the labor supply of women since they face more problems entering the urban labor market in the current economic environment. Consequently, women's schedules have required more accommodation giving rise to long-term structural discrimination in the labor market and high female unemployment rates.

Labor force participation has still remained low in Albania during the last decade, especially for young people and women. Statistics from INSTAT (2014) show that the labor-force participation rate for women between 15-64 years old is 50.1% whereas for men is 70.2%. Employed women are still finding it more difficult to balance career and family as the availability of state-financed social services is quite low and the number of pre-school facilities has not recovered to the levels previous to the collapse of the communist regime (See Figure 2.3). As a result, two decades after the fall of the communist regime, one may observe that the mutual support between family members has assumed a great role in attenuating the shock of poverty inherited from the communist regime, and that informal channels of support function better than the state social provisions and assistance (Danaj, 2014).

Therefore, a distinguishing feature of this study is that it is quite interesting to analyze the

⁸See Figure 2.3 for more detail on the situation of pre-school in Albania during the period analyzed in this study.

effect of fertility on parental labor force participation in a developing country characterized by a combination of low levels of employment and also low fertility rate. Figure 2.2 presents these trends since 1990 until 2012 for the case of Albania, where we can clearly observe a decrease in both variables. The timing of these events suggests that there might exist a causal relationship between them.

2.3 The Data

The data used in this paper come from surveys conducted in Albania in four different moments in time: the 2002 Living Standard Measurement Survey (LSMS), the 2005 Living Standard Measurement Survey (LSMS), the 2008-09 Demographic Health Survey (DHS) and the 2012 Living Standard Measurement Survey (LSMS). The LSMS surveys were undertaken by the Albanian National Institute of Statistics (INSTAT) with the technical assistance of the World Bank. The 2008-09 DHS was conducted by the INSTAT but also by the Institute of Public Health (IPH). These surveys correspond to different wave rounds, but it is important to note that they do not form a panel database. They are nationally representative surveys for the whole country, as well as at regional and at the urban/rural level. Each survey contains a wide range of information on individual's demographic and socioeconomic characteristics (i.e. fertility, health, employment, migration and education among many others). In each one of these databases, the country is stratified into four regions (strata), Tirana, Coastal, Central and Mountain. All the surveyed individuals belong to the 36 districts in the country, with all regions nearly equally represented.

These surveys interviewed both women and men with respect to their fertility decisions, thereby permitting the unique opportunity to incorporate information in my study regarding the fertility history of each household. In other words, each one of these surveys contains detailed information which helps identify children by birth of order and match them with their biological mother and father. For instance, in the 2008-09 DHS survey, all the children ever born are already matched to each one of their corresponding mother and father, which makes the DHS a perfect database for studying the fertility history of the parents. In contrast, in the LSMS surveys, the fertility history information is not as clearly identified as in the DHS, which means that the children are not already linked to their corresponding parents. However, the advantage of the LSMS surveys is that they provide a direct mother and father identifier code, which

is used to link the children to their corresponding parents. In addition, the Albanian LSMSs databases are a special case in terms of fertility compared to other LSMS surveys because they contain very detailed information on the family structure of each one of the households, including information about the children no longer living in the household, which makes possible the construction of the entire fertility history of each parent in the databases.⁹

Following similar studies in the literature, I limit the analysis to mothers aged between 20 and 49 years old with at least two children whose oldest child was at most 18 years old at the time of the surveys. Firstly, women who are younger than 20 years old are excluded from the analysis because none of them have two children. Since the instrumental variable used in this study is based on a comparison of the sex of the parent's previous children, it is necessary to include only individuals with two or more children. Secondly, mothers with children older than 18 years old at the time of the survey are also excluded from the sample. This restriction ensures that the first child has not yet moved from home at the time of the survey.¹⁰ Additionally, mothers with a child below the age of 1 are also excluded from the analysis in order to avoid potential bias due to lower labor market activity of mothers during the initial months following childbirth.

Taking into account these restrictions, with these four surveys I construct two repeated-cross section subsamples. The first includes mothers with two or more children whose oldest child is younger than 18 years old. The second includes fathers (husbands of the married mothers) with two or more children whose oldest child is younger than 18 years old. I do not carry out a separate analysis on the subsample of married women because in my final sample the married women represent 98% of the total (see Table 2.2), which makes no significant difference with respect to the whole sample. Thus, the final samples are made up of 7,480 observations of mothers and 5,496 observations of fathers with at least two children under the age of 18.¹¹

⁹The 2002 LSMS contains a specific module on fertility, where it is possible to identify the maternity history of every woman in the database. This type of module is not included in the 2005 and 2012 LSMS surveys, but for these cases the Migration module and Sons and Daughters Living Away module are used to construct the fertility history.

¹⁰In Albania, it is increasingly likely for a child over 18 years old to have moved to a different household. For instance, using the LSMS surveys, which provide information on the sons and daughters who have left the household, I find that the average age at which a child in Albania left his home is 23.66 during the period studied. Moreover, the age at which a child left the household is quite different between males and females. For example, the average age at which a girl left home is 21.92 years old, while for boys it is 25.41. These descriptive statistics indicate that Albanian girls leave their homes quite earlier compared to boys and this is generally due to marriage.

¹¹It is important to note that the difference in observations between the married mothers (7,480) and the married fathers (5,496) is due to the number of observations as provided in the official 2008-09 Albanian DHS. In other words, there are more observations of women than men. A detailed description of the LSMS surveys and the DHS

2.3.1 Descriptive Statistics

Descriptive statistics for mother's fertility in Albania are given in Table 2.1. As the independent variable of primary interest for this analysis is fertility, the variable *number of children ever born per mother* is used as a measure of fertility.¹² Using the sex of the oldest two children, I define the siblings sex composition pairs as *same sex* if they have the same gender and *mixed sex* if otherwise. The former is also decomposed to define sibling pairs of *two girls* and *two boys*.

Table 2.1 indicates that the average number of children ever born for Albanian mothers with at least two children is 2,60. Some families have as many as 9 children. Among all mothers with two children, about 41% considered having a third child, an event indicated by the variable *More than 2 children*. Both the number of children and the proportion of mothers with three or more children are quite similar to other studies in the literature (e.g. Angrist and Evans, 1998). Moreover, just 50% of the households with at least two children had either a girl or a boy as a firstborn and also just 50% of all two child households had children of the same sex. However, the occurrence of two boys in a row is slightly higher (25%) than the occurrence of two girls (24%).

Demographic and labor-force participation descriptive statistics for mothers and fathers are shown in Tables 2.2 and 2.3, respectively. These tables include variables such as parent's age, age at first birth, years of education, household size, residential region and indicators for religion among others. The labor-force participation variables are based on questions concerning the employment status of the parent during the last 7 days and 30 days previous to the moment of the surveys. The four surveys (LSMS and DHS) contain labor supply indicator variables which measure whether the parents worked for somebody other than a household member (*worked off-farm*), whether they worked on a farm owned by a household member (*worked on-farm*) and whether the parents worked on own account (*self-employed*). It is important to notice that these working activities are not exclusive, which means that each parent might be doing more than one. Thus, I define another indicator variable that combines these three variables and takes the value of 1 if the parents performed at least one of these working activities during the last 7 days (*worked in last 7 days*) and 0 otherwise.

survey for Albania and on the construction of the subsamples and some key variables is provided in appendix A. See Table 2.D.1 in the Appendix for more information.

¹²This measure of fertility, *number of children ever born*, also includes the deceased children.

Just the LSMS surveys also contain other labor supply variables that measure whether the parents have a *second occupation*, their usual number of *hours worked per week* and also their amount of *labor income per month* in old leks.¹³ The latter two variables are set to zero for those parents who reported not working during the last 7 days.

The descriptive statistics in Tables 2.2 and 2.3 indicate that the mean age is 35 for the sample of mothers and almost 40 for the fathers (husbands or partners of mothers). On average, women become mothers at the early age of 23 compared to 28 for men. Parents in Albania complete a similar number of years of education, being 10 years for the mothers and 11 for the fathers. Half of the households in the sample live in urban areas and around 38% of them live in an extended family.

About 44.5% of the mothers with at least two children were working during the last 7 days previous to the surveys, whereas this number rises up to 78.7% for the sample of fathers. This difference of 34.2 percentage points in favour of labor-force participation of men reflects their much more active involvement in the labor market of Albania. In addition, fathers are much more likely of having a second occupation (5.8%) compared to mothers (1.4%). From the sample of mothers with at least two children, 16.3% were *working off-farm*, 21.7% were *working on-farm* and only 6.6% were *self-employed*. While, for the husbands or partners of mothers, 39.3% were *working off-farm*, 23.6% were *working on-farm* and 19.7% were working on own account. This summary statistics indicate that men in Albania dominate in the non-agricultural sector, while women make up a higher proportion of employed persons in the agricultural sector.

Overall (i.e., including zeros), the average number of *hours worked per week* is 16.6 among the mothers, but it is 40 hours among those that effectively work. On the other hand, fathers in the sample worked 34.2 hours per week (i.e., including zeros) and this number rises up to 44.8 for those that worked a positive number of hours. Lastly, the average monthly labor income for mothers with at least two children is 50,166 old leks (221,591 without zeros), whereas fathers earn on average 191,841 old leks (331,172 without zeros), which is quite more compared to mothers.¹⁴ Once again, it is important to notice the gender inequality in terms of labor-force participation and earnings in Albania during the period analysed in this study.

¹³The 2002 and 2005 LSMS surveys also contain information on the number of *weeks worked per year*, but unfortunately this variable is not available for the rest of the surveys and for this reason it is not used as an additional dependent variable in this study

¹⁴Monthly labor income is in real terms with base year 2007 using the national CPI provided by the Albanian National Institute of Statistics (INSTAT).

2.4 Empirical Strategy: Instrumental Variables

2.4.1 Son Preference in Albania and Siblings Sex-Composition

In this context, I exploit Albanian parental preference for having sons combined with the siblings sex-composition as an exogenous source of variation in the fertility decision. For this empirical strategy to identify the parameter of interest, the instrument sibling sex-composition must satisfy two basic conditions: (i) must be strongly correlated to the fertility choice (ii) it must have no correlation with factors directly affecting parental labor-force participation other than through its impact on fertility.

Several demographic analyses have confirmed that son preference is a distinctive feature of Albania's population ([UNFPA, 2012](#); [Gjonça, Aassve and Mencarini, 2008](#)). It is well documented that if the first child of an Albanian couple is a daughter, they will try to have at least one more child ([Zickel and Iwaskiw, 1994](#)). Son preference in Albania is partially driven by the desire and need to perpetuate the family name, but is also reinforced by a collection of customs that make a son desirable economically. For instance, Albanian sons are considered more economically productive than daughters because in general they obtain higher wages.

Old-age support from mature sons is another influential factor in parental decisions to have sons in Albania. They are still more preferred than daughters because they provide a greater source of protection and expected support for the parents at old age, a need reinforced by the uncertainties of the economic and social environment since the exit from the communist regime. Moreover, Albanian tradition and culture dictates that sons should provide financial support and protection for parents. In particular, Albanian custom obliges the youngest son (and his wife) to take care of his parents in their old age - indeed the youngest son is referred to in Albania as "the son of old age" ([King and Vullnetari, 2006](#); [King, Mata-Codesal and Vullnetari, 2013](#)).

On the contrary, the obligation of daughters is much weaker, in part because they are considered as transient members of their native family and also because they are thought to be responsible for their husband's parents. The newly married wife automatically switches her family to become part of her husband's family, and also part of that family's system of honour and economic support. According to custom, and in the absence of social insurance, parents are better served by ensuring the presence of a son and daughter-in law than the presence of a

daughter and a son-in-law.

Given this patriarchal Albanian context, the idea behind the identification strategy is that parents whose first two children are girls should have higher fertility in pursue of the male child compared to parents with other sibling's sex-compositions. Similarly, parents whose two first children are boys should be less likely to have an additional child because they have already reached the goal of having at least one son.

Therefore, a set of dummy variables indicating whether a mother's first two children are *two girls*, *two boys* or whether they have the *same sex* can be used as instrumental variables for higher order fertility. This identification strategy is akin to [Angrist and Evans \(1998\)](#), but since parental sex preferences in the Albanian context are different from the ones in the United States, I argue that the instrumental variable (*two girls*) should be more relevant for influencing progression to higher parity compared to the instrument *same sex* used by the authors. However, as the phenomenon of parental preference for balancing the sex-composition of the children has been widely documented in previous studies, I also use this alternative source of exogenous variation for comparative reasons.

2.4.2 The Relevance Condition

Table 2.4 reports raw estimates of the impact of child sex and the sibling's sex-composition on fertility. Panel A presents sex preference in households with at least one child, conditional on the sex of the first child. The figures show that around 50,6% of mothers have a boy as a firstborn, while approximately 49,3% correspond to mothers whose first child is a girl. The next three columns show the fraction of mothers with either a boy or a girl as a firstborn that had a second birth for the entire sample and by rural/urban area. It can be observed that mothers with a female firstborn child are more likely to have a second birth than those with a male firstborn child. The difference is statistically significant at the 1% level and is bigger in rural areas. This estimate suggests that subsequent fertility is a function of the sex of this first child and it is the first evidence of son preference in Albania.

Panel B of Table 2.4 documents parental sex-mix preference in households with at least two children, conditional on the sex of the first two children. Estimates in column (1) suggest that Albanian mothers with two children of the same sex are much more likely to have a third child than those with a mixed-sex siblings composition. More precisely, only 34,7% of mothers with

a mixed sex siblings composition have a third child compared to 48.7% of mothers with a same sex siblings composition. This difference is statistically significant at the 1% level and rises up to 14 percentage points, which is much bigger than the difference found in previous papers.

Furthermore, the correlation between sibling's sex-composition and the probability of additional childbearing is even stronger for mothers whose two first children are girls, reaching a statistically significant difference of 26 percentage points compared to mothers with a different sex-composition. For instance, estimates in column (1) indicate that around 61.4% of Albanian mothers with two girls have a third child compared to only 36.4% for those with two boys or even less for those with a mixed sex combination. Columns (2) and (3) indicate that fertility rate is higher in rural areas compared to urban areas and that parental preferences for sibling's sex-composition differ slightly by rural-urban decomposition, but the same pattern remains in both areas. These correlations suggest that Albanian parents have a clear preference for sons, which is slightly more pronounced in rural areas compared to urban areas.¹⁵

Overall, even though the total fertility rate has declined during the period studied, fertility patterns following *two girls* demonstrate the persistence of parental preference for sons. This results also indicates that, for Albanian parents, balancing the sex-composition of their children is not as preferred as having sons because if they did, they should want to have a girl if they already had two boys and this does not seem to be happening.¹⁶

2.4.3 The Exclusion Restriction

In addition to the relevance condition, the exclusion restriction requires that the sibling sex-composition has no correlation with other factors directly affecting labor-force participation other than through its impact on fertility. This condition means that it is of primary interest that the instrument has an as close a resemblance as possible to a random assignment. Basically, sibling sex-composition is supposed to be unrelated to any unobserved factors that might

¹⁵Figure 2.4 shows a graphical representation of Table 2.4 each one of the four databases used in this study. Mothers who give birth to a girl followed by another girl have a much higher fertility than those who give birth to mixed sex siblings or two boys in a row. On the contrary, mothers who give birth to *two boys* in a row are less likely to have an additional child compared to those with an otherwise sex-mix.

¹⁶It is important to notice that the relationship between sibling's sex composition and further childbearing is confirmed in each one of the databases used in this study (LSMS and DHS surveys). This is important because, unlike the 2005 and 2012 LSMS surveys where information about fertility is partly based on my own household matching between the children and their corresponding parents, the 2002 LSMS and 2008-09 DHS contain detailed fertility histories for each parent, including information on the dates of birth and sex of each child among others. This serves as a cross check for the relevance condition of the instrument. Tables 2.D.2 and 2.D.3 in the Appendix report estimates of the impact of child sex and sibling's sex-composition on fertility for each one of the four surveys.

affect parental labor supply because in essence it is virtually randomly assigned by a biological natural phenomenon.

Even though the sex of a child is plausibly randomly assigned, there exist several concerns that might put this statement into doubt and threaten the identification strategy of this analysis. One important concern is the prevalence of the sex-selective abortions in Albania. In view of a manifest preference for sons, the ratio of men to women has recorded a significant rise since the mid-1990s due to the legalisation of abortion and also the modernization of the available reproductive equipment, reaching 114 male births per 100 female births around 2005 (UNFPA, 2012). Hence, if the parents who prefer to have sons or the mothers who want to return to work are willing to engage in sex selection, the instrument will not be randomly assigned. Despite the fact that Albania has a high male-female ratio, the evidence from the sample of mothers and fathers used in this paper suggests that the sex-ratio at first and second order births is very close to the natural rate, which makes this issue a minor concern (see Tables 2.1 and 2.4).¹⁷

Another simple way to check whether the instrument is as good as random is to examine whether parents differ in demographic characteristics by the sex-compositions of their two first children (Angrist and Evans, 1998; Agüero and Marks, 2008). The idea is that if there is no correlation between the instrument and other variables affecting parental labor supply, then there should be no systematic differences in demographic characteristics between parents who have same-sex and mixed-sex sibling compositions, and similarly between those who have two girls and another sibling composition or two boys and otherwise. Examples of such variables are *age*, *age at first birth*, *years of education*, *rural area* and *religion*.

Table 2.5 reports the difference in means in demographic characteristics between parents (mothers and fathers) with different sibling's sex-composition (same-sex, two girls and two boys). Columns (1) and (4) indicate that mothers (fathers) with children of the same sex and those with children of mixed-sex, have similar demographic characteristics before the treatment, that is, the arrival of the third child. The rest of the estimates in Table 2.5 show that in general there are no statistically significant differences in demographic characteristics between parents

¹⁷Several methods to discriminate among unborn girls co-exist today in the world. The most recent methods are based on pre-conception selection and require access to the elaborate equipment necessary to perform sperm sorting, pre-implantation genetic diagnosis (PGD) or in-vitro fertilization (IVF). Cost and accessibility factors, however, restrict these technologies to developed countries and to the most affluent populations. These services were not generally offered by the public health service in Albania during the period studied. On the contrary, sex selective abortions offer a far easier and more accessible route to avoid female births (UNFPA, 2012).

with two boys and those with another sibling composition, and also between those with two girls and otherwise, except for the variable *age*. Parents with two girls and just the fathers with two boys seem to be slightly older than those with another sibling composition. The magnitude of these differences represents only 1% of the sample means of age for mothers and fathers, which is quite small. However, the variable *age* is used as a control in all the specifications of the paper in order to avoid any concerns.

Another possible threat to the validity of the exclusion restriction in this identification strategy is put forward by [Rosenzweig and Wolpin \(2000\)](#). They argue that having mixed sex siblings may violate the exclusion restriction by directly affecting both the marginal utility of leisure and child rearing costs and, thus, labor-force participation. For example, if spending on various child-related goods is different for parents who have a son versus those who have a daughter, the analysis will be contaminated by the direct impact of a child's sex on childcare costs. Using Indian data, they find that expenses for clothing of the third child are significantly lower if the older siblings are of same sex. They attribute this effect to "hand-me-down" savings, which are more likely to arise when there are children of the same sex in the household for items such as clothing or footwear. Since these items represent a sizeable fraction of the Indian household's expenditures (11% of the household income), they note that the sex composition of children plausibly alters female labor supply through mechanisms other than through fertility change alone.

Table 2.6 reports data on budget shares of child-related goods and mean household expenditure differences by sibling sex composition for Albania.¹⁸ The statistics show that Albanian households devote about 55% of their budget to food, 5.4% to health, 4.6% to clothing and 2.5% to children's education. In the clothing category, children's clothing expenditure represents only 2% of the entire budget. This estimate seems to be too small compared to the Indian case in order to account for a meaningful reduced form relationship between the sibling sex composition and parental labor supply.¹⁹ In addition, the evidence in Table 2.6 shows

¹⁸The data used for analyzing the expenditure patterns in Albania come only from the 2002, 2005 and 2012 Albanian LSMS surveys. Unfortunately, there are no available expenditure data on the 2008 DHS survey. The LSMS surveys contain very detailed information on household food and non-food expenditures. These databases have also the advantage of making possible the separation of child clothing and education expenditures from adult's expenditures.

¹⁹[Cruces and Galiani \(2007\)](#) make a similar argument for the cases of Argentina and Mexico. They indicate that sex composition is unlikely to have a noticeable effect on clothing and footwear expenditure for these Latin American countries due to the fact that only around 4.8% of the budget was used for clothing and footwear in the year 2000 in Mexico, while 6.7% for the case of Argentinean households (for all members).

that expenditure patterns of households in Albania are not significantly affected by the sex composition of children.²⁰ For instance, in one of the cases where the difference between budget shares is statistically significant, the sign contradicts the presence of economies of scale, since households whose first two children are girls spend a higher share of their budget on clothing. Also, even though Albanian parents with same-sex siblings seem to spend a lower proportion of their budget in children's education, the magnitude of this difference is too small to make a real impact, given the fact that children's education represents only 2.5% of the entire household budget. These statistics suggest that there is not a clear expenditure pattern of the Albanian parents by sibling sex composition. Therefore, the instrument seems not to be related to any indirect income effects that might question its exogeneity.²¹

Lastly, it can be argued that another possible threat to the exogeneity of the instrument is the fact that having two girls as the oldest children can make it easier for the parents to increase their labor supply as the older girls can do household chores and take care of the younger siblings. However, it is worth pointing out that this may be a bigger concern for the results of the mothers, but is unlikely to explain results for the fathers.²² In addition, parents may practice selective neglect of children based on gender. In the light of the patriarchal culture of Albania, parents could value taking care of two girls less and decide to keep working. Or, they may be less reluctant to give up a job when they have two girls, anticipating that they may not have a son to take care of them when they are old. Therefore, having two girls might have a direct effect on parental labor supply. In order to deal with this issue, in the estimation strategy I control for the gender of the first and second child.

²⁰Bütikofer (2010) contributes to the debate over the usage of the same-sex sibship as an identifying instrument by analyzing whether families with same-sex siblings composition face larger economies of scale in consumption in a variety of countries. She finds no significant differences between the estimated equivalence scales of families with different siblings sex composition in richer countries including the case of Albania. This analysis provides further evidence that household economies of scale do not crucially differ with siblings sex-composition.

²¹Huber (2012) developed a test to assess the validity of an instrumental variable in just-identified models and applied it to the Angrist and Evans's database finding evidence for the validity of *same-sex*.

²²In order to further address this concern, I also check the robustness of the main results for the mothers when the sample is limited to families where the oldest child is less than 13. As in Albania the average age for finishing primary school is 13, it is reasonable to argue that girls would not be asked to take care of siblings if they are younger than 13 years old. In the Appendix, estimates in Table 2.D.5 indicate that the main results remain unchanged after limiting the sample of mothers to those whose oldest child is younger than 13 years old.

2.4.4 The Econometric Framework

Potential Drawbacks of OLS

This section discusses the ordinary least square (OLS) and the instrumental variables (IV) techniques and presents the regression models relating parental labor-force participation and fertility. To begin with, I examine the effect of fertility on parental labor-force participation by using OLS through the following linear model:

$$LFP_{ijs} = \beta_0 + \beta_1 F_{ijs} + \varepsilon_{ijs} \quad (2.1)$$

Here, LFP_{ijs} is a measure of labor-force participation of the parent i residing in district j and observed in survey year s (i.e. dummy indicator for *worked in last 7 days* or *hours worked per week* among others); F_{ijs} is the endogenous fertility variable measured through the number of children ever born per parent i ; and ε_{ijs} is the error term associated with unobserved heterogeneity for the parent. The parameter of interest, β_1^{OLS} , represents the mean effect of having an additional child (third or higher order) on parental labor-force participation.

It is of most importance to notice that when we estimate the model in equation 3.2 by ordinary least squares (OLS), the estimator β_1^{OLS} is likely to be biased due to reverse causality and omitted variables. In particular, the magnitude of β_1^{OLS} is likely to be biased upwards due to unobserved ambition or ability which positively influences the outcome variable but is probably negatively correlated with having an additional child.

Instrumental Variables Estimation: LATE

In order to disentangle the causal mechanism linking fertility and parental labor supply, I use sibling sex-composition as an instrument that induces plausibly exogenous variation in F_{ijs} (the number of children ever born). Under reasonably general assumptions (independence and monotonicity), the estimate β_1^{IV} captures the local average treatment effect (LATE), first discussed by Imbens and Angrist (1994). The β_1^{IV} estimate can, then, be interpreted as the average effect of F_{ijs} on LFP_{ijs} for those parents whose fertility has been affected by the sex-mix of their previous children. Following the terminology in Angrist and Evans (1996), in order to better understand for which subgroup of parents with two children the average treatment effect can be consistently estimated, it is useful to classify them into the following 3 sub-populations:

those who will have a third child even following a son (*always takers*), those who will never have a third child even following two daughters (*never takers*), and those who will have a third child following two daughters but would otherwise stop (*compliers*).

Hence, β_1^{IV} can consistently estimate the average effect for individuals who have one more child because their first two children are girls (i.e. compliers) provided that the instrument satisfies monotonicity. Basically, what is needed to be assumed is that there are no *defiers*, those that change their behaviour in the opposite direction due to the instrument. In other words, monotonicity requires that having never had a son only makes one more likely to have a third child, a reasonable assumption given the pervasive son preference in Albania.²³ The IV approach fails to identify the effect among two sub-populations: the *always takers*, who generally have lower costs to childbearing than compliers, and the *never takers*, who will generally have higher costs to childbearing than compliers.

Therefore, the second-stage regression model that links the endogenous fertility measure \hat{F}_{ijs} with labor supply variables for the parents is the following:

$$LFP_{ijs} = \beta_0 + \beta_1 \hat{F}_{ijs} + \beta_2 b_1 + \beta_3 b_2 + \beta X'_{ijs} + \mu_s + \lambda_j + \varepsilon_{ijs} \quad (2.2)$$

where LFP_{ijs} measures labor-force participation of the parent i residing in district j and observed in survey year s ; \hat{F}_{ijs} is the endogenous fertility variable measured through the number of children ever born per parent i ; X'_{ijs} is a set of control variables that are plausibly exogenous to fertility, such as *parent's age*, *parent's age squared*, *parent's age at first birth* and *parent's years of education*; b_1 and b_2 are indicators for the sex of the first and the second child of parent i , respectively, and ε_{ijs} is the error term associated with unobserved heterogeneity for the parent. The variables b_1 and b_2 are included to control for potential additive effects of child gender, which might affect labor supply for reasons other than fertility. For example, as mentioned in Angrist and Evans (1998), this effect could arise if parents behave differently towards boys and girls, or whether a father's commitment to the family is contingent on the sex of the child. For this reason, the sex of the first-born and the second child, b_1 and b_2 , are included in the equation to minimize any omitted variable bias caused by additive effects of child sex.

²³ However, when replying upon siblings-sex composition as an instrument for fertility, the condition of monotonicity might be violated. Even though son preference is well documented for the case of Albania, a possibly minor fraction of parents might still prefer to have at least two children of the same sex (e.g. three boys), such that monotonicity would not hold. Following De Chaisemartin (2017), in the Appendix 2.B, I show that the 2SLS estimator is still valid, even under the presence of defiers, provided the "compliers-defiers" condition is satisfied.

I also control for the gender of the first child in order to address potential concerns about the exclusion restriction related to the fact that having two girls might have a direct affect on parental labor supply.

Initially, the first-stage regressions for the just-identified models that link the potentially endogenous fertility variable to only one instrument are the following:

$$\hat{F}_{ijs} = \alpha_0 + \alpha_1(\text{same sex})_{ijs} + \alpha_2 b_1 + \alpha_3 b_2 + \alpha X'_{ijs} + \mu_s + \lambda_j + v_{ijs} \quad (2.3)$$

$$\hat{F}_{ijs} = \alpha_0 + \alpha_1(\text{two girls})_{ijs} + \alpha_2 b_1 + \alpha_3 b_2 + \alpha X'_{ijs} + \mu_s + \lambda_j + \eta_{ijs} \quad (2.4)$$

$$\hat{F}_{ijs} = \alpha_0 + \alpha_1(\text{two boys})_{ijs} + \alpha_2 b_1 + \alpha_3 b_2 + \alpha X'_{ijs} + \mu_s + \lambda_j + \zeta_{ijs} \quad (2.5)$$

where the sibling sex-composition instrument is measured by three dummy variables, *same sex*, *two girls* and *two boys*, indicating whether the sex of the first child is the same as that of the second child. Given the Albanian parental preference for sons, F_{ijs} is expected to be positively correlated with *same sex* and *two girls*, but negatively correlated with *two boys*. Both the first and second stage regressions include fixed effects for districts λ_j and survey years μ_s . Also, in order to avoid potential biases in the estimation of the standard errors, an arbitrary structure of covariance is allowed by computing clustered robust standard errors at the primary sample unit (PSU) level.²⁴

Furthermore, since the *same sex* binary variable is easily decomposed into two variables indicating the sex composition of the first two children, namely *two boys* and *two girls*, an over-identified model with these two instruments and one endogenous variable can be estimated by the following first-stage regression:

$$\hat{F}_{ijs} = \alpha_0 + \alpha_1(\text{two girls})_{ijs} + \alpha_2(\text{two boys})_{ijs} + \alpha_3 b_1 + \alpha X'_{ijs} + \mu_s + \lambda_j + v_{ijs} \quad (2.6)$$

In this specification, either b_1 or b_2 must be dropped from the list of covariates because b_1, b_2 , *two boys* and *two girls* are linearly dependent. Thus, I chose to drop b_2 . It is important to clarify that the results are not sensitive to this choice, or to the elimination of both b_1 and b_2 , as is shown in the first-stage results presented in the following section.

²⁴The primary sampling units (PSUs) are geographically defined area units, which are selected with a probability proportional to size and they are very similar to a village level unit. The PSUs are already constructed in all the LSMS surveys and also in the DHS.

2.5 Results

2.5.1 First-Stage

Apart from the raw estimation analysis used to test the relevance condition of the instruments, this section examines in more depth this condition through the first-stage estimation. Table 2.7 reports the first-stage results linking sex-mix and fertility, where columns (1) to (3) show the results for the just-identified models without covariates, columns (4) to (6) show the results after adding covariates, and column (7) shows the results for the over-identified model with covariates. The top half of the table (Panel A) gives the estimation results for the sample of mothers, while the bottom half (Panel B) gives the results for the sample of fathers (husbands of the married women).

The results from the first-stage estimation in column (1) suggest that mothers with two children of same sex have on average 0.22 more children than mothers with children of mixed sex. The instrument *same sex* seems to be much stronger for the case of Albania compared to the one reported by Angrist and Evans (1998) and Cruces and Galiani (2007), who find that mothers with two children of the same sex have only 0.07 more children on average.²⁵

Moreover, the results in columns (2) and (3) show that the effect of sibling sex composition on fertility is even much stronger if the first two children are girls, but weaker if they are boys. More specifically, having two girls in an Albanian household increases on average the number of children ever born by 0.47. This coefficient is twice as large as the coefficient for *same sex*, which means that the instrument *two girls* is better capturing the exogenous variability in fertility. On the contrary, having two boys decreases the number of children ever born by about 0.16, which suggest that the parents are less likely to have additional children following sons.

The next three columns (4) to (6) present similar results after controlling for several covariates such as *parent's age*, *parent's age squared*, *parent's age at first birth*, *parent's years of*

²⁵Following Angrist and Evans (1998), if the indicator variable *More than 2 children* is used as a dependent variable in the first-stage estimation, I obtain very similar results which are significant at the 1% level and in each case the F-statistic on the excluded instruments is higher than 10. Compared to Angrist and Evans (1998), who find that the effect of the *Same sex* instrument to be around 6 percentage points, my first-stage estimates using the indicator *More than 2 children* seem to be much stronger. For instance, I find that Albanian mothers with two first children of the same sex are estimated to be 13.1 percentage points more likely to have a third child. This estimate rises up to 25.5 percentage points if the first two children are girls, which is very big. On the contrary, I find that mothers with two boys are on average 7.7 percentage points less likely to have a third child. These results are available upon request.

education and also whether the first and the second child are boys, *boy1st* and *boy2st*. After the inclusion of these covariates, the instruments *same sex* and *two girls* still have a positive and statistically significant impact on fertility and the magnitude of the estimates remains quite unchanged. More specifically, having children of the same sex increases fertility by 0.22, but having two girls leads to even a bigger increase which reaches up to 0.43.²⁶ All these results are significant at the 1% level of significance and in each case the F-statistic on the excluded instruments is higher than 10, considered to be the rule of thumb threshold by [Stock, Wright and Yogo \(2002\)](#).

Conversely, the first-stage estimate in column (6) that uses *two boys* as an instrument becomes zero in magnitude and it is not statistically significant anymore. In addition, the R-squared and the F-statistic of the excluded instrument from this specification are zero, which clearly indicates that *two boys* is a weak instrument in the context of Albania. For this reason, it will not be used in the second-stage estimation.

Lastly, the first-stage estimates of the over-identified model in column (7), which uses both *two girls* and *two boys* as instruments, also indicate that having two girls increases the number of children by 0.43 on average and this result is statistically significant at the 1% level. On the contrary, having two boys does not seem to produce an effect on the number of children and the result is also statistically insignificant, proving to be a weak instrument once more. Thus, the over-identified specification is neither a good candidate for the second-stage estimation.

As a whole, the first-stage results in Table 2.7 confirm once more that Albanian parents have a strong preference for having sons. In particular, having two girls increases the likelihood of having an additional child, while having two boys seems to have no effect on fertility after controlling for several covariates. Even though the instrument *same sex* seems to have a positive and statistically significant effect on fertility, it does not mean that Albanian parents prefer to balance their children's sex-composition because this variable is only capturing the variability generated by *two girls*. Therefore, the second-stage estimation will be based mainly on the exogenous variability of the instrument *two girls*. The *same sex* instrument will be used only for

²⁶This is also equivalent to an increase of 24 percentage points in the probability of having a third child when the indicator *More than 2 children* is used as a measure for fertility. This coefficient is three times bigger than the one found by [Angrist and Evans \(1998\)](#) -7.1 percentage points in the U.S.- and five times bigger than the one found by [Cruces and Galiani \(2007\)](#) -5.3 percentage points for married women in Argentina and 4.6 percentage points for married women in Mexico-.

comparative analysis.²⁷

2.5.2 Main Results: OLS and Second-Stage

This section presents the main results of the effect of fertility on parental labor-force participation. The OLS and second-stage results are estimated separately for the mothers and the fathers because women and men in Albania differ substantially in the type and amount of labor they supply. Tables 2.8 and 2.9 report OLS estimates in column (1) and two sets of IV estimates (i.e. *same sex* and *two girls* as instruments) in columns (2) and (3) for the sample of mothers and fathers, respectively.²⁸ All the specifications include the same set of control variables as in Table 2.7. The upper panel of these tables reports the estimates for the different measures of labor force participation (e.g. *worked in last 7 days*, *work off-farm*, *work on-farm*, *self-employed*), while the bottom panel reports estimates for *hours worked per week*, *second occupation* and *monthly labor earnings*.²⁹

The OLS estimates in Tables 2.8 and 2.9 indicate that the correlation between fertility and likelihood of working is very small and statistically insignificant for mothers, but positive and statistically significant for fathers. In particular, if the analysis is separated by type of work, the presence of an additional child has a negative and statistically significant correlation with parents's likelihood of working off-farm, being self-employed and also monthly earnings, but it has a positive and statistically significant correlation with both parent's likelihood of working on-farm. Fertility also seems to be positively correlated with father's likelihood of having a second occupation.

Although the OLS estimates suggest that parent's withdrawal from the labor market or their further participation as a consequence of an increase in family size depend on the type of job they are performing, they should be taken with caution because they could well be

²⁷The corresponding first-stage estimates for the sample of fathers presented in Panel B of Table 2.7 are almost exactly the same as the ones for the sample of mothers.

²⁸The first-stage specifications of columns (4) and (6) in Table 2.7, which use *same sex* and *two girls* as instruments for fertility, are the ones selected for the second-stage estimation because these are the ones that proved to be the strongest instruments after controlling for covariates.

²⁹The variables presented in the bottom panel of Tables 2.8 and 2.9 such as *hours worked per week*, *second occupation* and *monthly labor income* are only available in the Albanian LSMS surveys, but not in the Albanian DHS survey. For this reason, the sample of mothers is reduced to 4862 observations and the sample of fathers to 4530 observations. The first-stage estimates for these reduced samples are very similar to the ones showed in Table 2.7, but very slightly smaller in magnitude. Once again *two girls* and *same sex* prove to be valid instruments for fertility, while *Two boys* seems to be a weak instrument. The first-stage estimates for these samples are reported in Table 2.D.4 of the Appendix.

biased. After correcting for the endogeneity problem, the second-stage results tell a completely different story. Once *same sex* and *two girls* are used as instruments, some of the IV estimates of the effect of fertility on parental labor-supply change sign and magnitude compared to the OLS estimates.

When using *same sex* as an instrument for fertility, the IV estimates show that having an additional child (third order or higher) increases the likelihood of participating in the labor market for both parents, but most of these coefficients are statistically insignificant. One reason for this result may be linked to the fact that *same sex* does not capture the exogenous variability of fertility as precisely as the *two girls* instrument in the context of Albania, which may probably lead to imprecise estimates.

The IV estimates using *two girls* as instrument indicate that there is a positive effect of fertility on mother's likelihood of working off-farm and this coefficient is statistically significant at the 5% level. More specifically, having an additional child (third order or higher) increases on average a mother's likelihood of working off-farm (+5 p.p), which represents a 30% increase in the average participation rate of mothers's work off-farm. But, the presence of an additional child does not seem to have a statistically significant impact on mother's likelihood of working on-farm or being self-employed. In addition, the two sets of IV estimates indicate that having an additional child seems to also have a positive and statistically significant impact on mother's hours worked per week. To be more precise, each additional child increases the working time of a mother by about 4.4 - 4.6 hours per week. This coefficient is quite big given that it represents 26% of the average working time of the mothers in the sample.³⁰

The labor-supply effects estimated using the sample of fathers are quite similar to those for the mothers, except for the outcome *hours worked per week*. The IV estimates in Table 2.9 using *two girls* as instrument indicate that having an additional child increases father's likelihood of working off-farm (+6.7 pp) and also their likelihood of having a second occupation (+4.0

³⁰A distinguishable characteristic of Albania during the years of this study is the combination of low fertility with low female participation rates. Female labor force participation has stagnated or even decreased during the last two decades despite a significant drop in fertility rates. The combination of these two stylized facts does not necessarily mean that there exists an inverse relation between fertility and female labor supply, as it has been stated in various developed countries. A study from [Da Rocha and Fuster \(2006\)](#) corroborates in part this mechanism. These authors find that labor market frictions determine how employment and fertility are associated across economies. In particular, they find that fertility and employment are positively associated across economies with low employment ratios (low probability of finding a job). On the other hand, they find that fertility and employment are negatively associated across economies with high probability of finding a job. Their findings also suggest that the sign of this relationship could change as the job-finding rate of females increases.

pp). These two coefficients are statistically significant at the 5% and 10% level of significance, respectively. For practical reasons, I do not present the coefficients on all the controls. However, I do present the coefficients for *boy1st* and *boy2nd* in Tables 2.D.6 and 2.D.7 of the Appendix 2.D in order to test whether the sex of the children affects parental labor supply for reasons other than fertility. In all the specifications in Tables 2.D.6 and 2.D.7, the coefficients of *b1* and *b2* are statistically insignificant for all the relevant outcomes of the analysis, which indicates that the instruments do not have a direct effect on parental labor-force participation.

Therefore, a plausible interpretation of the positive effect found between fertility and parental labor supply in Albania is that having an additional child induces parents with at least two children to work more for somebody rather than a household member, which means that they increase their labor supply in the non-agricultural sector. But, why would parents decide to augment their chances of being employed in a kind of job that might be considered as more stable or with a greater formality as a response to an increase in family size?

One way of interpreting these findings is by looking at the payment of a job as a bundle of services, such as wages, schedule flexibility, social status, security etc. Thus, an increase in fertility would not only increase the cost of time intensive activities but also the attractiveness of some jobs (scheme of payments) that are compatible with bigger family size versus others. For the case of mothers, this type of job might be more temporary than permanent due to the fact that part-time is easier to combine with motherhood tasks. Conversely, for the case of fathers this type of job might be more permanent than temporary due to the better benefits, such as higher earnings among others.

2.6 Discussion

2.6.1 Two plausible mechanisms

The regression analysis implies a significant and positive effect, as Albanian parents, men and women alike, are more likely to be working off-farm as a consequence of further childbearing (third order or higher) relative to parents with just two children.³¹ Through which channels

³¹These results confirm the common view in labor economics that children are an important determinant of parental labor supply. However, in contrast to most of the findings in the previous literature, children might not only restrict the opportunity of parents to participate in the labor market but as well create an incentive to search for employment. In order to shed some light on the mechanisms behind these results, a simple theoretical model that links childcare costs and family structure with fertility and parental labor supply is derived in the Appendix

does this positive effect operate? In this section, I propose two main plausible explanations: first, childcare provided by non-parental adults (e.g. grandparents) in extended families; and, second, greater financial costs of maintaining more children.

Extended families are composed by several generations consisting of grandparents, aunts, uncles, and cousins all living nearby or in the same household. A typical example is a married couple that lives with either the husband's or wife's parents. Thus, if indeed the parents have help in the household from the grandparents, they might not be forced to leave the workforce when they have an additional child. Under this scenario, the effect of fertility on parental labor supply depends on the extent of substitutability between the time spent by grandparents and parents in childcare and household chores. In response to an increase in fertility, grandparents may adjust their time allocation by supplying more time in childcare and household chores and, thus, parents are freed from such tasks and not forced to drop out of the labor market.³²

Indeed, several demographic studies have documented the rather particular nature of the Albanian family structure, characterized by extended families (e.g. [Gjonça, Aassve and Mencarini, 2008](#); [King, Mata-Codesal and Vullnetari, 2013](#))³³ In terms of care duties, caring for one's older parents is also strongly-felt duty in Albania. According to the long-established tradition and custom, middle-aged and elderly parents are to live with one of their sons, usually the youngest, and his wife. Indeed the youngest son is referred to as "the son of old age", whose role and duty is to look after his parents in their later years. Moreover, to respect and care for one's parents is not only a duty but highly honourable in Albanian society ([King and Vullnetari, 2006](#)). Therefore, in the context of Albania where extended families are quite prevalent and professional childcare is rather rare and expensive, non-parental household members such as grandparents may be the perfect candidates for taking care of the children.

2.C.

³²Several previous studies have examined the effect of family structure on female labor supply and they have found that the presence of grandparents in the household increases the labor supply of women who have children (e.g. [Wong and Levine, 1992](#); [Sasaki, 2002](#); [Gong and Van Soest, 2002](#); [Maurer-Fazio et al., 2011](#); [Compton and Pollak, 2014](#); [Garcia-Moran and Kuehn, 2017](#); [Posadas and Vidal-Fernández, 2013](#); [Shen, Yan and Zeng, 2016](#); [Landman, Seitz and Steiner, 2017](#)).

³³During most of the 20th century, the extended family in Albania was the basic single residential unit and economic entity. Basically, it was formed by a married couple with their married sons and their offspring, as well as any unmarried daughters. The size of these families was very large, and by the end of WWII some numbered as many as 60 to 70 members ([Gjonça, Aassve and Mencarini, 2008](#)). Even though the profound post-communist transformation and emigration have started to break down this tradition, extended families with three generations are still very common in rural Albania. For instance, the datasets used in this paper indicate that around 38% of the households live in extended families. This number is slightly higher in rural areas where it increases up to 42%.

Similarly, the positive labor supply effects attributable to an increase in family size can also be rationalized on the bases of greater financial costs of feeding more children. Naturally, fertility can affect parental labor supply in two directions: on the one hand, it can provoke parents to withdraw from the labor market due to child caring and higher child costs; on the other hand, given that modern child caring and child costs are assumed to be very expensive, an increase in fertility may reduce the standard of living in the household and also stimulate parents to participate in the labor market in order to finance basic expenditures on their children.

Therefore, for the case of Albania the latter constraint might be stronger than the former. It is possible to suggest that poorer, less educated and younger parents are the ones that need to provide more for their families compared to richer, more educated and older parents. This mechanism could be particularly true for developing countries such as Albania because in this context households generally act under tighter budget constraints where a large share of income is devoted to food and other basic needs. For instance, in Table 2.6 it can be observed that Albanian households dedicate more than half of their budget to food expenditure (54.8 %).

2.6.2 Heterogeneity Analysis

In order to shed light on the mechanisms proposed, this section examines whether the effect of fertility on parental labor-force participation may be sensitive to certain sub-populations.³⁴ For this purpose, I perform an heterogeneity analysis by exploring whether the effect of fertility on parental labor supply varies with: *parent's education level*, *parent's birth cohort* and *family structure*. Moreover, it seems essential from a policy point of view to be able to identify which one of the sub-populations has the greatest response to the effects of fertility on parental labor market outcomes, in particular in relation to income or wealth level.

First, Table 2.10 presents the heterogeneity analysis of the effect of fertility on parental labor supply by *parent's educational level*. For this analysis, all the parents in the sample are divided into two groups by taking as a reference their primary school completion: below or

³⁴Gronau (1988) and Angrist and Evans (1998) suggest that the relationship between labor supply and childbearing is likely to differ by women's educational level. They find that the labor supply of less educated women is more sensitive (more negative) to the presence of children than the labor supply of more educated women. Angrist and Evans (1998) also explore how the female labor-market consequences of childbearing vary with the earnings of the husbands. They find that the effect of fertility on women's labor supply declines in magnitude (less negative) with husbands's earnings.

completed primary level (≤ 8 years of schooling) and above primary level (> 8 years of schooling). This separation is based on two simple facts: the average schooling years for mothers and fathers with at least two children in Albania are 10 and 11, respectively; and parents who have completed primary school (8 years) represent about 50% of the observations in each sample.

As can be seen in Table 2.10, the IV coefficients on *hours worked per week* for mothers belonging to the lowest educational level have a positive sign and are statistically significant. In addition, the IV coefficient on *second occupation* for fathers belonging to the lowest educational level also has a positive and statistically significant effect. On the contrary, the labor supply effects for parents with a higher level of education are imprecisely estimated. Thus, it seems reasonable to argue that fertility increases labor-force participation for less educated parents and, hence, possibly less economically well parents in order to respond to an increased demand in the household, namely need for income.

Second, Table 2.11 explores whether the effect of fertility on parental labor-force participation is likely to vary with *parent's birth cohort*. The birth cohort threshold is set at the median birth year in order to allow the division sample into two subsamples of similar size. On the one hand, the sample of mothers is divided between mothers born before 1972 (included) and mother's born after 1972. On the other hand, the sample of fathers is divided between fathers born before 1966 (included) and those born after 1966.³⁵

Mothers of younger cohorts show a positive and statistically significant effect of fertility on *hours worked per week* at the 1% level. Also, the IV estimates for younger fathers show a positive and statistically significant effect on father's likelihood of having a *second occupation* as a consequence of childbearing. On the contrary, the coefficients on fertility for mothers and fathers of older cohorts are lower in magnitude and statistically insignificant. Hence, I can confirm that there exists a degree of heterogeneity in the link between fertility and labor market outcomes by parent's birth cohort, where the positive effect seems to be driven mostly by younger parents. This result also suggests that the positive effect of fertility on parental labor-supply is not driven by parents whose older children in the household are taking care of the younger siblings, which is reassuring in terms of the exclusion restriction.

³⁵The average year of birth in the sample of mothers aged 20 to 49 years old with at least two children is 1972, while the average year of birth in the sample of fathers (husbands or partners of the married mothers) is 1966. Mothers born before 1972 (included) represent 49.5 % of the sample, while mothers born after 1972 represent 50.5 % of the sample. Similarly, fathers born before 1966 (included) represent 50.2 % of the sample, while fathers born after 1966 represent 49.8 % of the sample.

Third, Table 2.12 reports the effect of fertility on parental labor supply conditional on the *family structure*. Given the Albanian context, it could be argued that extended families in this country are a given characteristic of society since they are already established by strong traditional and social norms. To give more evidence on this assumption, I use the 2002 Albanian LSMS which contains detailed information on the fertility history of mothers from old birth cohorts in order to calculate the proportion of parents in extended families living with their sons by their son's birth order. Indeed, Figure 2.5 shows that most of the parents in extended families live with their youngest married son (61.1 %).

However, there still remains the concern that the fertility decision and the co-residence decision might be jointly determined. Given the strong son preference in the context of Albania, grandparents may encourage parents with two daughters to have a third child in exchange for moving in and helping out in providing child care. In other words, grandparents may bargain with parents for having a third child, which could affect their willingness to co-reside and take care of the children. Also, grandparents may just be more willing to move in if the parents are more likely to have three children.

In order to test this, I perform a falsification exercise on the first-stage to check whether siblings sex composition is a significant determinant of the co-residence decision. The results are reported in Table 2.13 and the specification are identical to those reported in Table 2.7. The effect of siblings sex composition on the likelihood of living in an extended family is small and insignificant in all regressions. There is even a negative correlation between having two girls and living in an extended family, which contradicts the idea of co-residence in exchange for more children and childcare. Overall, these results suggest that children's sex mix is not a predictor for living in an extended family structure, which reinforces the idea that extended families are a product of a strong social norm.

For this heterogeneity analysis both samples (mothers and fathers) are divided into two types of households: those that have grandparents older than 50 years old living in the same household as the parents (*extended families*) and those that do not (*non-extended families*). Conforming to the first mechanism proposed, the IV estimates in Table 2.12 show that mothers living in extended families experience the largest positive effects of fertility on labor supply. In contrast, there is no statistically significant link between additional fertility and labor supply for mothers living in non-extended families. Another interesting and important result is that

the positive effect of fertility on father's labor supply is mostly driven by those living in non-extended families. Since parents in non-extended families do not count on childcare provided by grandparents or other relatives, it is very likely for the mother to stay in the household to take care of the children, whereas the father probably searches for a second occupation.

Last, in Table 2.14 I examine whether the effect of fertility on family expenditure per capita depends on family structure and the wealth level of the household. For this estimation, households are pooled together into five expenditure quintiles i.e. all households in quintile 1 belong to the lowest expenditure level. As one would expect, parents in poorer quintiles are more likely to be younger, to have children at a younger age, to have a third child and to live in rural areas, but they are less likely to live in an extended family. Moreover, having an additional child reduces the family expenditure per capita in all quintiles. This result clearly suggests that an increase in fertility leads to a tighter budget constraint in the household.

When these estimates are analyzed by family structure, I find that the negative effect of fertility on expenditure per-capita is smaller in magnitude (less negative) for extended families and it is only statistically significant in two quintiles. On the contrary, non-extended families experience larger negative effects of fertility on family expenditure per-capita and these results are statistically significant in all quintiles. Also, the estimates for non-extended families show a stronger negative effect in the lowest expenditure quintiles.

To sum up, these heterogeneity analysis results suggest that the positive labor supply effect can be attributed to parents which are less educated (poorer), younger and who live in extended families. These results seem to support the view that higher fertility associated with greater child costs and lower childcare costs provided by grandparents in extended families are two important mechanisms that influence the labor supply decision of parents in Albania.

2.7 Conclusion

The purpose of this paper is to investigate whether there is a causal effect of fertility on parental labor-force participation in the context of a developing country in the Balkans such as Albania. In order to address the endogeneity in the fertility decision, I exploit Albanian parental preference for having sons combined with the siblings sex-composition instrument proposed by [Angrist and Evans \(1998\)](#) as an exogenous source of variation in the fertility decision. Using a repeated cross-section of Albanian mothers and fathers with at least two

children, I find that Albanian parents have a strong and persistent preference for having sons. In particular, the results from the first-stage specifications suggest that having *two girls* increases on average the number of children ever born by 0.43, which is equivalent to an increase of 24 percentage points (almost 60% of the sample mean) in the likelihood of having an additional child.

While most of the previous studies on this topic have found a negative relationship between fertility and labor supply, especially for women's labor supply, there are also some studies that contradict these findings and indicate a positive relation (Porter and King, 2012; Priebe, 2010; Bloom et al., 2009). Hence, in the same line as the latter group of studies, the point estimates reported in this paper suggest that fertility has a positive effect on parental labor-force participation in the context of Albania.

More precisely, I find that having an additional child (third or higher order) increases on average a mother's likelihood of *working off-farm* by 5 percentage points and it also increases the working time by around 4.4 *hours per week* on average. Similarly relevant is the finding that fathers also change positively their labor-market behaviour in response to a change in family size. In particular, having an additional child increases father's likelihood of *working off-farm* by 6.7 percentage points, and also their likelihood of having a *second occupation* by 4 percentage points.

The second focus of this study is to shed light on the channels through which these results operate. I propose two plausible mechanisms behind the positive effect of fertility on parental labor supply. The first one suggests that childcare provided by non-parental adults (e.g. grandparents) may act as a substitute for parental childcare in developing countries, thus, allowing them to increase labor supply as a consequence of further childbearing. The second one suggests that households have to face a tighter budget constraint as a response to a increase in fertility which reduces the standard of living in the household and forces parents to participate in the labor market. Conforming to the mechanisms proposed, the heterogeneity analysis suggests that the labor-market consequences of fertility are more likely to be driven by poorer, less educated, and younger parents, who also happen to live in extended families.

Lastly, the findings in this paper might have important implications in terms of public policy. In many developed countries much attention has been given to the role of child care costs/subsidies, including direct provision of public pre-school, on female labor supply (Blau

and Robins, 1998; Connelly, 1992; Kimmel, 1998; Gelbach, 2002). The main finding of this paper implies that with lower childcare costs provided by the presence of grandparents in extended families but also greater child costs, parental labor supply is not reduced as a consequence of an increase in family size, suggesting that as an informal childcare provision mechanism, extended families play an important mitigating role for the time cost of children.

However, as family structure in developing countries becomes more nuclear with economic development, the time cost of children may increase for parents which may lead to further fertility decline and fewer opportunities for non-parental household members to help with childcare or household chores. In view of this decline in intergenerational co-residence over time and also the inadequate provision of public childcare in the context of Albania, one possible public policy intervention should be the implementation of child care subsidies including direct provision of day-care nurseries, public kindergartens and pre-schools in order to promote labor supply and sustainable economic development.

Chapter 2 - Figures and Tables

Table 2.1: Descriptive Statistics for Mother's Fertility in Albania (2002 - 2012)

Variables	Observations	Mean	Standard Deviation	Min	Max
<i>Sample: Mothers aged 20 to 49 years old with 2 or more children younger than 18 years old</i>					
Fertility (Number of children ever born)	7480	2.600	(0.871)	2	9
More than 2 children (=1 if mother had 3 or more children)	7480	0.418	(0.493)	0	1
First child boy (=1 if first child was a boy)	7480	0.492	(0.499)	0	1
First child girl (=1 if first child was a girl)	7480	0.508	(0.499)	0	1
Second child boy (=1 if second child was a boy)	7480	0.514	(0.499)	0	1
Two boys (=1 if first two children were boys)	7480	0.254	(0.435)	0	1
Two girls (=1 if first two children were girls)	7480	0.247	(0.431)	0	1
Same sex (=1 if first two children have the same sex)	7480	0.502	(0.500)	0	1
Mixed sex (=1 if first two children have different sex)	7480	0.497	(0.500)	0	1

Notes: The data used are the 2002 LSMS, the 2005 LSMS, the 2008-09 DHS and the 2012 LSMS

Table 2.2: *Descriptive Statistics for Mother's Labor-Supply in Albania (2002 - 2012)*

Variables	Observations	Mean	Standard Deviation	Min	Max
Mother's Labor Supply					
Worked in last 7 days	7445	0.444	(0.496)	0	1
Worked off-farm	7445	0.163	(0.369)	0	1
Worked on-farm	7445	0.217	(0.412)	0	1
Self-employed	7445	0.066	(0.249)	0	1
Second occupation	4903	0.014	(0.120)	0	1
Hours/week (total)	4903	16.60	(21.24)	0	112
Hours/week (off-farm)	4903	7.17	(16.20)	0	70
Hours/week (on-farm)	4903	7.52	(15.64)	0	70
Hours/week (self)	4903	2.35	(11.15)	0	112
Monthly labor income	4903	50166.4	(114456)	0	1217661
Other characteristics of mothers					
Age	7480	35.04	(5.691)	20	49
Age at first birth	7480	23.22	(3.489)	14	41
Years of education	7404	10.47	(3.135)	0	21
Rural	7480	0.482	(0.499)	0	1
Household size	7480	5.188	(1.478)	1	16
Extended family	7480	0.381	(0.485)	0	1
Marital status (married)	7480	0.980	(0.137)	0	1
Muslims	7480	0.816	(0.387)	0	1
Catholics	7480	0.081	(0.273)	0	1
Orthodox	7480	0.067	(0.251)	0	1
Coastal region	7480	0.264	(0.440)	0	1
Central region	7480	0.282	(0.450)	0	1
Mountain region	7480	0.291	(0.454)	0	1
Tirana region	7480	0.161	(0.367)	0	1

Notes: The data used are the 2002 LSMS, the 2005 LSMS, the 2008-09 DHS and the 2012 LSMS

Table 2.3: *Descriptive Statistics for Father's Labor-Supply in Albania (2002 - 2012)*

Variables	Observations	Mean	Standard Deviation	Min	Max
Father's Labor Supply					
Worked in last 7 days	5496	0.787	(0.409)	0	1
Worked off-farm	5496	0.393	(0.488)	0	1
Worked on-farm	5496	0.236	(0.424)	0	1
Self-employed	5496	0.197	(0.398)	0	1
Second occupation	4566	0.058	(0.234)	0	1
Hours/week (total)	4566	34.23	(23.06)	0	112
Hours/week (off-farm)	4566	18.50	(23.43)	0	99
Hours/week (on-farm)	4566	9.10	(18.71)	0	99
Hours/week (self)	4566	7.63	(19.09)	0	112
Monthly labor income (old leks)	4566	191841.9	(333208.7)	0	12000000
Other characteristics of fathers					
Age	5737	39.92	(5.855)	22	64
Age at first birth	5737	28.33	(4.177)	15	58
Years of education	5541	11.08	(3.426)	0	20
Rural	5737	0.478	(0.499)	0	1
Household size	5737	5.210	(1.424)	1	16
Extended family	5737	0.364	(0.481)	0	1
Marital status (married)	5737	0.998	(0.032)	0	1
Muslims	5737	0.798	(0.401)	0	1
Catholics	5737	0.082	(0.275)	0	1
Orthodox	5737	0.069	(0.254)	0	1
Coastal region	5737	0.262	(0.440)	0	1
Mountain region	5737	0.285	(0.451)	0	1
Tirana region	5737	0.164	(0.370)	0	1

Notes: The data used are the 2002 LSMS, the 2005 LSMS, the 2008-09 DHS and the 2012 LSMS

Table 2.4: *Fraction of Households that had Another Child by Parity and Sex Composition in Albania (2002- 2012)*

Sex of the first child in households with one or more children	Fraction of the sample	Fraction that had another child		
		All (1)	Urban (2)	Rural (3)
Panel A: Mothers aged 20 to 49 years old with 1 or more children younger than 18 years old				
(1) one girl	0.493	0.809 (0.005)	0.781 (0.008)	0.841 (0.007)
(2) one boy	0.506	0.764 (0.006)	0.752 (0.008)	0.777 (0.008)
Difference (2) - (1)	0.012 (0.010)	-0.044 (0.008)***	-0.028 (0.011)***	-0.063 (0.011)***
Observations	9509	9509	5049	4460
Sex of first two children in households with two or more children	Fraction of the sample	Fraction that had another child		
		All (1)	Urban (2)	Rural (3)
Panel B: Mothers aged 20 to 49 years old with 2 or more children younger than 18 years old				
one boy, one girl	0.237	0.352 (0.011)	0.265 (0.014)	0.441 (0.016)
one girl, one boy	0.260	0.343 (0.010)	0.247 (0.013)	0.449 (0.016)
two boys	0.254	0.364 (0.011)	0.277 (0.014)	0.458 (0.016)
two girls	0.247	0.614 (0.011)	0.501 (0.016)	0.736 (0.014)
(1) mixed combination	0.497	0.347 (0.007)	0.256 (0.009)	0.445 (0.011)
(2) both same sex	0.502	0.487 (0.008)	0.387 (0.011)	0.595 (0.011)
Difference (2) - (1)	0.005 (0.011)	0.140 (0.011)***	0.131 (0.014)***	0.150 (0.016)***
Observations	7480	7480	3874	3606

Notes: The data used are the 2002 LSMS, the 2005 LSMS, the 2008-09 DHS and the 2012 LSMS. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%. Standard errors are in parenthesis.

Table 2.5: *Differences in Means for Demographic Characteristics by Sex Composition*

Variables	Mothers (N=7480)			Fathers (N=5496)		
	Same-sex (1)	Two boys (2)	Two girls (3)	Same-sex (4)	Two boys (5)	Two girls (6)
Age	0.1252 (0.1316)	-0.1683 (0.1510)	0.3394 (0.1523)**	0.0045 (0.1546)	-0.3907 (0.1769)**	0.4085 (0.1795)**
Age at first birth	-0.0160 (0.0807)	-0.0414 (0.0926)	0.0206 (0.0934)	-0.1199 (0.1102)	-0.2101 (0.1362)	0.0546 (0.1281)
Years of education	-0.0186 (0.0728)	0.0020 (0.0835)	-0.0270 (0.0843)	0.0086 (0.0920)	-0.0392 (0.1054)	0.0520 (0.1069)
Rural	-0.0014 (0.0115)	0.0011 (0.0132)	-0.0030 (0.0133)	0.0122 (0.0131)	0.0241 (0.0150)	-0.0083 (0.0153)
Muslim	-0.0121 (0.0089)	-0.0109 (0.0102)	-0.0051 (0.0103)	-0.0082 (0.0105)	-0.0109 (0.0121)	0.0001 (0.0123)
Catholic	0.0051 (0.0063)	0.0057 (0.0072)	0.0009 (0.0073)	0.0065 (0.0072)	0.0078 (0.0083)	0.0007 (0.0084)
Orthodox	0.0020 (0.0058)	0.0085 (0.0066)	-0.0059 (0.0067)	0.0003 (0.0067)	0.0087 (0.0076)	0.0094 (0.0078)

Notes: Differences in means (mean of the relevant group minus mean of the rest of the population) and their standard errors (in parentheses). * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%. The sample of mothers consists of 7480 Albanian women aged 20 to 49 years old with two or more children aged 18 or younger. The sample for fathers consists of 5496 husbands of the married women with two or more children aged 18 or younger. The data used are the 2002 LSMS, the 2005 LSMS, the 2008-09 DHS and the 2012 LSMS

Table 2.6: *Differences in HH Budget Shares by Sex Child Composition in Albania (2002 - 2012)*

Variables	Budget Share (All)	Same-sex	Two boys	Two girls
	(1)	(2)	(3)	(4)
HH Food Expenditure (S)	0.548	-0.001 (0.004)	-0.008* (0.004)	0.006 (0.004)
HH Health Expenditure (S)	0.054	-0.003 (0.002)	0.002 (0.003)	-0.007** (0.003)
HH Clothing Expenditure (S)	0.046	0.001 (0.001)	-0.000 (0.001)	0.001 (0.001)
Children's Education Expenditure (S)	0.025	-0.001 (0.001)	-0.002* (0.001)	0.001 (0.001)
Children's Clothing Expenditure (S)	0.020	0.001* (0.000)	-0.000 (0.000)	0.002*** (0.000)
Children's Education Expenditure (S.p.c)	0.010	-0.001*** (0.000)	-0.000* (0.000)	-0.001* (0.000)
Children's Clothing Expenditure (S.p.c)	0.008	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Children's Clothes (S.p.c.)	0.005	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Children's Footwear (S.p.c.)	0.003	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)

Notes: Differences in means (mean of the relevant group minus mean of the rest of the population) and their standard errors (in parentheses). * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%. The data used are the 2002 LSMS, the 2005 LSMS and the 2012 LSMS. The sample consists of 7480 Albanian mothers aged 20 to 49 years old with two or more children aged 18 or younger. (S) refers to the share of a particular expenditure over total expenditure, while (S.p.c) refers to the respective share divided by the number of children aged 18 or younger still living in the household.

Table 2.7: First-Stage Specifications - Albania (2002 - 2012)

Dependent Variable: Fertility (Number of children ever born)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Panel A: Mothers aged 20 to 49 years old with 2 or more children younger than 18 years old</i>							
Boy first	—	—	—	-0.2461*** (0.0178)	-0.4655*** (0.0267)	-0.0255 (0.0227)	-0.0283 (0.0234)
Boy second	—	—	—	-0.2175*** (0.0177)	-0.4370*** (0.0267)	0.0031 (0.0226)	—
Two girls	—	—	0.4723*** (0.0254)	—	—	0.4412*** (0.0346)	0.4381*** (0.0267)
Two boys	—	-0.1621*** (0.0209)	—	—	0.0029 (0.0226)	—	0.0024 (0.0226)
Same sex	0.2289*** (0.0192)	—	—	0.2202*** (0.0173)	—	—	—
R-squared	0.1197	0.1090	0.1569	0.2982	0.2665	0.2982	0.2982
Observations	7480	7480	7480	7404	7404	7404	7404
Mean dep. var.	2.60	2.60	2.60	2.60	2.60	2.60	2.60
F-stat excl. inst.	140.1	61.2	343.3	161.7	0.02	268.7	134.4
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	Yes	Yes	Yes	Yes
<i>Panel B: Husband's of the married women with 2 or more children younger than 18 years old</i>							
Boy first	—	—	—	-0.2393*** (0.0204)	-0.4619*** (0.0309)	-0.0115 (0.0262)	-0.0153 (0.0277)
Boy second	—	—	—	-0.2240*** (0.0205)	-0.4465*** (0.0304)	0.0057 (0.0266)	—
Two girls	—	—	0.4703*** (0.0280)	—	—	0.4591*** (0.0400)	0.4465*** (0.0304)
Two boys	—	-0.1615*** (0.0235)	—	—	-0.0003 (0.0268)	—	-0.0014 (0.0269)
Same sex	0.2252*** (0.0216)	—	—	0.2225*** (0.0201)	—	—	—
R-squared	0.1160	0.1059	0.1530	0.2743	0.2415	0.2743	0.2743
Observations	5737	5737	5737	5460	5460	5460	5460
Mean dep. var.	2.60	2.60	2.60	2.60	2.60	2.60	2.60
F-stat excl. inst.	113.9	43.12	275.3	131.4	0.07	220.2	110.2
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the PSU level are in parentheses. Other covariates in the models are the following: Age, Age Squared, Age at First birth, Years of Education and also indicators for Boy 1st and Boy 2nd. Each model is estimated with survey and district fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.

Table 2.8: OLS and IV Estimates of Mother's Labor-Supply Models in Albania (2002 - 2012)

Method			OLS	IV	IV
Instrument for Fertility: (Number of children ever born)			—	Same-sex	Two girls
Survey Fixed-Effects			Yes	Yes	Yes
District Fixed-Effects			Yes	Yes	Yes
Controls			Yes	Yes	Yes
Dependent variable	Mean	Obs.	(1)	(2)	(3)
<i>Panel A: All databases (LSMS and DHS)</i>					
Worked in last 7 days	0.445	7404	-0.0063 (0.0073)	0.0474 (0.0474)	0.0360 (0.0340)
Worked off-farm	0.164	7404	-0.0291*** (0.0045)	0.0052 (0.0354)	0.0503** (0.0258)
Worked on-farm	0.218	7404	0.0347*** (0.0066)	0.0125 (0.0396)	-0.0019 (0.0280)
Self-employed	0.067	7404	-0.0086** (0.0037)	0.0404 (0.0270)	-0.0107 (0.0182)
<i>Panel B: Only LSMS databases</i>					
Hours per week (total)	16.69	4862	-0.203 (0.3867)	4.697* (2.5509)	4.409** (1.8866)
Hours per week (off-farm)	7.21	4862	-1.429*** (0.255)	-0.078 (1.902)	2.683* (1.449)
Hours per week (on-farm)	7.56	4862	1.924*** (0.321)	2.303 (1.779)	1.570 (1.332)
Hours per week (self)	2.36	4862	-0.470** (0.181)	2.841** (1.444)	0.225 (1.024)
Second occupation	0.014	4862	0.004 (0.002)	0.010 (0.014)	0.002 (0.011)
Log(Monthly Labor Income)	50498.2	4862	-0.3972*** (0.0778)	0.4172 (0.5740)	0.8679 (0.4384)

Notes: Standard errors clustered at the PSU level are in parentheses. All the models include the following covariates: Age, Age Squared, Age at First birth and Years of Education. IV Models in columns (2) and (3) also include indicators for Boy 1st and Boy 2nd. The sample includes mothers aged 20 to 49 years old with 2 or more children younger than 18 years old. Each model is estimated with survey and district fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.

Table 2.9: OLS and IV Estimates of Father's Labor-Supply Models in Albania (2002 - 2012)

Method			OLS	IV	IV
Instrument for Fertility: (Number of children ever born)			—	Same-sex	Two girls
Survey Fixed-Effects			Yes	Yes	Yes
District Fixed-Effects			Yes	Yes	Yes
Controls			Yes	Yes	Yes
Dependent variable	Mean	Obs.	(1)	(2)	(3)
<i>Panel A: All databases (LSMS and DHS)</i>					
Worked in last 7 days	0.790	5460	0.0124* (0.0067)	0.0581 (0.0445)	0.0257 (0.0314)
Worked off-farm	0.394	5460	-0.0234*** (0.0080)	0.0365 (0.0537)	0.0671* (0.0401)
Worked on-farm	0.218	5460	0.0347*** (0.0066)	0.0125 (0.0396)	-0.0019 (0.0280)
Self-employed	0.198	5460	-0.0137** (0.0063)	0.0215 (0.0437)	-0.0072 (0.0319)
<i>Panel B: Only LSMS databases</i>					
Hours per week (total)	34.39	4530	0.1056 (0.4501)	-0.2895 (2.8443)	-1.1751 (2.0713)
Hours per week (off-farm)	18.60	4530	-1.609*** (0.426)	-4.3926 (2.8683)	-0.8306 (2.1766)
Hours per week (on-farm)	9.15	4530	2.744*** (0.390)	2.3598 (2.2198)	0.5399 (1.6274)
Hours per week (self)	7.65	4530	-0.741** (0.331)	2.7154 (2.3503)	-0.0305 (1.7875)
Second occupation	0.058	4530	0.008* (0.004)	0.0321 (0.0289)	0.0406* (0.0230)
Log(Monthly Labor Income)	193092.2	4530	-0.3816*** (0.1144)	-0.6700 (0.7176)	-0.0883 (0.5305)

Notes: Standard errors clustered at the PSU level are in parentheses. All the models include the following covariates: Age, Age Squared, Age at First birth and Years of Education. IV Models in columns (2) and (3) also include indicators for Boy 1st and Boy 2nd. The sample of fathers includes the husbands of the married women aged 20 to 49 years old with 2 or more children younger than 18 years old) Each model is estimated with survey and district fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.

Table 2.10: OLS and IV Estimates of Parental Labor-Supply - Heterogeneity Analysis by Educational Attainment

Method Instrument for Fertility:	Primary Educ. Level (≤ 8 years of schooling)					Above Primary Educ. Level (>8 years of schooling)				
			OLS —	IV Same-sex	IV Two girls			OLS —	IV Same-sex	IV Two girls
Dependent variable	Mean	Obs.	(1)	(2)	(3)	Mean	Obs.	(4)	(5)	(6)
<i>Panel A: Mothers</i>										
<i>Worked off-farm</i>	0.054	3900	-0.019*** (0.004)	0.014 (0.028)	0.033 (0.021)	0.286	3504	-0.059*** (0.011)	-0.023 (0.080)	0.073 (0.060)
<i>F-stat FS</i>				90.22	161.8				77.94	119.3
<i>Hours per week</i>	14.56	2578	0.140 (0.461)	5.523* (3.012)	4.092* (2.185)	19.09	2284	-1.755** (0.728)	3.398 (4.660)	4.514 (3.632)
<i>F-stat FS</i>				58.72	95.82				62.51	85.07
<i>Panel B: Fathers</i>										
<i>Worked off-farm</i>	0.283	2420	-0.024** (0.010)	0.030 (0.077)	0.060 (0.049)	0.482	3040	-0.017 (0.013)	0.034 (0.077)	0.068 (0.063)
<i>F-stat FS</i>				48.10	104.2				90.03	117.6
<i>Second occupation</i>	0.054	2048	-0.000 (0.006)	0.055 (0.040)	0.052* (0.028)	0.062	2482	0.018** (0.007)	0.008 (0.044)	0.026 (0.037)
<i>F-stat FS</i>				44.69	85.45				72.56	87.25
Survey Fixed-Effects			Yes	Yes	Yes			Yes	Yes	Yes
District Fixed-Effects			Yes	Yes	Yes			Yes	Yes	Yes
Controls			Yes	Yes	Yes			Yes	Yes	Yes

Notes: Standard errors clustered at the PSU level are in parentheses. All the models include the following covariates: *Age*, *Age Squared*, *Age at First birth* and *Years of Education*. IV Models in columns (2) and (3) also include indicators for *Boy 1st* and *Boy 2nd*. The sample of mothers includes women aged 20 to 49 years old with 2 or more children younger than 18 years old. The sample of fathers includes the husbands of the married women with 2 or more children younger than 18 years old. Each model is estimated with survey and district fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%. The F-Statistic of the excluded instruments corresponds to the first-stage estimation.

Table 2.11: OLS and IV Estimates of Parental Labor-Supply - Heterogeneity Analysis by Birth Cohort

Method Instrument for Fertility:	Younger Parent					Older Parent				
			OLS —	IV <i>Same-sex</i>	IV <i>Two girls</i>			OLS —	IV <i>Same-sex</i>	IV <i>Two girls</i>
Dependent variable	Mean	Obs.	(1)	(2)	(3)	Mean	Obs.	(4)	(5)	(6)
<i>Panel A.1. Mothers born after 1972</i>						<i>Panel A.2. Mothers born before 1972 (incl.)</i>				
Worked off-farm	0.111	3253	-0.024*** (0.007)	0.032 (0.056)	0.058 (0.043)	0.205	4151	-0.031*** (0.006)	-0.011 (0.045)	0.048 (0.032)
<i>F-stat FS</i>				71.82	108.15				94.65	173.4
Hours per week	11.95	2019	-1.002 (0.687)	14.292*** (5.066)	10.584*** (3.723)	20.05	2843	-0.007 (0.484)	-0.010 (3.043)	1.599 (2.180)
<i>F-stat FS</i>				50.98	69.69				72.02	124.18
<i>Panel B.1. Fathers born after 1966</i>						<i>Panel B.2. Fathers born before 1966 (incl.)</i>				
Worked off-farm	0.349	2719	-0.009 (0.013)	0.083 (0.096)	0.070 (0.069)	0.439	2741	-0.036*** (0.010)	0.008 (0.064)	0.063 (0.050)
<i>F-stat FS</i>				61.55	96.03				65.95	123.3
Second occupation	0.056	2194	0.009 (0.007)	0.095 (0.058)	0.079* (0.042)	0.061	2336	0.008 (0.005)	-0.005 (0.032)	0.015 (0.027)
<i>F-stat FS</i>				47.54	76.03				67.18	105.75
Survey Fixed-Effects			Yes	Yes	Yes			Yes	Yes	Yes
District Fixed-Effects			Yes	Yes	Yes			Yes	Yes	Yes
Controls			Yes	Yes	Yes			Yes	Yes	Yes

Notes: Standard errors clustered at the PSU level are in parentheses. All the models include the following covariates: *Age*, *Age Squared*, *Age at First birth* and *Years of Education*. IV Models in columns (2) and (3) also include indicators for *Boy 1st* and *Boy 2nd*. The sample of mothers includes women aged 20 to 49 years old with 2 or more children younger than 18 years old. The sample of fathers includes the husbands of the married women with 2 or more children younger than 18 years old. Each model is estimated with survey and district fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%. The F-Statistic of the excluded instruments corresponds to the first-stage estimation.

Table 2.12: OLS and IV Estimates of Parental Labor-Supply - Heterogeneity Analysis by Family Structure

Method Instrument for Fertility:	Extended Family					Non-Extended Family				
			OLS —	IV <i>Same-sex</i>	IV <i>Two girls</i>			OLS —	IV <i>Same-sex</i>	IV <i>Two girls</i>
Dependent variable	Mean	Obs.	(1)	(2)	(3)	Mean	Obs.	(4)	(5)	(6)
<i>Panel A: Mothers</i>										
Worked off-farm	0.162	2804	-0.037*** (0.008)	0.020 (0.060)	0.077* (0.046)	0.165	4600	-0.025*** (0.006)	-0.001 (0.044)	0.038 (0.032)
<i>F-stat FS</i>				71.02	104.76				92.89	167.24
Hours per week	17.37	1729	-0.663 (0.715)	7.532* (4.303)	6.982* (3.564)	16.31	3133	-0.043 (0.462)	3.620 (3.294)	3.536 (3.278)
<i>F-stat FS</i>				49.54	57.64				66.87	67.53
<i>Panel B: Fathers</i>										
Worked off-farm	0.359	1967	-0.013 (0.015)	0.107 (0.091)	0.061 (0.071)	0.414	3493	-0.027*** (0.009)	-0.005 (0.071)	0.072 (0.049)
<i>F-stat FS</i>				61.30	77.33				71.78	133.3
Second occupation	0.060	1607	0.009 (0.007)	0.044 (0.048)	-0.012 (0.040)	0.057	2923	0.008 (0.005)	0.027 (0.038)	0.071** (0.028)
<i>F-stat FS</i>				50.66	57.09				63.27	109.8
Survey Fixed-Effects			Yes	Yes	Yes			Yes	Yes	Yes
District Fixed-Effects			Yes	Yes	Yes			Yes	Yes	Yes
Controls			Yes	Yes	Yes			Yes	Yes	Yes

Notes: Standard errors clustered at the PSU level are in parentheses. All the models include the following covariates: *Age*, *Age Squared*, *Age at First birth* and *Years of Education*. IV Models in columns (2) and (3) also include indicators for *Boy 1st* and *Boy 2nd*. The sample of mothers includes women aged 20 to 49 years old with 2 or more children younger than 18 years old. The sample of fathers includes the husbands of the married women with 2 or more children younger than 18 years old. Each model is estimated with survey and district fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%. The F-Statistic of the excluded instruments corresponds to the first-stage estimation.

Table 2.13: *Falsification Exercise: Determinants of Living in an Extended Family - Albania (2002 - 2012)*

Dependent Variable: Living in an Extended Family				
	(1)	(2)	(3)	(4)
Boy first	0.0065 (0.0131)	0.0247 (0.0165)	0.0159 (0.0110)	0.0064 (0.0159)
Boy second	-0.0002 (0.0167)	0.0180 (0.0156)	0.0092 (0.0114)	
Two girls	-0.0187 (0.0228)			-0.0186 (0.0156)
Two boys		-0.0176 (0.0227)		0.0004 (0.0166)
Same-sex			-0.0091 (0.0114)	
R-squared	0.2982	0.2665	0.2982	0.2982
Observations	7404	7404	7404	7404
Mean dep. var	0.381	0.381	0.381	0.381
Survey FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

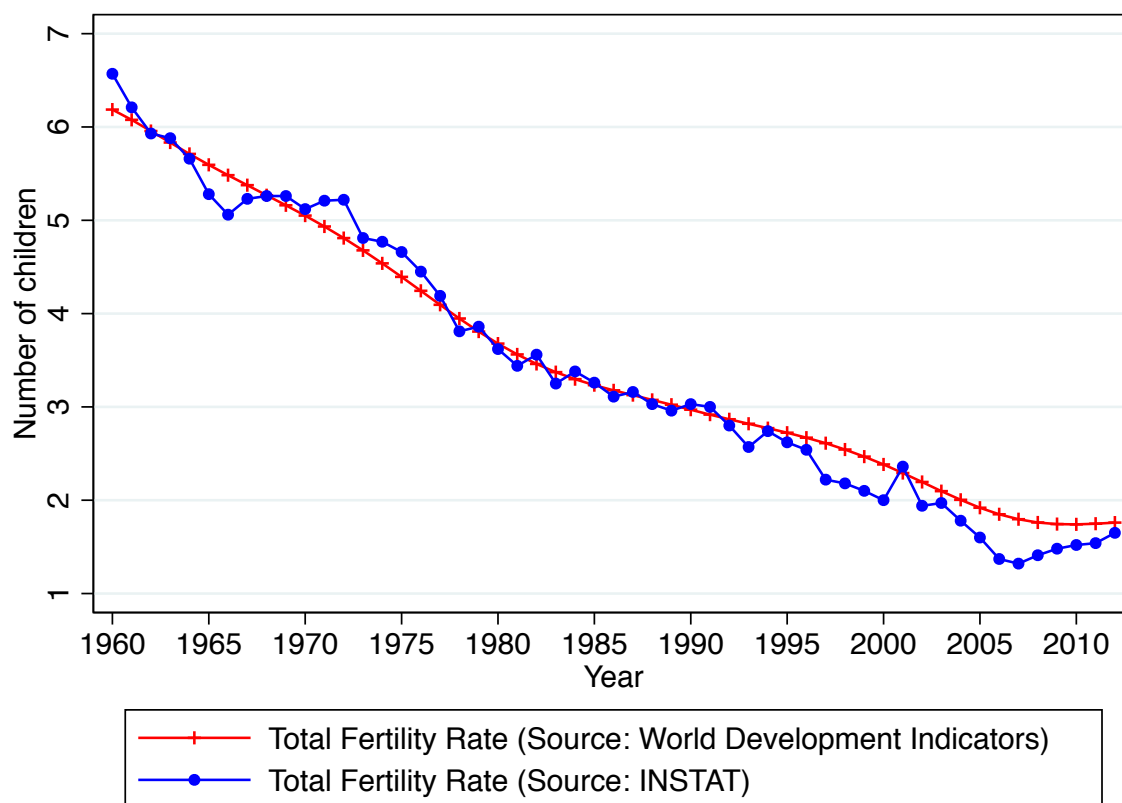
Notes: Standard errors clustered at the PSU level are in parentheses. Other covariates in the models are the following: *Age, Age Squared, Age at First birth, Years of Education* and also indicators for *Boy 1st* and *Boy 2nd*. Each model is estimated with survey and district fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.

Table 2.14: IV Estimates of the Effect of Fertility on Family Spending per-capita - Heterogeneity Analysis by Expenditure Quintiles and Family Structure - Albania (2002 - 2012)

Variables	All	Expenditure Quintiles				
		Q1 (Poorest)	Q2 (Poorer)	Q3 (Middle)	Q4 (Richer)	Q5 (Richest)
Mother's age	35.04	34.38	34.82	35.03	35.35	35.91
Mother's age at 1st birth	23.22	23.06	23.02	23.11	23.33	23.70
Father's age	39.91	38.77	39.85	39.98	40.36	40.95
Father's age at 1st birth	28.33	27.86	28.14	28.23	28.60	29.01
% Rural	0.482	0.699	0.656	0.492	0.309	0.209
% Third child	0.418	0.488	0.445	0.430	0.384	0.310
% Extended Family	0.381	0.313	0.384	0.386	0.394	0.451
Dependent variable: Ln (Per-capita Expenditure)						
Panel A: Extended Family						
Fertility						
IV (<i>same sex</i>)	-0.0890 (0.0826)	0.1579 (0.1811)	-0.0252 (0.1036)	-0.1985*** (0.0632)	-0.0309 (0.1505)	-0.2110 (0.1321)
IV (<i>two girls</i>)	-0.1563** (0.0674)	-0.0948 (0.1184)	-0.0979 (0.0683)	-0.2180*** (0.0585)	-0.1487 (0.0950)	-0.2967** (0.1162)
Observations	1699	248	340	401	450	260
Panel B: Non-Extended Family						
Fertility						
IV (<i>same sex</i>)	-0.1600** (0.0649)	-0.2860*** (0.0795)	-0.2088*** (0.0634)	-0.2170*** (0.0367)	-0.1769*** (0.0683)	-0.1614 (0.1187)
IV (<i>two girls</i>)	-0.2054*** (0.0465)	-0.2444*** (0.0578)	-0.2314*** (0.0429)	-0.1898*** (0.0270)	-0.2254*** (0.0384)	-0.1884* (0.0985)
Observations	3081	784	684	666	646	301
Survey Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

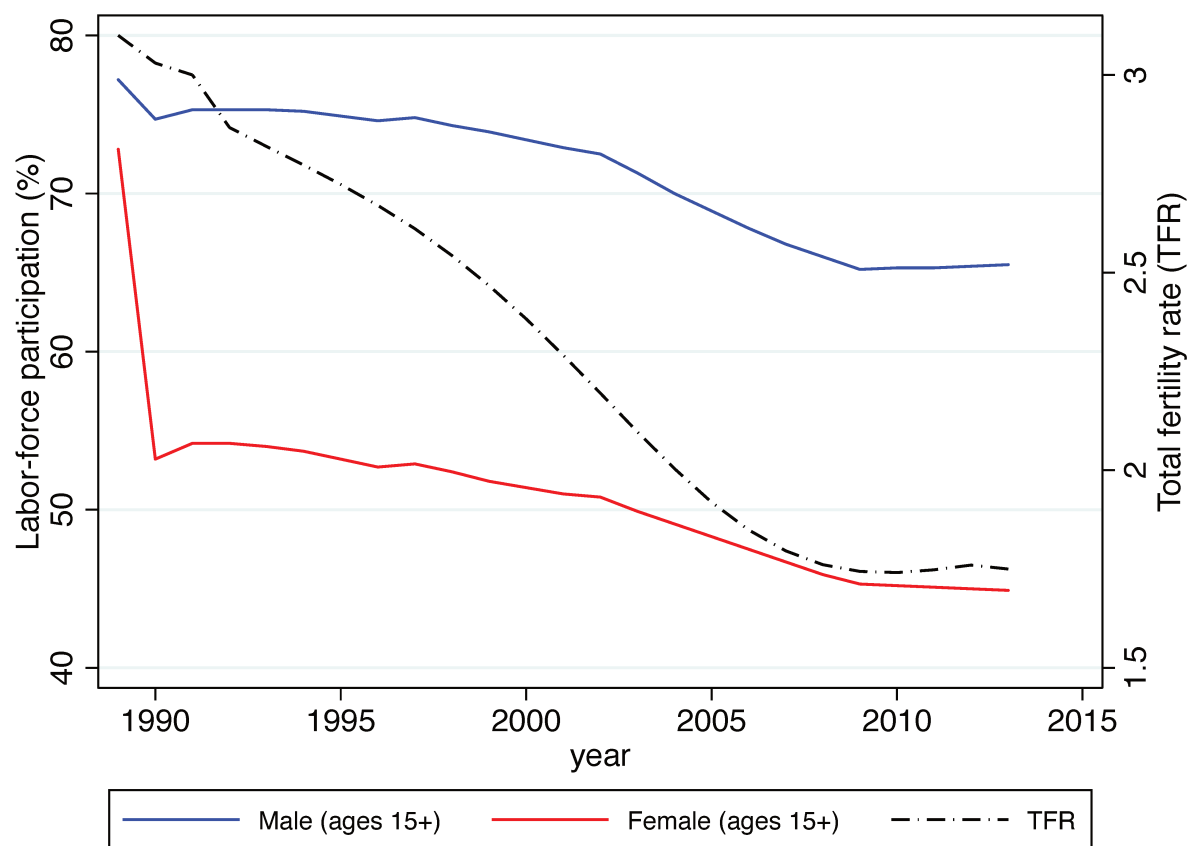
Notes: Standard errors clustered at the PSU level are in parentheses. The IV models include the following covariates: Age, Age Squared, Age at First birth, Years of Education, Boy1st and Boy 2nd. Each model is estimated with survey and district fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.

Figure 2.1: *Evolution of the Total Fertility Rate in Albania (1960 - 2012)*



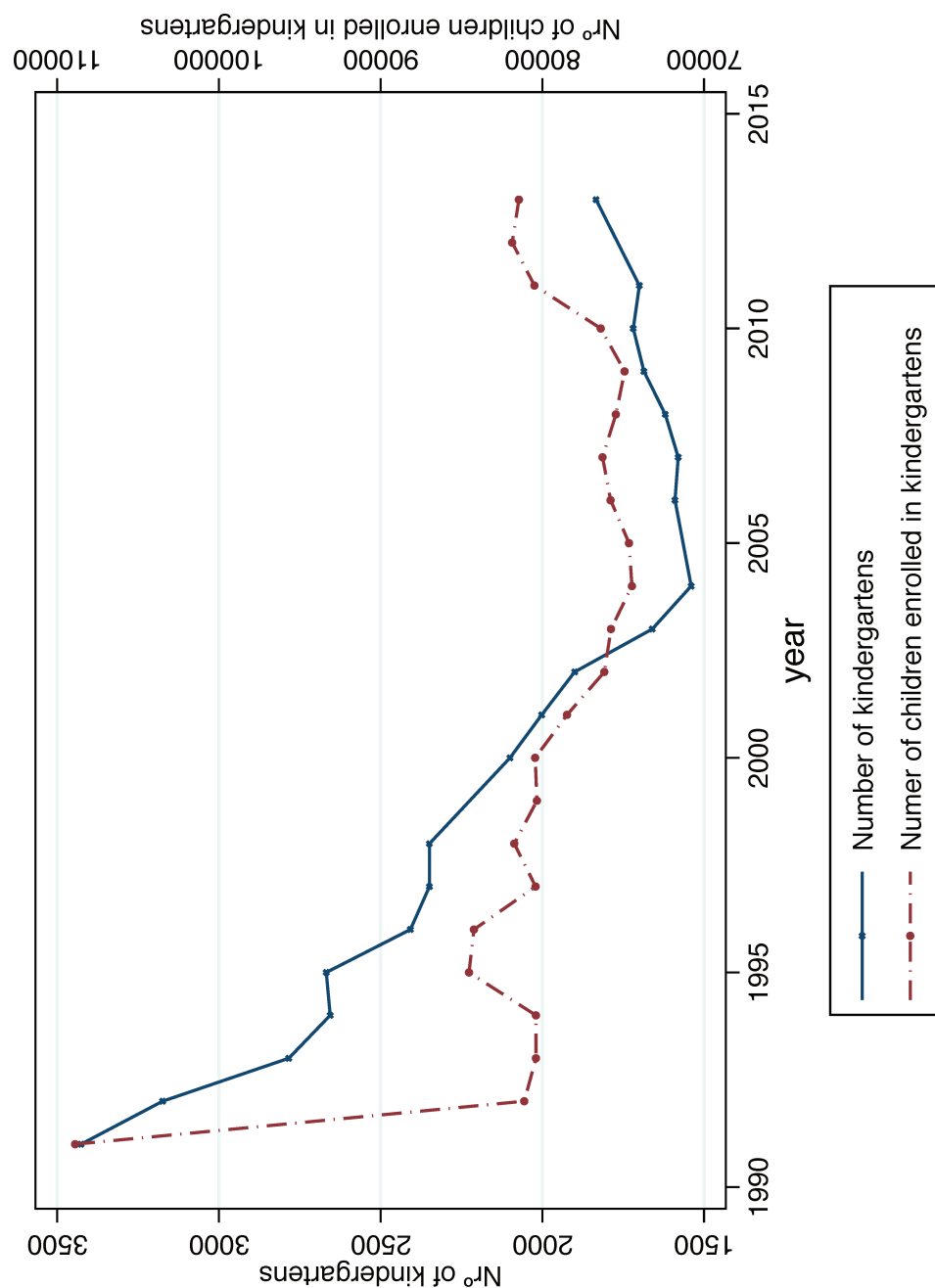
Notes: According to Gjonça et al (2008), the total fertility rate rose during the 1950s reaching a peak of almost 7 children per woman by 1960. This was followed in the 1970s by a steady decline, with a total fertility rate of less than 4 in 1980 and just over 3 children per woman in 1990. The 1990s saw a continuing reduction which reached the replacement level of 2.2 children per woman in 2002, which at the same time was the highest fertility rate in Europe. During the 2000s the total fertility rate has decreased even further reaching a level of 1.65 children per women in 2012, which currently characterizes Albania as a country with a very low fertility rate.

Figure 2.2: *Labor-Force Pariticipation and Total Fertility Rate - Albania (1990 - 2012)*



Notes: Source: Albanian Institute of Statistics (INSTAT)

Figure 2.3: Pre-school in Albania (1990 - 2012)



Source: Albanian Institute of Statistics (INSTAT) and Ministry of Education and Sports (MAS)
 Notes: Data for public and private kindergartens in total, since 2006-07 year school.

Figure 2.4: Graphical Version of Fraction of Households that had Another Child by Parity and Sex Composition - Decomposed by Survey

(Mothers 20 to 49 years old with at least 2 children younger than 18 years old)

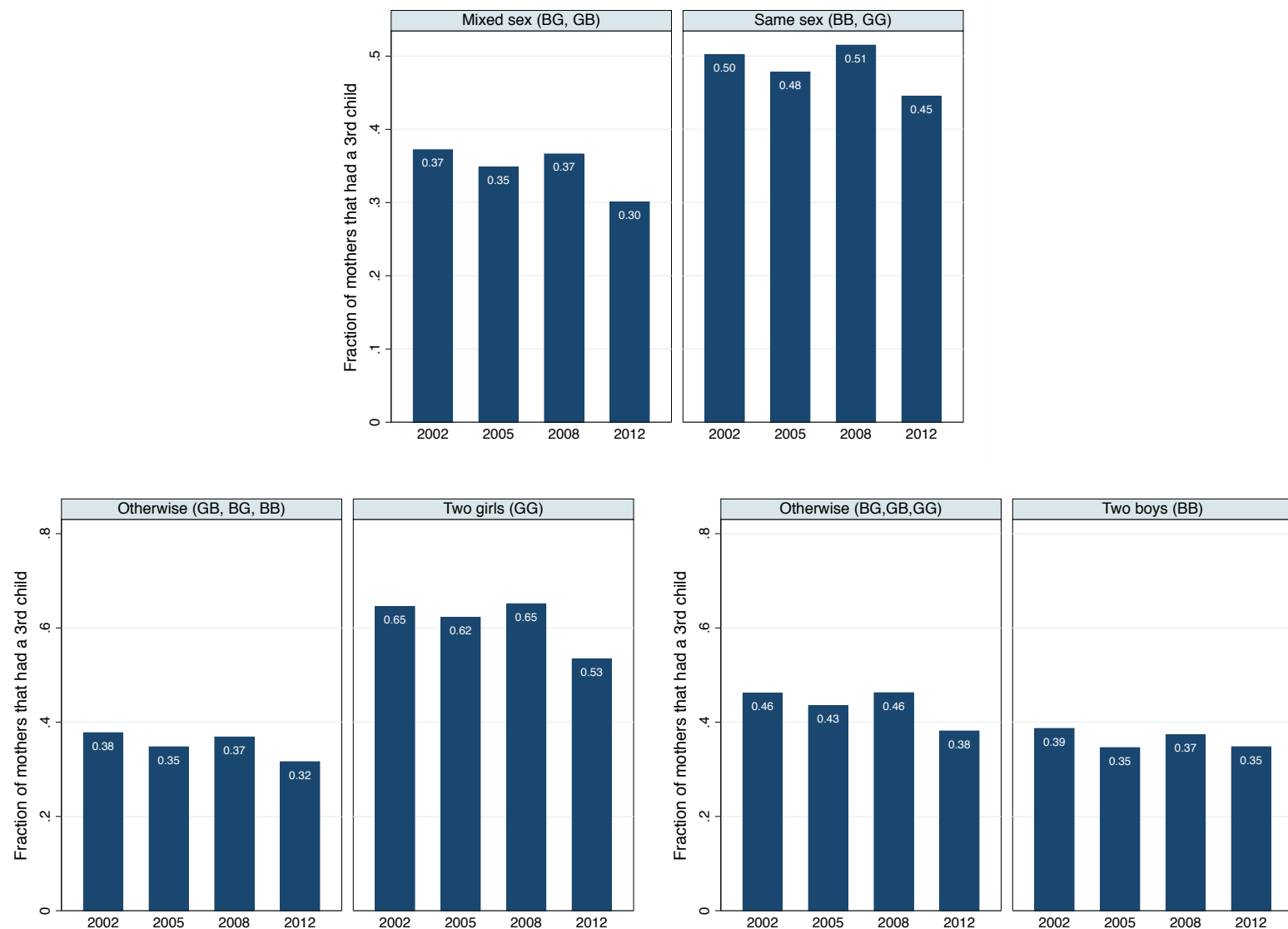
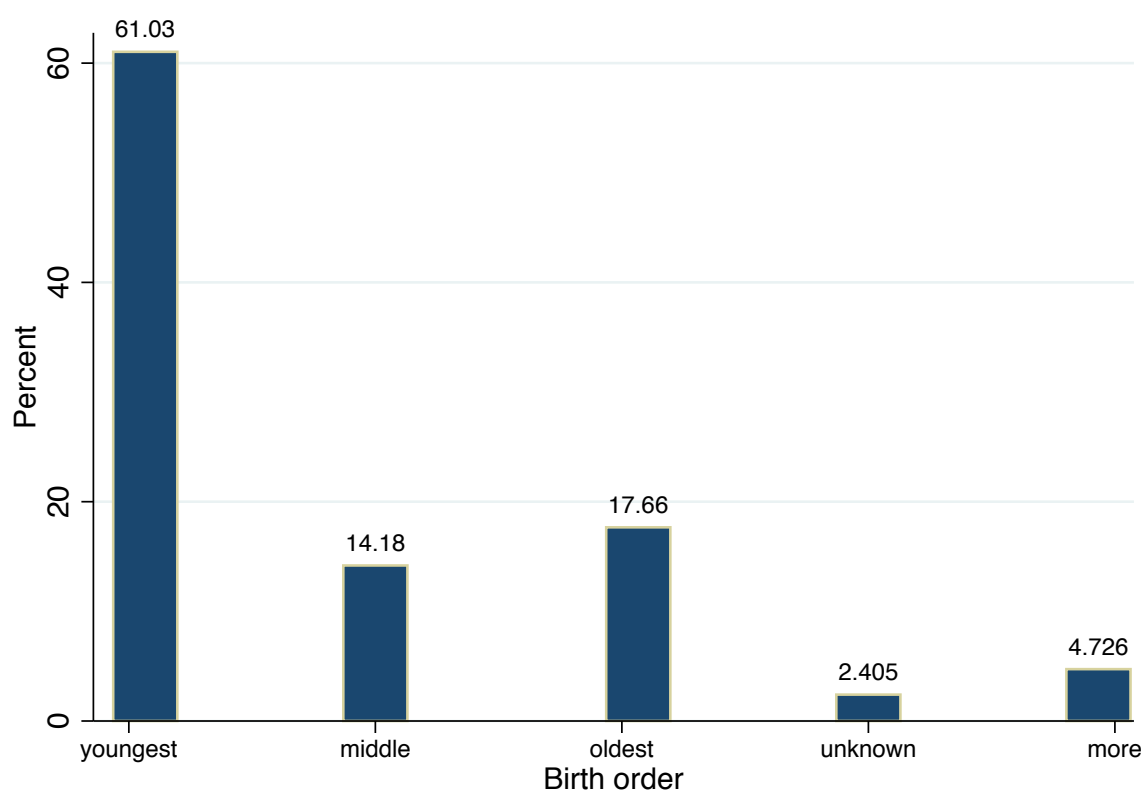


Figure 2.5: *Proportion of parents in extended families living with their married son by their son's birth order*



Source: 2002 Albania LSMS

Source: 2002, 2005 and 2012 Albanian LSMS; 2008-09 Albanian DHS

Appendix

2.A Data sources and data construction

The 2002, 2005 and 2012 LSMS were undertaken by the Albanian National Institute of Statistics (INSTAT) with the technical assistance of the World Bank; the 2008-09 DHS was also conducted by the INSTAT but also by the Institute of Public Health (IPH). All these databases are nationally representative surveys for the whole country, as well as at regional and at the urban/rural level. Each survey contains a wide range of information on individual's demographic and socioeconomic characteristics (i.e. fertility, health, employment, migration and education among many others). The LSMS surveys also comprise a detailed expenditure section which is used in this study to obtain figures on clothing and schooling expenditures on children. The DHS is especially characterized for having broader information about health. In each one of the databases, the country was stratified into four regions (strata), Tirana, Coastal, Central and Mountain. All the surveyed individuals belong to the 36 districts in the country, with all regions nearly equally represented. They are also based on a two-stage sampling cluster design.

The 2002 LSMS is the first survey of this type conducted in Albania. This survey includes 3,600 households, from which 8,395 are men and 8,126 are women. Three years later, the same survey was redone but unfortunately the same households were not followed in time in order to form a panel.³⁶ The 2005 LSMS covers 3,638 households, from which 8,713 are men and 8,589 are women. The 2008-09 DHS comprises a total of 7,999 households, from which 7,584 are women and 3,013 are men. And lastly, the 2012 LSMS contains in total 6,671 household observations, from where 12,747 correspond to men and 12,588 correspond to women. The respective original individual observations for each survey are reported in Table 2.D.1 below.³⁷

These surveys interviewed both women and men with respect to their fertility decisions, thereby permitting the unique opportunity to incorporate information in my study regarding the fertility history of each household. In the 2008-09 DHS, all the children ever born are already matched to each one of the mothers and fathers they correspond to, which makes the

³⁶The data from the 2002 LSMS has been used to construct a panel survey with two additional waves, one in 2003 and another one in 2004. The problem with this panel survey is that the questionnaire was readministered only to a sub-sample of the 2002 LSMS households, which reduces a lot the sample size in order to allow for the adopted empirical strategy in this study.

³⁷For further information on the Albanian LSMS surveys, see the documentation of the World Bank available online: <http://microdata.worldbank.org/index.php/catalog/lsms>.

DHS a perfect database for studying the fertility history of the parents. In addition, their gender, age, date of birth, month of birth, year of birth and birth order among other characteristics are very well indicated. In contrast, in the LSMS surveys, the fertility history information is not as clearly presented as in the DHS, which means that the children are not already linked to their corresponding parents. However, the advantage of the LSMS surveys is that they provide a direct mother and father identifier code, which is used to link the children to their corresponding parents.

In addition, the LSMSs databases contain very detailed information on the family structure of each one of the households, including information about the children no longer living in the household, which makes possible the construction of the entire fertility history of each parent. This means that, in the LSMS surveys, it is possible to identify which one of the individuals is the head of the household, which is the spouse, which are the children and also which are the other members of the family such as grandparents, sisters, brothers, uncles, aunts etc. Apart from this, it is possible to identify the age and gender of each one of the household individuals, which is very useful to construct the number of children per parent and also the birth order of each one of them. Therefore, for the LSMSs databases, I identified one by one all the children corresponding to each mother and father, their gender, age, year of birth and birth order with respect to their siblings. Then by reshaping each database, I could match all the children to their corresponding parents within households by using the mother and father identifier code and at the same time I could order the children from youngest to eldest by birth order. As a cross-check, I also attached individuals in a household labeled as "child" in the primary relationship code to a female householder or the spouse of a male householder. In households with multiple families, detailed relationship codes as well as subfamily identifiers were used to pair children with mothers.

On the other hand, the Household Roaster Module of the LSMSs databases contains detailed information on each one of the members who were living in the household at the time of the survey but it does not contain information on other members such as sons or daughters who had already left the household. Therefore, in order to construct the entire fertility history of each woman I had to use information about the sons and daughters living away, which is available in the Migration Module of the LSMSs databases.

However, a small concern about the Migration Module is that this information might be

subject to measurement errors. More precisely, the major concern comes from the plausible underreporting of daughters. The first reason behind this concern might be to the fact that Albanian parents expect at least one of their sons to take care of them at old age, while the rest of the children especially girls leave the household. Albanian tradition mandates that it is the role of the youngest son to take care of the parents in their later years. Therefore, it could be that there are many households that report only sons because it has been impossible to track the rest of the siblings. The second reason could be to pre-natal sex selection. In some developing countries including Albania, the presence of strong son preference could affect the sex composition of children, either through stopping rules or selective abortion. However, by restricting the sample of women to mothers with at least two children whose oldest child is at most 18 years old at the time of the survey, I rule out these two concerns and I find that the fraction of households that had a boy or a girl as a first child is quite balanced for each of the four databases (see Table 2.D.2). Hence, due to these reasons and also because it is more relevant to study the effect of fertility on labor supply for the parent's whose children still live in the household, I restrict my sample to mothers aged 20 to 49 with at least two children, whose oldest child was at most 18 years old at the time of the survey and for the case of the fathers I restrict my sample to the husbands of the married women.

2.B LATE without monotonicity

Estimates in this paper might not capture the LATE of compliers because of the presence of defiers. Defiers could be present in this study because some parents might be sex-biased. This means that some parents might just want two daughters and no more children afterwards or they might have preferences for having at least two children of the same sex and choose to have a third child if the first two are of mixed sex; such parents would be defiers.

De Chaisemartin (2017) shows that the 2SLS estimator still estimates a LATE under a weaker condition than monotonicity. In other words, the 2SLS estimator is still valid even if there are defiers, provided the "compliers-defiers" condition is satisfied. If a subgroup of compliers accounts for the same percentage of the population as defiers and has the same LATE, 2SLS estimates the LATE of the remaining part of compliers. Under this condition, the part of compliers and the defiers cancel one another out, and the 2SLS coefficient is equal to the effect of the treatment on a subpopulation of compliers which the author calls "surviving-compliers". Essentially, this conditions requires that compliers and defiers' LATEs are not too different.

In this paper, the share of defiers (those parents whose two first children are *not* girls and decide to have a third child) cannot be more than 35.3%. (See Panel B of Table 2.4) Following De Chaisemartin (2017), let $P(F)$ denote the percentage of defiers in the population. Let $LATE_C$ denote the LATE of compliers, and let $LATE_F$ denote the LATE of defiers. Finally, let FS denote the coefficient of the instrument in the first-stage regression of the treatment on the instrument, and let W denote the coefficient of the treatment on the 2SLS regression. The author shows that if:

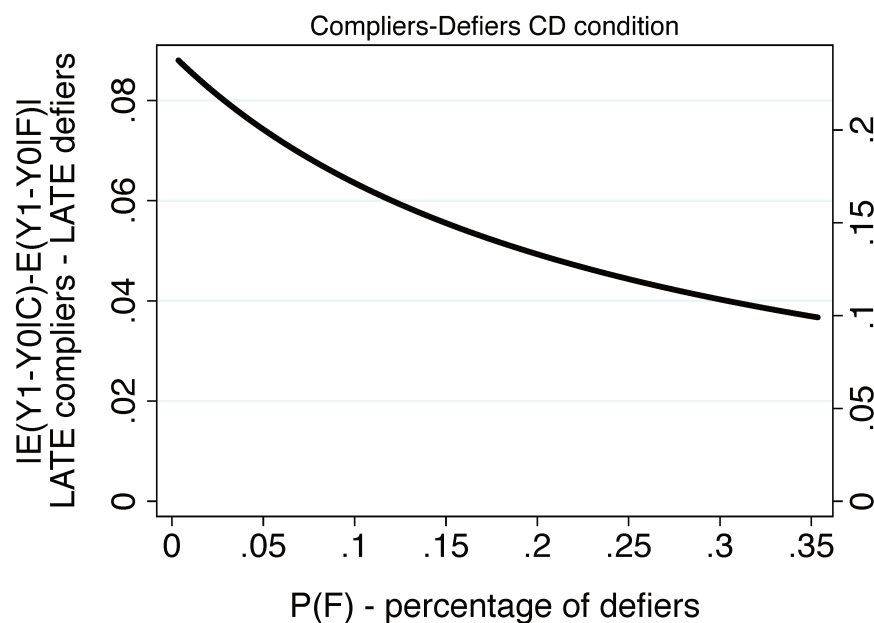
$$LATE_C - LATE_F \leq |W| * \frac{FS}{(FS + P(F))} \quad (2.7)$$

then the "compliers-defiers" CD condition holds and the 2SLS estimator is valid. This condition is more likely to be satisfied when the instrument has large first and second stages, and when defiers are unlikely to account for a large fraction of the population.

Figure 2.B.1 in the paper applies this result to the data in this paper, and plots values of $P(F)$ (x-axis) and $LATE_C - LATE_F$ (y-axis) for which $LATE_C - LATE_F \leq W * \frac{FS}{(FS + P(F))}$. Those are all the values below the black line. For instance if $P(F) \leq 0.05$ (meaning that there are 5% of defiers) and if $LATE_C - LATE_F \leq 0.074$ (meaning that the LATEs of compliers differ by no more than 7.4 percentage points), then the CD condition will hold in this application.

In the 2008 Albanian Demographic and Health Survey, women were asked their ideal sibship composition. Among women whose first two children are boys or have a mixed combination, around 1,7% (12 out of 698) had 3 children or more and also declared that their ideal sex sibship composition would have been at least two boys and no girl, or no boy and two girls. There are also women whose two first children are girls and decide not to have more children because they have a sex bias for girls. These women represent 1.1% of women with two girls that decide not to have more children because their ideal number of children is effectively two girls. This evidence indicates that the share of defiers in this application is around 3%, which is quite small suggesting that the LATE on the compliers does not change significantly when assuming the weaker condition on monotonicity.

Figure 2.B.1: *LATE without monotonicity - Compliers-Defiers (CD) condition*



Notes: For all values of $P(F)$ and $LATE_C - LATE_F = E(Y1 - Y0|C) - E(Y1 - Y0|F)$ below the black line, the “compliers-defiers” condition is satisfied.

2.C Theoretical Framework

The model presented in this section is intended to serve as a framework and to motivate the cross-section estimation of the effect of fertility on parental labor supply. Dynamic considerations are ignored for simplicity reasons. In order to explore the factors affecting the relationship between fertility and their parent's labor supply in a developing country context, I adapt a unitary household framework that incorporates features of the [Becker and Lewis \(1973\)](#) and [Becker and Tomes \(1976\)](#) quantity/quality model into [Blau and Robins \(1998\)](#) and [Connelly \(1992\)](#) models of home production.

This model is intended to apply to households in which children requiring continuous care are present and in which the mother and the father, as well as one other potential childcare provider are present (i.e. grandparent, relative etc). In this context, household member's utilities are represented by a unique utility function. The household members are assumed to make choices regarding consumption of market goods (C), childcare quality (Q), leisure (L) and the number of children in the family (N) in such a way as to maximize their utility U subject to a series of constraints:

$$\begin{aligned}
 \max \quad & U = U(C, L_M, L_F, Q, N) \\
 \text{subject to} \quad & L_M + H_M + t_M = 1 && \text{(Mother's time constraint)} \\
 & L_F + H_F = 1 && \text{(Father's time constraint)} \\
 & Q = Q(t_M, t_0, t_K) && \text{(Childcare quality)} \\
 & \frac{\partial Q}{\partial t_M} = Q_M > 0, \frac{\partial Q}{\partial t_K} = Q_K > 0, Q_{MM} < 0, Q_{KK} < 0 \\
 & C = E + w_F H_F + w_M H_M - (p_K t_K + p_N) N && \text{(HH budget constraint)} \\
 & t_M + t_K + t_0 = 1 && \text{(Child's time constraint)}
 \end{aligned}$$

where the subscripts M and F represent mother and father, respectively.

The total time available to the mother is normalized to 1 and can be divided between working hours (H_M), leisure time (L_M) and time spent looking after the children (t_M). Notice that the father's time is only spent working (H_F) or enjoying leisure time (L_F), in other words, the father does not spend time caring for the children.

The three potential sources of childcare quality are the mother (t_M), the potential informal

provider (i.e. grandparents) (t_0) and the formal childcare available in the market (t_K). On the one hand, it is assumed that a formal market childcare quality is available for purchase at hourly price (p_K) per unit of quality. On the other hand, the childcare provided by the mother and the informal provider (i.e. grandparents) is free.

The consumption possibilities of the family are limited by the amount of exogenous income available (E) (i.e. remittances and alike), the amount of income the mother and father earn (their respective wages (w_F) and (w_M) times their respective working hours (H_F) and (H_M)) and child costs. Child costs are separated into childcare costs and direct child costs with former ones being modelled as the time devoted to childcare as in [Blau and Robins \(1998\)](#) [Connelly \(1992\)](#) and [Kimmel \(1998\)](#). (p_K) is the hourly price of formal childcare purchased in the market. (p_N) is the price of market inputs required by children and ($P_N N$) represents direct child costs such as food or non food expenditures (i.e. clothing, school uniforms).

The child's time constraint indicates that the time the child is looked after is equal to the time the mother is looking after her plus the number of hours she is looked after somebody else, either in formal childcare or in informal childcare. This constraint rules out the possibility that the family leaves their children on their own.

2.C.1 Maximization Problem

In order to understand how the family chooses between leisure, consumption, childcare purchase and the number of children, I solve the utility maximization problem. The exogenous determinants are the price of childcare; the price of market inputs required by children, the wage rates of the parents and the family's non-labor income.

Case 1: Nuclear Family

First, I consider the case where there is no informal childcare provider. The family has to choose between maternal childcare and formal childcare provision in the market. The household's maximization problem is the following:

$$\max_{H_M, H_F, t_K, N} U = U(\underbrace{E + w_M H_M + w_F H_F - p_K t_K N - p_N N}_C, \underbrace{t_K - H_M}_{L_M}, \underbrace{1 - H_F}_{L_F}, Q(1 - t_K, t_K), N) \quad (2.8)$$

Solving the maximization problem in this case gives one of the following first order conditions:

$$\left(\frac{\partial U}{\partial Q} \frac{\partial Q}{\partial t_M}\right) - \frac{\partial U}{\partial L_M} = \left(\frac{\partial U}{\partial Q} \frac{\partial Q}{\partial t_K}\right) - p_K N \frac{\partial U}{\partial C} \quad (2.9)$$

This F.O.C suggests that those households buying childcare in the market will increase the number of hours of this care until the net marginal benefit of an extra hour of formal childcare equals the net marginal benefit of an extra hour of maternal childcare.

The RHS of this equation means that an extra hour of formal childcare increases the household's utility by increasing the quality of the child (or childcare). This increase in household's utility constitutes the marginal benefit of formal care. The marginal cost of paid care is the decrease in utility induced by the reduction in consumption as a consequence of paying p_K for an extra unit of care. The LHS of this equation means that an extra hour of maternal childcare increases the household's utility by increasing the quality of the child (or childcare). This is the marginal contribution to childcare quality of an extra hour of mother's care. The marginal cost is the decrease in household's utility as a consequence of the mother's decrease in pure leisure. Notice that the household does not face the trade-off between maternal childcare and formal childcare when deciding on the father's optimal amount of working hours.

Dividing (1) by $\partial U / \partial C$, I get an expression for the reservation wage of the mother:

$$\frac{U_{L_M}}{U_C} = w_M = \frac{U_Q}{U_C} (Q_M - Q_K) + p_K N \quad (2.10)$$

This interior solution shows that marginal rate of substitution between goods and leisure of the mother equals the wage and the wage in turn equals the net marginal benefit of maternal care, which depends on the difference in maternal and formal childcare quality, on the price of formal childcare which is the money savings from an hour of maternal childcare, as well as on the number of children.

When the wage is lower than net marginal benefit of maternal care ($w_M < \frac{U_Q}{U_C} (Q_M - Q_K) + p_K N$), the mother will not participate in the labor market. While, mothers for whom the wage is higher than the net marginal benefit of maternal care, do participate in the labor market. The higher is the mother's wage, the higher is the likelihood that the household uses formal childcare, which increases the likelihood of the mother to participate in the labor market.

If the quality of maternal child-care is considered higher than that of the formal childcare

(i.e. $Q_M - Q_K > 0$), the difference between the wage and childcare costs has to compensate for this fact in order to provide a strong enough incentive for mothers to work. In order to increase the incentive for mothers to work, either improvements in formal childcare quality Q_K or reduction in the price of formal childcare p_K or reduction in the number of children N (all other things being equal) would be appropriate. The interior solution of this simple model of utility maximization of the household suggests that the formal childcare price, the formal childcare quality and the number of children are determinants of the employment decision of mothers with dependent children.

Another first order condition from this maximization problem is the following:

$$p_K t_K + p_N = \frac{(\partial U / \partial N)}{(\partial U / \partial C)} = \frac{U_N}{U_C} \quad (2.11)$$

The marginal benefit of having an additional child is equal to the marginal cost, which is given by the disutility of a decrease in the household's consumption equal to total child costs (formal childcare costs and direct child costs). An increase in the number of children has a direct effect on the total cost of childcare by increasing the cost of formal childcare and also the direct child cost. If there are economies of scale in home produced child quality (maternal childcare), then an increase in N will lower the probability of participating in the labor market for mothers.

Case 2: Extended Family

A peculiarity of the market for non-maternal childcare is that all households do not face the same prices in the market. [Heckman \(1974\)](#) and [Blau and Robins \(1998\)](#) have each emphasized the potential some families have for lower cost (at least in monetary terms) informal care, most often provided by a relative such as a grandparent. The existence of extended families where grandparents live in the same house as the parents and the children is also a very common feature of developing countries.

In this case, childcare is provided by three sources: the mother (t_M), the informal childcare provider such a relative of grandparent (t_0) and the market formal childcare (t_K). In addition, I assume that maternal childcare and grandparent childcare are perfect substitutes for the children, meaning that both types of childcare provide the same level of quality to the children. I also assume that non-parental childcare (e.x. grandparents's childcare) is provided at a

price p_0 , which is lower than the price p_K for the market formal childcare. This non-parental childcare price represents the cost related to having grandparents taking care of the children.

The household maximization problem is the following:

$$\max_{H_M, H_F, t_K, N} U = U(\underbrace{E + w_M H_M + w_F H_F - p_K t_K N - p_N N - p_0 t_0 N}_C, \underbrace{t_K + t_0 - H_M}_{L_M}, \underbrace{1 - H_F}_{L_F}, Q(1 - t_K, t_K, t_0), N) \quad (2.12)$$

I also assume that if informal childcare (t_0) is available at a lower price than the market childcare, the family will always use. Therefore, the number of hours of informal childcare used by the household is treated as a fixed parameter, which only appears in the child's time constraint and in the childcare production function.

In this case, there is a corner solution to the household utility maximization problem. The presence of other potential caretakers in the household such as grandparents or other relatives, lowers the amount parents pay for childcare in the market which in turn increases the labor participation of the mother. Even when an increase in the number of children increases the total cost of childcare, the availability of informal non-maternal childcare in the household allows substituting from maternal childcare to informal childcare, which in turn also allows the mother to participate in the labor market. Therefore, $H_M = t_0$ and $H_F > 0$. To sum up, if the grandparent is available to care for the children, childcare is completely informal and mother's participation in the labor market increases.

2.D Supplementary Tables

Table 2.D.1: *Sample Sizes: LSMS and DHS - Albania (2002-2012)*

Survey	Households	Women	Men	Individuals	Women $20 \leq age \leq 49$	Women with +2 children below 18 y.o. $20 \leq age \leq 49$	Husbands of the married women $20 \leq age \leq 49^*$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
2002 LSMS	3,600	8,126	8,395	16,521	3,418	1,558	1,523
2005 LSMS	3,638	8,589	8,713	17,302	2,897	1,553	1,508
2008-09 DHS	7,999	7,584	3,013	10,597	6,066	2,542	1,011
2012 LSMS	6,671	12,588	12,747	25,335	3,995	1,827	1,695
Total	21,908	36,887	32,868	69,755	16,376	7,480	5,737

Notes: Columns (1) to (5) report the number of observations as provided in the official LSMS and DHS data sets. The last two columns report the number of observations after data cleaning procedures were applied. (*) The age in column (7) corresponds to the women.

Table 2.D.2: *Fraction of Households that had Another Child by Parity and Sex Composition in Albania (2002-2012) - Decomposed by Survey*

	(2002) LSMS		(2005) LSMS		(2008-09) DHS		(2012) LSMS	
	Fraction of the sample	Fraction that had another child	Fraction of the sample	Fraction that had another child	Fraction of the sample	Fraction that had another child	Fraction of the sample	Fraction that had another child
<i>Panel A - Mothers aged 20 to 49 with 1 or more children younger than 18 years old</i>								
106 (1) one girl	0.489	0.812 (0.012)	0.480	0.797 (0.013)	0.503	0.827 (0.009)	0.495	0.792 (0.010)
(2) one boy	0.510	0.781 (0.013)	0.519	0.770 (0.013)	0.496	0.794 (0.010)	0.504	0.708 (0.010)
Difference (2) - (1)	0.020 (0.022)	0.030 (0.018)**	0.039 (0.022)	0.027 (0.018)	-0.007 (0.017)	0.032 (0.013)**	0.008 (0.020)	0.083 (0.017)***
Observations	1956	1956	1982	1982	3135	3135	2436	2436

Notes: * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.

Table 2.D.3: *Fraction of Households that had Another Child by Parity and Sex Composition in Albania (2002-2012) - Decomposed by Survey*

Sex of first two children in HH with 2 or more children	(2002) LSMS		(2005) LSMS		(2008-09) DHS		(2012) LSMS	
	Fraction of the sample	Fraction that had another child	Fraction of the sample	Fraction that had another child	Fraction of the sample	Fraction that had another child	Fraction of the sample	Fraction that had another child
<i>Panel B - Mothers aged 20 to 49 with 2 or more children younger than 18 years old</i>								
one boy, one girl	0.211	0.378 (0.026)	0.256	0.329 (0.023)	0.241	0.371 (0.019)	0.236	0.326 (0.022)
one girl, one boy	0.267	0.366 (0.023)	0.253	0.368 (0.024)	0.258	0.361 (0.018)	0.261	0.276 (0.020)
two boys	0.288	0.386 (0.022)	0.254	0.345 (0.023)	0.244	0.372 (0.019)	0.239	0.347 (0.022)
two girls	0.231	0.645 (0.025)	0.235	0.621 (0.025)	0.254	0.651 (0.018)	0.262	0.534 (0.022)
(1) mixed combination	0.479	0.372 (0.017)	0.509	0.348 (0.016)	0.500	0.366 (0.013)	0.497	0.300 (0.015)
(2) both same sex	0.520	0.501 (0.017)	0.490	0.478 (0.018)	0.499	0.514 (0.014)	0.501	0.445 (0.016)
Difference (2) - (1)	0.041 (0.025)	0.129 (0.024)***	0.019 (0.025)	0.129 (0.024)***	0.000 (0.019)	0.148 (0.019)***	0.003 (0.023)	0.145 (0.022)***
Observations	1558	1558	1553	1553	2542	2542	1827	1827

Notes: * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.

Table 2.D.4: First-Stage Specifications (Albania 2002, 2005 and 2012 LSMS)

Dependent Variable: Fertility (Number of children ever born)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>Panel A: Mothers aged 20 to 49 years old with 2 or more children younger than 18 years old</i>							
Boy first	—	—	—	-0.2415*** (0.0213)	-0.4652*** (0.0317)	-0.0262 (0.0243)	-0.0368 (0.0287)
Boy second	—	—	—	-0.2042*** (0.0215)	-0.4279 (0.0316)	0.0211 (0.0279)	—
Two girls	—	—	0.4694*** (0.0301)	—	—	0.4507*** (0.0413)	0.4296*** (0.0316)
Two boys	—	-0.1489*** (0.0254)	—	—	0.0209 (0.0279)	—	0.0201 (0.0279)
Same sex	0.2320*** (0.0228)	—	—	0.2248*** (0.0206)	—	—	—
R-squared	0.1127	0.1003	0.1492	0.2958	0.2647	0.2958	0.2958
Observations	4938	4938	4938	4862	4862	4862	4862
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	Yes	Yes	Yes	Yes
<i>Panel B: Husband's of the married women with 2 or more children younger than 18 years old</i>							
Boy first	—	—	—	-0.2444*** (0.0225)	-0.4789*** (0.0336)	-0.0099 (0.0287)	-0.0336 (0.0302)
Boy second	—	—	—	-0.2108*** (0.0225)	-0.4454*** (0.0333)	0.0237 (0.0291)	—
Two girls	—	—	0.4782*** (0.0309)	—	—	0.4691*** (0.0434)	0.4454*** (0.0333)
Two boys	—	-0.1484*** (0.0258)	—	—	0.0248 (0.0291)	—	0.0237 (0.0291)
Same sex	0.2355*** (0.0232)	—	—	0.2345*** (0.0217)	—	—	—
R-squared	0.1159	0.1031	0.1530	0.2839	0.2512	0.2838	0.2839
Observations	4726	4726	4726	4530	4530	4530	4530
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the PSU level are in parentheses. Other covariates in the models are the following: *Age, Age Squared, Age at First birth, Years of Education* and also indicators for *Boy 1st* and *Boy 2nd*. Each model is estimated with survey and district fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.

Table 2.D.5: OLS and IV Estimates of Mother's Labor-Supply Models in Albania (2002 - 2012)
(Mothers aged 20 to 49 years old with 2 or more children younger than 13 years old)

Method			OLS	IV	IV
Instrument for Fertility: (Number of children ever born)			—	<i>Same-sex</i>	<i>Two girls</i>
Survey Fixed-Effects			Yes	Yes	Yes
District Fixed-Effects			Yes	Yes	Yes
Controls			Yes	Yes	Yes
Dependent variable	Mean	Obs.	(1)	(2)	(3)
<i>Panel A: All databases (LSMS and DHS)</i>					
<i>Worked in last 7 days</i>	0.409	4390	-0.0263** (0.0115)	0.1114 (0.0801)	0.0590 (0.0570)
<i>Worked off-farm</i>	0.144	4390	-0.0265*** (0.0071)	0.0695 (0.0568)	0.1047** (0.0430)
<i>Worked on-farm</i>	0.211	4390	0.0133 (0.0100)	0.0052 (0.0642)	-0.0266 (0.0441)
<i>Self-employed</i>	0.056	4390	-0.0068 (0.0053)	0.0763* (0.0412)	0.0030 (0.0269)
<i>Panel B: Only LSMS databases</i>					
<i>Hours per week (total)</i>	15.34	2949	-0.4517 (0.5889)	12.0095** (5.1338)	7.8316** (3.3599)
<i>Hours per week (off-farm)</i>	6.356	2949	-1.3845*** (0.4208)	6.6192* (3.6856)	7.5173*** (2.5979)
<i>Hours per week (on-farm)</i>	7.295	2949	1.5401*** (0.4834)	1.7394 (3.5686)	-0.8690 (2.1904)
<i>Hours per week (self)</i>	2.111	2949	-0.2615 (0.2792)	4.8982* (2.6642)	1.2340 (1.6549)
<i>Second occupation</i>	0.014	2949	0.0078* (0.0047)	0.0314 (0.0291)	0.0032 (0.0197)
<i>Monthly Labor Income</i>	43824.28	2949	-6365.61*** (2321.12)	54116.80** (24697.23)	39320.05** (16289.33)

Notes: Standard errors clustered at the PSU level are in parentheses. All the models include the following covariates: Age, Age Squared, Age at First birth and Years of Education. IV Models in columns (2) and (3) also include indicators for *Boy 1st* and *Boy 2nd*. The sample includes mothers aged 20 to 49 years old with 2 or more children younger than 13 years old. Each model is estimated with survey and district fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.

Table 2.D.6: Secular Additive Effects of Child Gender on Labor Supply - Albania (2002-2012)

	<i>Worked in last 7 days</i>				<i>Hours worked per week</i>			
	OLS (1)	IV Same-sex (2)	IV Two girls (3)	IV Two girls (4)	OLS (5)	IV Same-sex (6)	IV Two girls (7)	IV Two girls (8)
<i>Sample: Mothers</i>								
Fertility	-0.0074 (0.0076)	0.0481 (0.0473)	0.0360 (0.0340)	0.0473 (0.0472)	-0.3171 (0.3984)	4.6978* (2.5509)	4.4091** (1.8866)	4.6397* (2.5424)
Boy first	-0.0031 (0.0115)	0.0101 (0.0159)	0.0073 (0.0135)	0.0099 (0.0159)	-0.2702 (0.5999)	0.8898 (0.8447)	0.8236 (0.7262)	0.8764 (0.8432)
Boy second	-0.0072 (0.0111)	0.0051 (0.0152)		0.0049 (0.0152)	-0.9292 (0.5815)	0.1113 (0.7923)		0.0993 (0.7909)
Observations	7,404	7,404	7,404	7,404	4,862	4,862	4,862	4,862
R-squared	0.1054	0.0986	0.1012	0.0988	0.0988	0.0887	0.0921	0.0894
Mean dep. var.	0.445	0.445	0.445	0.445	16.69	16.69	16.69	16.69
<i>Sample: Fathers</i>								
Fertility	0.0129* (0.0067)	0.0592 (0.0453)	0.0257 (0.0314)	0.0592 (0.0453)	0.1513 (0.4551)	-0.2749 (2.8674)	-1.1751 (2.0713)	-0.2749 (2.8674)
Boy first	0.0125 (0.0106)	0.0233 (0.0149)	0.0155 (0.0127)	0.0233 (0.0149)	0.7374 (0.6907)	0.6380 (0.9502)	0.4289 (0.8354)	0.6380 (0.9502)
Boy second	0.0047 (0.0102)	0.0153 (0.0145)		0.0153 (0.0145)	0.4927 (0.6637)	0.4020 (0.9096)		0.4020 (0.9096)
Observations	5,460	5,460	5,460	5,460	4,530	4,530	4,530	4,530
R-squared	0.1589	0.1516	0.1582	0.1516	0.0933	0.0931	0.0914	0.0931
Mean dep. var.	0.790	0.790	0.790	0.790	34.39	34.39	34.39	34.39
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the PSU level are in parentheses. All the models include the following covariates: *Age, Age Squared, Age at First birth and Years of Education*. The sample of mothers includes women aged 20 to 49 years old with 2 or more children younger than 18 years old. The sample of fathers includes the husbands of the married women with 2 or more children younger than 18 years old. Each model is estimated with survey and district fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.

Table 2.D.7: Secular Additive Effects of Child Gender on Labor Supply - Albania (2002-2012)

	<i>Worked off-farm</i>				<i>Second Occupation</i>			
	OLS (1)	IV Same-sex (2)	IV Two girls (3)	IV Two girls (4)	OLS (5)	IV Same-sex (6)	IV Two girls (7)	IV Two girls (8)
<i>Sample: Mothers</i>								
Fertility	-0.0325*** (0.0046)	0.0052 (0.0354)	0.0503* (0.0258)	0.0043 (0.0353)	0.0039 (0.0028)	0.0108 (0.0150)	0.0029 (0.0114)	0.0107 (0.0149)
Boy first	-0.0067 (0.0081)	0.0024 (0.0116)	0.0130 (0.0100)	0.0021 (0.0115)	-0.0042 (0.0035)	-0.0026 (0.0052)	-0.0044 (0.0044)	-0.0026 (0.0052)
Boy second	-0.0083 (0.0080)	-0.0090 (0.0112)		-0.0102 (0.0112)	0.0020 (0.0034)	0.0034 (0.0045)		0.0034 (0.0045)
Observations	7,404	7,404	7,404	7,404	4,862	4,862	4,862	4,862
R-squared	0.1850	0.1793	0.1575	0.1796	0.0150	0.0133	0.0149	0.0133
Mean dep. var.	0.164	0.164	0.164	0.164	0.014	0.014	0.014	0.014
<i>Sample: Fathers</i>								
Fertility	-0.0262*** (0.0081)	0.0354 (0.0545)	0.0671* (0.0401)	0.0354 (0.0545)	0.0072 (0.0044)	0.0320 (0.0291)	0.0406* (0.0230)	0.0320 (0.0291)
Boy first	-0.0054 (0.0129)	0.0089 (0.0174)	0.0162 (0.0154)	0.0089 (0.0174)	-0.0073 (0.0071)	-0.0015 (0.0094)	0.0005 (0.0084)	-0.0015 (0.0094)
Boy second	-0.0084 (0.0128)	-0.0144 (0.0175)		-0.0144 (0.0175)	-0.0091 (0.0073)	-0.0039 (0.0092)		-0.0039 (0.0092)
Observations	5,460	5,460	5,460	5,460	4,530	4,530	4,530	4,530
R-squared	0.1138	0.1048	0.0932	0.1048	0.0541	0.0480	0.0429	0.0480
Mean dep. var	0.394	0.394	0.394	0.394	0.058	0.058	0.058	0.058
Survey FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the PSU level are in parentheses. All the models include the following covariates: *Age, Age Squared, Age at First birth and Years of Education*. The sample of mothers includes women aged 20 to 49 years old with 2 or more children younger than 18 years old. The sample of fathers includes the husbands of the married women with 2 or more children younger than 18 years old. Each model is estimated with survey and district fixed effects. * Indicates statistical significance at 10%. ** Indicates statistical significance at 5%. *** Indicates statistical significance at 1%.

Chapter 3

Returning Home After Conflict Displacement: Labor Supply and Schooling Outcomes Among Kosovar Households¹

¹I am very grateful to my advisors, Karen Macours and Oliver Vanden Eynde, for their special support and guidance on the elaboration of this paper and to Christophe Bergouignan from Université Montesquieu - Bordeaux IV for providing me the 1999 Kosovo Socio-Demographic and Health Survey (DSHS). I also gratefully acknowledge all the helpful comments and suggestions from Gustavo Bobonis, John Giles, Pascaline Dupas, Suresh Naidu, Pamela Jakiela, Jeremie Gignoux, Christian Fons-Rosen, Iván Torre and all the participants of the Casual Friday's Development Seminar at Paris School of Economics, the European Doctoral Programme Conference and the Oxford Development Conference. I am also very thankful to Guadalupe Kavanaugh who provided excellent research assistance. The views expressed in this paper are those of the author and do not necessarily represent those of Paris School of Economics. I am responsible for all remaining errors.

Abstract

This paper uses the 1998-1999 Kosovo war and the following massive displacement of people as a natural experiment in order to estimate the impact of conflict displacement on labor market and education outcomes of Kosovars after they returned from exile. I exploit the interaction of the spatial variation in conflict intensity -as measured by casualties and bombings- and distance to the Albanian border as a source of exogenous variation in the displacement status. Results indicate that displaced Kosovar men are less likely to be employed in the agricultural sector and to work on their own account, while displaced Kosovar women are more likely to be inactive. Loss of assets (e.g. land, livestock) in an agrarian skill-based economy and also loss of social networks in an informal labor market might have further decreased the probability to find employment relative to stayers. However, shortly after the return home, the results also indicate that displaced Kosovar men and women are more likely to be working off-farm, especially in the construction and public administration sectors, which indicates a relatively quick recovery. In addition, displaced Kosovar girls are more likely to be enrolled in primary school, but I find no effect on education for boys. The refugee camp experience might have provided better conditions to young Kosovar girls compared to the precarious pre-war “parallel” education system.

JEL Classification: I20, J22, O12, O15

Keywords: conflict displacement, education, labor, instrumental variables

3.1 Introduction

Every year millions of people around the world are being forced to abandon their homes due to conflict, either as refugees or internally displaced persons (IDPs). According to the [UNHCR \(2017\)](#) Global Trends Report, by the end of 2016 the number of forcibly displaced individuals worldwide as a result of persecution, conflict, violence, or human rights violations reached 65.6 million, which has been the highest on record.² It is not just the scale of global forced displacement that is disconcerting but also its rapid acceleration in the recent years ([Martin, 2016](#); [IMDC, 2016](#); [Crawford et al., 2015](#)).

Migration and displacement may look very similar ways of movement of people, but while the former can be considered an optimization problem for the household, the latter is an exogenous shock to the household.³ Displacement is a direct side-effect of armed conflict, where individuals are forced to abandon their original place of residence due to life threatening situations. This condition puts individuals and families in a very vulnerable situation where they lose their social network, physical assets and often family members. However, little is known about the short-to long-term impacts of displacement on livelihoods. In order to identify policies that might mitigate the challenges and adverse conditions that the displaced people face, it is necessary to evaluate the effects of displacement on individuals so that post-war aid can be better targeted.

There is already an extensive economic literature on the impacts of voluntary migration and the impacts of war and violence, but the literature on the economics of forced displacement is still in its early stages.⁴ One of the principal reasons for the limited number of studies using quantitative methods is the lack of reliable data. Similarly, methodological difficulties in

²During 2016, 10.3 million people were newly displaced by conflict or persecution. This includes 6.9 million individuals displaced within the borders of their own countries and 3.4 million new refugees and new asylum-seekers. The [UNHCR \(2017\)](#) Global Trends Report can be found here: <http://www.unhcr.org/5943e8a34.pdf>.

³The discussion on the determinants of forced displacement, its definition and how forced migrants compare to “voluntary” migrants is out of the scope of this paper. See [Czaika and Kis-Katos \(2009\)](#); [Engel and Ibáñez \(2007\)](#); [Cortes \(2004\)](#) and [Stark \(2004\)](#) for this discussion.

⁴The literature on the impacts of conflict has found mixed consequences. For instance, a number of studies have found that civil war has little or no lasting effects on an area ([Brakman, Garretsen and Schramm, 2004](#); [Chen, Loayza and Reynal-Querol, 2008](#); [Davis and Weinstein, 2002](#); [Miguel and Roland, 2011](#)). Other studies have found that conflicts in fact might have positive impacts, especially in terms of political participation ([Valente, 2013](#); [Bellows and Miguel, 2009](#); [Blattman and Annan, 2010](#)). There is also evidence of long-run negative impacts from conflict on labor market and education outcomes ([Swee, 2015](#); [Akbulut-Yuksel, 2014](#); [Leon, 2012](#); [Blattman and Annan, 2010](#); [Shemyakina, 2011](#); [Kondylis, 2008](#)).

establishing exogeneity in the displacement shock complicates claims of causality. However, this literature is starting to gain attention in the last years as the micro data sets on conflict areas are becoming more available. There are few examples of quantitative estimates of the effect of displacement and the consequences seem to be mixed.⁵ For instance, [Sarvimäki, Uusitalo and Jantti \(2009\)](#) find increased mobility among displaced Finns due to WWII and consequently higher long-run incomes. Nevertheless, most of the previous literature suggests that there are serious negative consequences of forced displacement for those forced to migrate. [Fiala \(2015\)](#) finds a sizeable reduction in consumption smoothing for displaced households in Uganda. [Eder \(2014\)](#) analyzing post-war Bosnia, shows that displaced individuals invest less on their children's education. [Kondylis \(2010\)](#) also using data from post-war Bosnia, finds higher unemployment for men and lower labor force participation for women. [Bauer, Braun and Kvasnicka \(2013\)](#), analyzing the integration of Germans from Easter Europe, conclude that the first generation of migrants has lower incomes and ownership rates. [Abdel-Rahim, Jaimovich and Ylönen \(2015\)](#), studying displacement in Nuba Mountains of Sudan, conclude that displaced households hold fewer assets and are less involved in production. [Verwimp and Muñoz-Mora \(2018\)](#) investigate the food security and nutritional status of formerly displaced households in Burundi and they find that individuals who remain much longer in a displacement status are worse off compared to those who returned earlier.

This paper contributes to this literature by analyzing the impact of conflict displacement on labor market and education outcomes for the case of post-war Kosovo. During the 1998-1999 Kosovo war and especially during the NATO bombing campaign (March-June 1999), around 13,140 individuals were killed or went missing in Kosovo and more than 1 million were displaced either as refugees or IDPs, which represents approximately 70% of Kosovo's pre-war population. However, after the end of the conflict in June 1999, the displaced individuals started returning immediately to their previous residences and by the end of 1999 almost 95% had returned. The aim of this study is to use the Kosovo war and this massive displacement of people as a natural experiment in order to compare the labor market and education outcomes of those individuals that were displaced and decided to return relative to those that stayed in Kosovo.

For this purpose, I use two post-conflict individual and household survey data (e.g. 1999

⁵[Ruiz and Vargas-Silva \(2013\)](#) provides a literature review on the effect of displacement on migrating individuals as well as on hosting communities.

Kosovo Demographic Social and Health Survey and *2000 Kosovo Living Standard Measurement Survey*) containing rich information on labor market and education outcomes, displacement status and other individual characteristics. Both household surveys have several remarkable features that make them convenient for measuring displacement. For instance, they record the place of residence before, during and after the conflict for each individual and they were both collected post-war, that is, after most of the displaced individuals returned to their homes. Displacement status is defined using the place of residence: a person who reports having migrated during the period of the conflict —while resettlement or returned refugee is considered a displaced person, regardless of whether or not she resettled in her municipality of origin.⁶ I also use two municipality level data on conflict intensity: the *1998-2000 Kosovo Memory Book* database on casualties from the Kosovo war and the *Human Rights Data Analysis Group database on NATO bombing airstrikes*, which both can be geo-matched to the household data at the municipality level.

Despite the fact that displacement is to a great extent a forced action, it is still partly a result of a decision and therefore it is an endogenous variable.⁷ In order to reduce unobserved selection and biases that may be present in the displacement decision, I use an instrumental variables approach where I exploit the interaction of the spatial variation in conflict intensity and distance to the Albanian border as a source of exogenous variation in the displacement decision.⁸ This empirical strategy uses two sources of variation. First, the severity of the conflict is a good candidate to serve as instrument since the pattern of the Serb invasion in 1998/99 was governed by the will to create an ethnically homogeneous Serb territory. Likewise, the patterns of the NATO bombing raids generated the necessary fear to make people flee their homes (Ball et al., 2002; OSCE, 1999). Second, distance has been generally assumed in the literature to discourage migration by raising transaction costs. However, in this context distance

⁶Municipalities are the second political division of Kosovo, below districts and above villages or settlements. There were 29 municipalities in 1991 Kosovo, and some were divided after the war to form 30 municipalities. In 1991, the median population in the municipalities was 54,544 and the mean was 65,206, with a minimum of 4,611 and a maximum of 199,654.

⁷Conflict displacement is often a non-random event. Households are generally forced to leave their homes by rebels or army forces that take possession of their land, expand territorial control, weaken population support for opponent groups or increase their own support base and income. Therefore, it is likely that characteristics such as wealth or local visibility makes some households more prone to being displaced than others (Verwimp and Muñoz-Mora, 2018; Justino, 2011).

⁸The interaction-based instrumental variables technique has been previously proposed and used in the literature by Nunn and Qian (2014); Esarey (2015); Nizalova and Murtazashvili (2016) among others.

is used to capture affinity with the customs and culture over the border and mainly ethnic heterogeneity. Basically, areas with higher ethnic heterogeneity (i.e. located further away from Albania) were also more likely to suffer forced displacement. Indeed, most of forced expulsions in 1999 were carried out by the Serb forces in large towns across the north-eastern region of the province, which also corresponds with the most ethnically heterogenous municipalities in terms of Albanian and Serb populations ([OSCE, 1999](#)). Hence, I argue that forced displacement in Kosovo was highly influenced by conflict intensity and distance to the Albanian border.

This identification strategy relies on the idea that the relationship between the severity of the conflict and the decision to be displaced depends on the distance to the Albanian border, but the relationship between the severity of the conflict and the outcomes of interest does not depend on the distance to the Albanian border. In practice, I use the war casualties and the number of NATO bombing days at the municipality level as two proxies for the level of conflict intensity. While, distance to the Albanian border is measured as the driving distance (in kilometers) from the village of residence to the south-west Kosovo-Albanian border of Morina.

However, one potential concern that might threaten the exogeneity assumption is that pre-war local economic conditions might predict local conflict intensity and distance to the Albanian border. For instance, locations closer to the Albanian border were more likely to have a higher proportion of Albanians before the war. This is of interest since ethnicity might have determined individual economic status through ethnic discrimination. Historical evidence suggests that casualties and bombings in Kosovo were not determined by pre-war economic performance at the local level since the primary aim of Serb attacks was territorial separation and ethnic cleansing. Similarly, NATO's objective was to attack only targets of military nature (i.e. military facilities, equipment, weapons etc) regardless of the economic performance of the different regions ([Grant, 1999](#); [ICTY, 2000](#)). Therefore, since the pattern of conflict in Kosovo was likely driven by geo-strategic motives rather than economic motives, it is plausible to argue that conflict intensity interacted with distance to the Albanian border is likely to be orthogonal to unobserved factors that might affect schooling and labor market outcomes.

Even though historical references suggest that in Kosovo there was no targeting of individuals and regions based on the local economic differences, in order to address this potential concern I control for pre-war labor-force participation and pre-war proportion of Albanians at the municipality level. In addition, I also perform several robustness checks and conduct

placebo tests with different samples in order to assess the validity of the exclusion restriction and also to reinforce the results obtained from the IV estimation.

The first-stage results indicate that further away from the Albanian border, an increase in conflict intensity increases the likelihood of being displaced. Indeed, according to the historical references Kosovar Albanians living in municipalities with more ethnic heterogeneity were more likely to be displaced by the Serb forces. The second-stage results show that conflict displacement impacted negatively but also positively the labor market and education outcomes of Kosovars who were forced to abandon their homes relative to those who stayed.

Firstly, I find that, in the short-run, conflict displacement had a negative impact on labor market outcomes of Kosovar men and women, particularly in terms of access to employment. More specifically, these results show that displacement is associated to a significant and large increase in women's inactivity and to a decrease in men's self-employment and their employment in the agricultural sector. Interestingly, I also find that shortly after the return in Kosovo, conflict displacement also had a positive impact on labor market outcomes. The medium-term results indicate that both displaced Kosovar men and women are also more likely to be working off-farm (i.e. construction and public administration sectors). One possible explanation for these findings is that loss of assets, land and livestock in an agrarian skill based economy must have made it very difficult for returned refugees to find employment. Another plausible mechanism behind these results is the loss of social networks in an informal labor market, which is fundamental for the job search in these type of transition economies.

Secondly, the results in terms of education outcomes show that, in the short-run, while displaced Kosovar girls are significantly more likely to be enrolled in primary school relative to those who stayed, displacement does not seem to have any effect on Kosovar displaced boys or teenage girls. One possible channel through which this effect might be operating is the refugee camp experience. Young female refugees, especially those who were in camps, might have had better access to basic education and better conditions than the IDPs and the stayer girls after taking into account the pre-war precarious context of the "parallel" education system in Kosovo.

This paper contributes to the growing literature on the economics of forced displacement at the microeconomic level. To the best of my knowledge, this study is the first to consistently analyze the casual effect of conflict displacement in the immediate post-conflict period in

Kosovo and to also provide empirical evidence on the potential mechanisms behind the results. It assess the effect of conflict-induced displacement on labor market and education outcomes accounting for potential selection issues by using a novel interaction-based instrument involving conflict intensity and distance. Lastly, Kosovo constitutes an interesting case study for this analysis as it is one of only a small number of countries for which detailed conflict intensity and conflict displacement information is available for the immediate period after the conflict.

The rest of the paper is organized as follows: Section 3.2 provides background information on the war and conflict displacement in Kosovo, Section 3.3 presents the databases used in the analysis and some descriptive evidence, Section 3.4 presents the instrumental variables empirical strategy and discusses the identifying assumptions, Section 3.5 presents the results, Section 3.6 sheds some light on the plausible channels on each one of the outcomes, while Section 3.7 briefly concludes.

3.2 Background

3.2.1 Kosovo War (1998-1999)

Kosovo is a partially recognised state in the Balkans with a long history of ethnic diversity and conflict. Just before the war, in 1998, it's population was around 2,1 million, of which 83% were Albanians, 10% were Serbs and 7% belonged to other ethnicities. Ethnic identity has always been analogous to religious identity, as Albanians are predominantly Muslims, while Serbs are Orthodox Christians ([Brunborg, 2002](#)).

From 1989 (when Kosovo's autonomous status within Serbia was partially revoked) till 1998, the majority of Kosovo Albanians lived in a situation similar to an apartheid, in which they were denied access to jobs and services, and were unable to exercise basic rights. As a result, the Kosovo Albanians established parallel systems of institutions for almost every aspect of daily life, including employment, education and health. Also, the continued discrimination and repression by the Serbs led to the emergence of an armed insurgency group of Albanians which was called the Kosovo Liberation Army (KLA). During the '90s, the KLA launched several attacks targeting Serbian law enforcement in Kosovo.⁹

⁹Between 1989 and the beginning of 1998, an estimated 350,000 Kosovo Albanians left the province at one stage or another, most of them going to countries in Western Europe.

Given this situation, in March 1998, Serb forces engaged in an indiscriminate military campaign of “ethnic cleansing” against KLA and Albanian civilians. Their aim was to expel all the Albanians from Kosovo in order to create an ethnically homogenous territory. After one year of continued ethnic tensions and violent confrontations between the Albanians and Serbs and after several failed attempts at a diplomatic solution, NATO intervened on March 24th 1999 with a bombing campaign against the Republic of Serbia, including attacks on targets in Kosovo. The NATO air campaign was justified in order to stop the actual and potential killings and expulsions of Kosovo Albanians by Serbian forces (Cutts, 2000).

Finally, after a 78-day air campaign, on June 9th 1999, the Republic of Serbia accepted a peace plan that required the withdrawal of all Serb forces from Kosovo, the safe and free return of all refugees and displaced people, and the establishment of a UN mission.

3.2.2 Forced Displacement from Kosovo

As a consequence of the ethnic cleansing and the NATO bombing campaign, Kosovo suffered one of the largest population displacements in Europe since WWII. On the one hand, reports by the *Humanitarian Law Centre* (HLC) in Belgrade and Kosovo estimate that approximately 13,535 civilians and soldiers were killed or missing. On the other hand, the *United Nations High Commissioner for Refugees* (UNHCR) estimates that around 1.4 million people were displaced from their homes, of which around 850,000 sought refugee protection out of Kosovo and around 600,000 were internally displaced persons.¹⁰ Figure 3.A.1 shows a map of the distribution of the displaced populations from Kosovo in neighbouring countries/territories and Figure 3.A.2 shows the cumulative refugee population over time and by country of destination from March till June 1999.¹¹

However, after the end of the war in June 1999, the refugees started returning immediately. Within three weeks, 500,000 people had returned, and by the end of 1999, more than 800,000 had returned to their homes (including people who had left before the NATO air campaign). In particular, out of approximately 850,000 Kosovo refugees during the war, by October 1999 around 65,500 individuals remained displaced and by May 2000 the number had dropped to

¹⁰*Kosovo Crisis Update*, June 11, 1999. Geneva, Switzerland: United Nations High Commissioner for Refugees

¹¹Of those that were expelled from Kosovo after the start of the air campaign, some 450,000 went to Albania, some 242,000 to the former Yugoslav Republic of Macedonia (FYR Macedonia), some 70,000 to Montenegro and some 96,000 participated of the Humanitarian Evacuation Programme (HEP) which allowed them to go to other countries such as Germany, USA, Turkey, France, Italy etc (Cutts, 2000).

around 40,000 (Cutts, 2000). Figure 3.A.3 shows the cumulative returned refugee population since the end of the Kosovo war.

It is important to note that the data used in this analysis is limited to displaced persons who by 1999-2000 returned to Kosovo, but clearly it does not include individuals that preferred not to come back before the collection of the household surveys. Returning home from conflict displacement is also a non-random event. In general, households that are poorly integrated in the host economy or with more assets at their original home may be more likely to return (Arias, Ibáñez and Querubin, 2014). According to the Kosovo Agency of Statistics (2014) the number of Kosovo residents that reported to have migrated during the 1998-1999 Kosovo war and decided not to come back was around 50,000 individuals by 2011. This number of non-returned individuals represents 5-6% of the total displaced refugee population, which is quite small in order to generate a problem of selected sample of the displaced individuals.

In order to better understand the nature of this selection, I also use the 1999 Kosovo DSHS, 2000 Kosovo LSMS and the 2012 Kosovo Remittances Survey to compare the educational attainment of the displaced who returned to Kosovo to that of emigrants from Kosovo to other countries who left the country due to the 1998-1999 war and never came back. I find that the proportion of individuals having achieved higher education is similar across groups, while the proportion of individuals having low (primary education) and medium education (secondary) differs across groups: around 40% of the displaced who returned have low education, compared to 31% of the emigrants from Kosovo to other countries; and around 38% of the displaced who returned have medium education, compared to 46% of the emigrants that never returned to Kosovo.

3.3 The Data

This study uses four data sources: two household-individual level surveys and two conflict intensity databases. Firstly, the individual level surveys are: the 1999 *Kosovo Demographic, Social and Health Survey* (DSHS) and the 2000 *Kosovo Living Standard Measurement Survey* (LSMS). Secondly, the conflict intensity datasets are municipality level data on war casualties from the 1998-2000 *Kosovo Memory Book* (KMBD), and municipality level data on reported bombing days from the *Human Rights Data Analysis Group* (HRDAG) database on NATO airstrikes.

The 1999 Kosovo DSHS was carried out by United Nations Population Fund (UNFPA), the

International Organization for Migrations (IOM), and the Statistical Office of Kosovo from November 1999 to February 2000, which is just after the conflict. While, the 2000 Kosovo LSMS was conducted by the World Bank from September to December 2000, which is over a year after the end of the NATO air campaign that terminated the conflict in Kosovo. These surveys are both representative on the national as well as on the regional level. The sampling procedure was stratified by region (7 regions in the DSHS and 5 regions or areas of responsibility in the LSMS) and by sector (rural and urban). The 1999 DSHS covered 27 out of 29 municipalities and interviewed a total of 7,343 randomly selected households and 40,918 individuals. The 2000 LSMS covered 29 out of 30 municipalities and was administered to a total of 2,880 randomly selected households and 17,917 individuals.¹² Both household surveys contain a rich set of information on demographics, education, labor activities, health, conflict displacement and other characteristics.

The Kosovo Memory Book Database (KMBD) is a joint project between the *Humanitarian Law Centre* (HLC) in Belgrade and the HLC in Kosovo. This project collected detailed information on casualties between 1998 and 2000 in connection to the war in Kosovo, which are document based on death records, statements by surviving family members and witnesses. This database contains the victims's vital information at the time of death, including name, age, ethnicity, location of the incident, date of the incident, type of casualty (civilian or military status) etc. Overall, the Kosovo Memory Book indicates that 13,140 individuals were killed or missing in Kosovo, with an average of 437 casualties per municipality. From the total number of victims, around 76% are civilians, while 24% are armed forces. Based on several analysis and findings, including a comparison with ten other databases in which no new death records were found, the KMBD was found to have more records than any other database in every period and for each municipality ([Krüger and Ball, 2014](#)). I use the KMBD municipality level data on war casualties and the 1991 population census to compute the war casualty rate- as the total number of casualties per 1,000 inhabitants at the municipality level.¹³ This variable offers a measure

¹²Note that until 1999, Kosovo had 29 municipalities. The municipality of Malisevo was part of four other municipalities (Klina, Orahovac, Suva Reka and Glogovac) and did not exist until July 2000, when it was re-established by the United Nations Mission in Kosovo (UNMIK).

¹³It is important to note that the 1991 Kosovo population census was boycotted by the Kosovar Albanian population. To compensate for this the FRY statistical office (FSO) in Belgrade estimated the size of the Albanian population on the basis of the 1981 census results taking into account population changes during the intercensal period 1981-1991. The 1991 population census data at the municipality level that is used in this study is taken from [Brunborg \(2002\)](#).

of conflict incidence at the local level. A municipal map of Kosovo which shows the spatial variation in war casualties per 1,000 Kosovo inhabitants is presented in Figure 3.2.

The NATO airstrikes database of the HRDAG records the number of reported bombing attacks occurring in each municipality per day during the NATO air campaign (March-June 1999). These bombing records are derived from a report published by the Human Rights Watch (HRW) in February 2000, which contains daily information on bombings based mostly on different Serbian government sources and Serbian newspapers, but they are also based on NATO's reports in the Operation Allied Force Update. No effort was made to quantify the severity of each airstrike, but reports of different airstrikes were counted separately. From this database, I compute bombing intensity as the total number of days a municipality was attacked with bombs and missiles during the NATO Air Campaign (78 days). Figure 3.3 shows the spatial variation in bombing intensity from March-June 1999 across municipalities.

These measures of local conflict incidence - i.e. war casualty rate and bombings at the municipality level - are used to instrument for displacement in the subsequent regression analysis.

3.3.1 Measuring Conflict Displacement

The 1999 Kosovo DSHS and the 2000 Kosovo LSMS have several attractive features which make them convenient for measuring forced displacement. Firstly, they both contain several self-reported outcomes which are used to explicitly identify each individual that was displaced during the 1998-1999 Kosovo war. In particular, forced migrants are identified using the following questions: *"How many times did you change residence since the beginning of the conflict (March 1998)?"*, *"What was the main reason for this displacement?"*; to which the five answers are: security, house inhabitable, work, study, other. Given this, I exclude from the analysis all individuals who declared having moved for a job, for studies and those that moved for other reasons, using the forcibly displaced as the treated group and the non-movers or stayers as the control group.¹⁴

Secondly, both household surveys record the place of residence pre, during and post-conflict

¹⁴In the 1999 Kosovo DSHS 97.33% declared being forcibly displaced while only 2.67% declared being willingly displaced. In the 2000 Kosovo LSMS 98% declared being displaced due to security reasons, 0.42% due to house inhabitable, 0.17% due to work, 0.03% due to study and 0.48% due to other reasons. Therefore, the sample excluded is clearly very small and negligible. Even if I do not exclude these few observations from the analysis and I put them in the treatment group as displaced individuals, the main results remain unaltered. (Results upon request).

for the forcibly displaced and the non-displaced. Specifically, this information is identified through the following question: “*Where did you live immediately before the conflict (March 1998)?*”, with the following categories: here (site of survey), other municipality, Former Yugoslavia, Albania, Serbia, Western Europe, other and not yet born. As the municipality of residence before the war is recorded for all individuals, this allows me to geo-match the measures of conflict intensity at the municipality level to each individual regardless of displacement status. These measures of local conflict intensity are used to instrument for displacement in the subsequent regression analysis.

Conversely to other studies, both surveys are successful in differentiating between refugees (displaced persons who went into exile during the time of the conflict) and internally displaced persons (IDPs) (individuals who resettled in camps or in other locations within Kosovo during the conflict). The individuals that declared being forcibly displaced were also asked “*Where were you living during most of this absence?*”; to which the answers are: other locality but same municipality, other municipality, Former Yugoslavia, Albania, Serbia, Western Europe and other. The 1999 Kosovo DSHS, in particular, also records whether an individual went to a refugee camp. Given this information, I only include individuals who resided in Kosovo pre-war, excluding individuals that lived in other parts of former Yugoslavia or some other country before March 1998.

Thirdly, both surveys were collected post-war, which coincides with the return of the majority of the refugees and IDPs to their homes. The 1999 Kosovo DSHS was implemented five months after the end of the war (November 1999), while the 2000 Kosovo LSMS was implemented almost one year and a half after the end of the war (September 2000). Figure 3.1 shows a timeline of the cumulative refugee flows (displaced refugees and returned refugees) during the war and the implementation of the household surveys in 1999 and 2000, which indicates that the vast majority of the refugees had returned in Kosovo before the start of the collection of both household surveys. Moreover, Figure 3.A.4 shows the patterns of daily returned refugees as recorded by UNHCR and Figure 3.A.5 shows the patterns of first and last displacement as recorded in the 1999 Kosovo DSHS. Both of them clearly indicate that most of the refugees had returned in Kosovo by the end of September 1999.

3.3.2 Descriptive Statistics

For the labor market outcomes analysis I restrict the sample to men and women aged 20-65 years old and for the education outcomes analysis, I restrict the sample to boys and girls aged 6-19 years old. Table 3.1 shows descriptive statistics of forced displacement by gender and age group for each database. In general, 60-70% of the individuals in each sample were forcibly displaced, which is very similar to the UNHCR estimates of 1.4 million displaced individuals from a population of 2.1 million (approx. 67%). Women and children were more likely to be displaced compared to men, indicating that a proportion of the men stayed in the province to fight in the war. In particular, 40-45% chose to move out of Kosovo (refugees), while 20-25% were displaced persons inside Kosovo (IDPs). Also, around 19% decided to go to a refugee centre, which means that around 2/3 of the displaced population went to host families. The return pattern of the displaced population indicates that 87-96% had the same municipality of residence as before the start of the war (March 1998). Due to the fact that several homes remained inhabitable after the war, some returnees remained displaced and could not return to their previous residence.

Table 3.2 presents descriptive statistics of the displacement status by ethnic group. More than 95% of the displaced population are ethnic Albanians, which clearly gives some evidence of the "ethnic cleansing" campaign of the Serbs against the Albanians. However, among the non-displaced population, Albanians also constitute the majority (70%), followed by Serbs (20%) and other ethnic groups (10%). Figure 3.A.6 and 3.A.7 show the spatial variation in the proportion of displaced individuals across municipalities for each household survey. While, Figure 3.A.8 shows the spatial ethnic distribution in 1991 across municipalities. The north-western, north-eastern and central regions have the highest proportions of displacement, which also coincides with the more ethnically heterogeneous municipalities.

Figure 3.A.9 shows labor market status by gender and age group for the 1999 DSHS and 2000 LSMS, respectively. The labor market variables are measured slightly different in each database. For instance, in the 1999 DSHS activity status is measured through categories: *employed*, *self-employed*, *contributing family worker*, *unemployed* (seeking work) and *inactive* (housewife, retired and other). Duration in each activity status is not specified in this survey. While, in the 2000 LSMS, *employed* is defined as having done any work (i.e. off-farm, on-farm, self-employed) during the last week. *Unemployment* is defined as having looked for a job during the last week.

The *inactive* are individuals neither in work, unemployed, nor attending school. The 2000 LSMS also reports usual weekly hours, which is used to measure hours of work.

In 1999, 88% of men aged 20-65 are economically active, but half (42%) are unemployed. In 2000, the proportion of economically active men lowers to 75%, of which 64% are employed and only 11% are unemployed. Only 32% of women in the same age group are economically active in 1999 and 36% in 2000, but their unemployment rate is a lower than the rate for men. Unemployment rates are very high among young adults aged 20-25, but these rates decrease with age. After age 40, unemployment rates are less than 40% for men and women alike. The inactive population is considerable at all ages for women (around 65%) and is composed mostly by housewives, while for men is much lower but increases substantially after the age of 50. Women work 34 hours per week, while men work 44 hours per week on average.

Some descriptive statistics on children's enrolment rates by gender are shown in Figure 3.A.10. Enrolment is measured as a dummy variable indicating whether a child is registered in primary, secondary or university during the 1999-2000 and 2000-2001 academic year, respectively. Both graphs indicate that around 78% of children aged 6-19 years old are enrolled in school. Primary or compulsory school age boys and girls (aged 6-14) are virtually all enrolled in school (more than 90%), with equality between genders. However, girl's seem to drop-out in the last years of primary school. Enrolment rates for secondary school children aged 15-19 drop to approximately 58% and the gender gap within this group is quite dramatic, with only 50-55% for girls versus 60-65% for boys. Approximately 20% of young people aged 19-25 are enrolled in higher education, with near-equality between genders. In this analysis, I focus only on primary and secondary enrolment.

3.4 Empirical Strategy

In order to measure the effect of conflict-displacement on labor market and education outcomes, the basic regression model can be represented by the following equation:

$$Y_{im} = \beta_1 D_i + \beta X'_{im} + \varepsilon_{im} \quad (3.1)$$

where Y_{im} represents the outcome of interest (e.g. work off-farm, work on-farm, self-employed, child enrolled in school etc) for individual i residing in municipality m after the war.

D_i is a dummy variable that indicates whether a person i was displaced due to the Kosovo war, X'_{im} is a vector of individual controls and ε_{im} is the unobserved individual heterogeneity.

Even though displacement is to a great extent a forced action, it is partly a result of a decision, and therefore it is an endogenous variable. This endogeneity issue can be clearly observed through the patterns of conflict displacement at the municipality level. Firstly, even in the most war-affected municipalities, the western part of Kosovo (see Figures 3.2 and 3.3), we do not observe the displacement of the entire population. For instance, only half of Djakovica municipality's population (50-56%) was displaced, even though this municipality is one that suffered the most from the war, either through casualties or bombing attacks. Secondly, in both surveys there are individuals who declared being displaced even if they resided in less war-affected municipalities, such as those in the north of Kosovo.

This patterns of conflict displacement suggest that it is possible for individuals to "self-select" into or out of displacement. As a result, those who leave could be different from those who stay in terms of unobservable characteristics that may also make them more (or less) successful in terms of post-war outcomes. In other words, there might be unobserved omitted variables, such as individual heterogeneity in preferences, ability etc, that might affect both displacement and outcomes. Also, pre-war socio-economic conditions might play an important role at the moment of displacement, resulting in reverse causality. [Kondylis \(2010\)](#), [Czaika and Kis-Katos \(2009\)](#) and [Ibáñez and Vélez \(2008\)](#) show that pre-war economic conditions are important determinants of the displacement decision, even when facing conflict and war violence. For example, if well-endowed households who are better able to cope with war have lower propensity of displacement, then the proportion of well-endowed individuals will be greater in high conflict intensity municipalities. Conversely, the opposite could also be true if well-endowed individuals have better outside opportunities (in employment or schooling) and are thus more likely to move. Failing to account for such endogeneity issues means that estimating the impact of displacement on outcomes by a simple OLS estimation might give biased and inconsistent estimates of β_1 . The following sub-section describes the identification strategy used in this study to disentangle the effect of displacement from the effect of conflict or war.

3.4.1 Identification: Instrumental Variables Approach

One way to address the potential endogeneity in the displacement decision is to use a recent methodological innovation based on interaction-based instrumental variables (Esarey, 2015; Nunn and Qian, 2014). This empirical strategy exploits the interaction of the spatial variation in conflict intensity and distance to the Albanian border as a source of exogenous variation in the conflict displacement decision. In order to identify the local average treatment effect (LATE), the instrument must satisfy two basic conditions: (1) to be correlated to displacement; (2) to satisfy the exclusion restriction, which means that it must not be correlated to factors directly affecting labor market and education outcomes.

The first obvious candidate to serve as an instrument is the severity of the conflict in the location of origin. Empirically, conflict intensity is measured through war casualties and bombings at the municipality level.¹⁵ In order to motivate the relevance condition, Figure 3.A.11 shows the estimated total refugee migration and casualties over time, while Figure 3.A.12 shows the estimated total refugee migration and bombing reports over time (March-June 1999). Figure 3.A.11 suggests that the observed pattern of casualties closely resembles the pattern of refugee flow during the whole period of the conflict, while Figure 3.A.12 indicates that NATO's activity coincides with the refugee flow only for the first part of the conflict (till the end of April). Bombing intensity increases substantially after the largest number of casualties and highest levels of refugee flow. Given that bombing intensity is consistent with the patterns of refugee flow only for the first period of the conflict, I will exploit this fact in order to disentangle the effect of displacement from the effect of conflict by using the bombing intensity measure only for this first period.¹⁶

The second candidate to serve as an instrument is distance to the Albanian border, since forced displacement was more intense further away from Albania and especially in municipali-

¹⁵Kondylis (2010) also uses conflict incidence as an instrument for conflict displacement in the context of the post-war in Bosnia and Herzegovina. She uses the municipality level population losses data (which reports the ICTY casualties estimates) and the 1991 census in order to compute the proportion of the pre-war population that went missing in each municipality during the conflict.

¹⁶Ball et al. (2002) also studies the statistical patterns of refugee flow and killings in Kosovo during the period March-June 1999 using only data from the Albanian border guard registries of people entering Albania through the village of Morina. The authors find that the killings and the exodus of refugees occurred in the same places at roughly the same times, implying that the common cause of both phenomena was a systematic military campaign by Serbian forces aiming to expel Kosovar Albanians from their homes. This study was used as evidence at the International Criminal Tribunal for the Former Yugoslavia (ICTY) in the case against Slobodan Milosevic.

ties with ethnic heterogeneity.¹⁷ Distance to the Albanian border is measured as the driving distance (in kilometers) from the village of residence to the south-west Kosovo-Albanian border of Morina. In general, distance has been assumed in the literature to discourage migration by raising transaction costs. However, in this context distance is used to capture mainly ethnic diversity. Basically, areas with higher ethnic heterogeneity were also more likely to suffer forced displacement. Indeed, according to a report from [OSCE \(1999\)](#) most of forced expulsions in 1999 were carried out by the Serb forces in large towns across the north-eastern region of the province, more precisely from Kosovska Mitrovica to Pec and from Pec to Pristina, which also corresponds with the most ethnically heterogeneous municipalities in terms of Albanian and Serb populations. In addition, even though there were several borders from where refugees could have left the province, in most of the cases the Serb forces closed the northern borders and diverted the convoys mainly south-west in order for the refugees to have no other choice but to go to Albania.¹⁸

A remaining econometric concern with these instruments is that using them separately might violate the exclusion restriction, in the sense that each instrumental variable might have an independent impact on post-war outcomes beyond any effects working through conflict displacement. For instance, pre-war local economic performance might predict local conflict incidence. Similarly, proximity (remoteness) to the Albanian border is likely to be associated to lower (higher) incomes during the pre-war period. However, given that changes in the level of violence had a larger effect on forced displacement for individuals residing in areas with higher ethnic heterogeneity and located further away from Albania, I argue that this concern can be addressed if the level of violence is interacted with distance to the Albanian border.

Basically, the idea behind this identification strategy is that the relationship between the severity of the conflict and the decision to be displaced is conditional on the distance to the Albanian border, but the relationship between the severity of the conflict and the outcomes of interest does not depend on the distance to the Albanian border. In other words, being located further away from the Albanian border strengthens the relationship between the severity of

¹⁷See Figures 3.A.6, 3.A.7 and 3.A.8.

¹⁸This type of displacement was particularly true for those refugees from the north of the province. For instance, many refugees from Kosovska Mitrovica and the surrounding area were not sent north to Leposavic, west towards Rozaje (Montenegro) or southwards down the main route to the Former Yugoslav Republic of Macedonia. Instead they were compelled to take very roundabout routes south-west along minor roads, eventually reaching Prizren and then Albania ([OSCE, 1999](#)).

the conflict and displacement because those Albanian Kosovars that were living closer to the Serbian border were more likely to be expelled and displaced from their homes compared to those that were living closer to the Albanian border. Thus, even if there was endogeneity between conflict intensity and the outcome of interest, the exclusion restriction would only be violated if the unobserved variables driving this endogeneity were also correlated with distance to the Albanian border (for more econometric details see [Nizalova and Murtazashvili \(2016\)](#); [Esarey \(2015\)](#)).

At the same time, there is little reason to believe that the impact of conflict intensity on the outcomes of interest is conditional on the distance to the Albanian border. Therefore, the interaction term (conflict intensity * distance to Albanian border) is a reasonable candidate instrument since it is likely to accurately predict displacement and at the same time is likely to be orthogonal to unobserved factors that might affect schooling and labor market outcomes. Casual inference using the interaction-based instrument relies on the assumption that, conditional on the controls, the interaction between conflict intensity and distance to the Albanian border only affects labor market and education outcomes through forced displacement. Since the validity of the instrument is central to this identification strategy, in the following sub-sections I provide some historical evidence and I also perform some robustness checks in order to assess its validity.

3.4.2 First-Stage Estimation

In order to account for the potential endogeneity in the displacement status, I use the interaction of conflict intensity and distance to the Albanian border as instrument for conflict displacement as follows:

$$Y_{im} = \beta_1 \hat{D}_i + \beta X'_{im} + v_{im} \quad (3.2)$$

$$\hat{D}_i = \alpha_1 (WCR_{m_o} * DA_{v_o}) + \alpha_2 WCR_{m_o} + \alpha_3 DA_{v_o} + \alpha X'_{im} + \epsilon_{im} \quad (3.3)$$

$$\hat{D}_i = \delta_1 (B_{m_o} * DA_{v_o})_{io} + \delta_2 B_{m_o} + \delta_3 DA_{v_o} + \delta X'_{im} + \epsilon_{im} \quad (3.4)$$

Equation 3.2 is the second stage of the 2SLS system and equations 3.3 and 3.4 are first stage specifications using the two different measures of conflict intensity. In each first-stage model, I regress the dummy for displacement status D_i of individual i on the interaction term between conflict intensity - measured as war casualty rate or bombings- and distance to the Albanian

border. WCR_{m_o} denotes the number of casualties per 1,000 inhabitants at the municipality of origin m_o and B_{m_o} denotes the number of days the municipality of origin m_o of individual i was attacked by NATO airstrikes. DA_{v_o} denotes distance from village of residence of individual i to the south-west Albanian border of Morina.

For the education outcomes analysis, controls include: age, ethnicity dummy (Albanian), dummies for parental educational attainment (medium and high), number of male and female adults in a household aged 20 to 65, number of siblings, distance to school, and dummy for rural location. Similarly, for the labor market outcomes, controls include: age, dummies for marital status, ethnicity (Albanian), dummies for educational attainment (medium and high), number of male and female adults in a household aged 20 to 65, number of dependent members by age group, and dummy for rural location. I also control for pre-war socio-economic conditions by including labor-force participation and proportion of Albanians in 1991 at the municipality level.

Tables 3.3 and 3.4 present the regression coefficients of the first-stage estimation for the children's sample and adult's sample, respectively. The results are shown separately for each database and the reported standard errors are clustered at the village level and municipality level. A more conservative inference requires to cluster the standard errors at the municipality level. However, in this analysis this may not be sufficient since I rely on less than 30 clusters (municipalities in Kosovo). In case of few clusters, clustered-robust standard errors may be under-estimated. Hence, I correct the inference with wild bootstrap methods as suggested by [Cameron, Gelbach and Miller \(2008\)](#) and [Cameron and Miller \(2015\)](#). This procedure allows to account for the correlation in the error terms of individuals born in the same municipality with few clusters. In the Appendix, I provide the P-values resulting from wild bootstrap for the second-stage results.¹⁹

Using the 1999 Kosovo DSHS database, it seems that the instrument ($WCR_{m_o} * DA_{v_o}$) is a good predictor for displacement status, while ($B_{m_o} * DA_{v_o}$) does not seem to be a valid instrument. While, for the 2000 Kosovo LSMS database, both instruments seem to be good predictors for displacement.²⁰ Even though these instruments are based on only 27/29 municipalities or

¹⁹Wild bootstrap P-values are obtained with the post-estimation command *boottest* by [Roodman \(2017\)](#), using Redmarcher weights, assuming the null hypothesis and setting replications to 1,000.

²⁰This difference in first-stage results is plausibly due to the different samplings in both databases. For instance, the municipalities of Zvecan and Malisevo are not included in the 1999 Kosovo DSHS database.

pre-war residence, they are highly significant for both females and males. The F-statistics of the excluded instruments are always above 10 when the standard errors are clustered at the village level for both children and adults samples, but they decrease slightly when the standard errors are clustered at the municipality level.

In general, these results indicate that near the Albanian border, an increase in conflict intensity (as measured by casualties or bombing) decreases the likelihood of being displaced; while far from the Albanian border, an increase in conflict intensity increases the likelihood of being displaced. This finding is in line with the historical fact that when the war started Kosovar Albanians living further away from Albania were more likely to be expelled from their homes because towns located in the north-eastern part of the province were more likely to be targeted by the Serb forces due to their ethnic heterogeneity. Overall, these results show that conflict intensity interacted with distance to the Albanian border is a good predictor of forced displacement in the context of the 1999 Kosovo war.

The instrumental variables approach estimates the impact of displacement for those individuals that were induced by the conflict and the residential characteristics, such as ethnic heterogeneity, to be forcibly displaced from their homes i.e. *local average treatment effect*. In other words, in this setting *compliers* are those individuals that were more likely to be forcibly displaced because their municipalities of residence suffered more from war casualties/bombings and also because these municipalities were more ethnically heterogeneous in terms of Albanian and Serb populations (i.e. located further away from the Albanian border). While it is not possible to observe whether individuals in a given municipality decided to move in response to an increase in conflict intensity and distance to the Albanian border, Tables 3.A.1 and 3.A.2 shed light on which municipalities were influenced by the interaction-based instrument by examining the size of the first-stage for different sub-populations.²¹

Column 1 reports the baseline first-stage relationship from the pooled sample of women and men for comparison purposes. Columns 2 and 3 divide the sample by whether the municipality had a higher labor force participation in 1991 than the median municipality. The correlation between the interaction-based instrument (conflict intensity * distance to the Albanian border) and conflict displacement is statistically significant in both samples but it is slightly larger in municipalities with more labor supply. Next, columns 4 and 5 divide the sample by whether

²¹This technique has been already used in Dell (2012) in order to better understand the characteristics of the compliers.

the municipality had a higher percentage of its population working in agriculture in 1991 than the median municipality. In this case, the interaction-based instrument has more power in municipalities with a higher proportion of the population working in agriculture. Lastly, columns 6 and 7 divide the sample by whether the municipality had a higher percentage of the population speaking Albanian in 1991 than in the median municipality. The correlation between the instrument and conflict displacement is statistically significant in both samples, but it is larger in municipalities with less Albanian speakers in 1991.

Overall, these results document that the interaction-based instrument -conflict intensity and distance to the Albanian border- has more power in municipalities with less Albanian population in 1991 but with more labor-force participation in 1991, especially in the agricultural sector. These characteristics coincide with the north-eastern region of the province which, before the war, was characterized for being more prosperous economically and also for having a population with more ethnic diversity.

3.4.3 Isolating Plausibly Exogenous Variation

In order to argue the exogeneity of the instrument, the exclusion restriction requires that the instrument has no correlation with other factors directly affecting labor market and education outcomes other than through its impact on displacement. In other words, the instrument needs to resemble as close as possible a random assignment across municipalities. The main concern that might threaten the exogeneity assumption is that pre-war local economic conditions might predict local conflict intensity and distance to the Albanian border. For instance, locations closer to the Albanian border were more likely to have a higher proportion of Albanians before the war (see Figure 3.A.8). This is of interest since ethnicity might have determined individual economic status through ethnic discrimination. While the exclusion restriction relies on the instrument being uncorrelated with unobserved determinants of the outcomes and hence is untestable, I shed light on its plausibility by providing some historical and empirical evidence.

Historical evidence on the 1998-1999 Kosovo war suggests that targeting of individuals (casualties) was not determined by the economic performance at the local level, as the primary aim of Serb attacks was an ethnically homogeneous and contiguous Serb territory (Ball et al., 2002; OSCE, 1999). Iacopino et al. (2001) study the patterns of forced displacement and human rights abuses using a household survey of 1180 ethnic Albanians living in 31 refugee camps

in Macedonia and Albania during the war. They find that the majority (68%) of participants reported that their families were directly expelled from their homes by Serb Forces. In addition, a report from the Organization for Security and Co-operation in Europe on patterns of human rights and humanitarian law violations in Kosovo confirms this idea (OSCE, 1999):

“After the start of the NATO bombing on the FRY on 24 March, Serbian police and/or VJ (Yugoslav Army), often accompanied by paramilitaries, went from village to village and, in the towns, from area to area threatening and expelling the Kosovo Albanian population. Others who were not directly forcibly expelled fled as a result of the climate of terror created by the systematic beatings, harassment, arrests, killings, shelling and looting carried out across the province. Kosovo Albanians were clearly targeted for expulsion because of their ethnicity. [...] Large numbers of civilians were also deliberately targeted and killed because of their ethnicity. No-one, it seems, was immune, as people of all ages, including women and children, were killed in large numbers.”

Similarly, the bombing attacks were not based on local economic disparities between regions, as NATO’s objective was to attack strongly Serbian targets of military nature (i.e. Serbian air defence sites, communication relays, military facilities and police force headquarters, ammunition dumps and supply routes, such as roads, bridges etc) in order to limit the ethnic cleansing (Grant, 1999; ICTY, 2000). Also, in the final report by the International Criminal Tribunal for former Yugoslavia (ICTY) on NATO’s bombing campaign, it is stated that in several occasions the bombing airstrikes resulted in collateral damage, where locations were mistakenly hit due to failures in target precision. This claim gives certain randomness to the bombing intensity measure.

3.4.4 Robustness Checks

One way to check whether the proposed instrument is as good as random across municipalities/villages is to examine whether individuals differ in pre-war economic performance by the severity of the conflict and the distance to the Albanian border. The idea is that if there is no correlation between the instrument and pre-war baseline characteristics, then there should be no systematic differences in pre-war demographic and economic characteristics across the municipalities/villages in Kosovo. In other words, in the absence of differences in conflict intensity and distance to the Albanian border, municipalities that suffered more from the war and were located further away from the Albanian border would not have been different on average from the rest of the municipalities in Kosovo.

In order to assess the validity of the IV estimates, I undertake three falsification tests on the first-stage to check if the instrument (conflict intensity * distance to the Albanian border)

captures the effect of economic differences across municipalities on conflict displacement. First, I test whether the instrument can predict pre-war migration patterns, which were most likely driven by economic motives. Second, I examine whether the interaction of conflict intensity and distance to the Albanian border is correlated to labor-force participation in 1991 and also to different measures of local economic activity in 1991. Lastly, I also test whether pre-war ethnicity explains any variation in conflict intensity and distance to the Albanian border.

As a first check, I use the municipality of birth and pre-war municipality for all individuals in order to test whether the interaction term can predict pre-1999 migration patterns. Due to lack of pre-war migration data, I consider that all individuals who lived in a different municipality at birth and just before the war are pre-war migrants.²² In this analysis, the municipality of origin is the municipality of birth and the municipality of destination is the pre-war municipality. The control group is formed, in this case, by those individuals that had never migrated before the war, regardless of their displacement status.

The results of this falsification test are reported in Table 3.5 and the specifications are identical to those reported in Table 3.4. The effect of the instrument (conflict intensity * distance to the Albanian border) on pre-war migration is close to zero and highly insignificant in all regressions and in both databases. Overall, these results suggest that conflict intensity in the municipality of birth interacted with distance to the Albanian border does not predict pre-war migrations, which were more likely to be driven by economic reasons. This falsification exercise sheds more light on the idea that conflict intensity was not motivated by the local pre-war economic performance of the municipalities.

As a second check, I use labor force participation (LFP) in 1991 and different measures of the local economic activity 1991 as proxies for pre-war economic performance. Labor force participation in 1991 is constructed by exploiting the Labor Module of the 2000 Kosovo LSMS, which asks individuals whether they were working in 1991 and in which type of activity they were involved (e.g. professional, administrative, clerical, services and agricultural). For this measure, I use only individuals whose residence at birth is the same as their residence previous to the war (i.e. those that have never migrated from their municipality- 85% of the individuals in the 2000 Kosovo LSMS) in order to avoid any measurement error due to migration.

To shed light on the plausibility of the identification assumption, Tables 3.6 and 3.7 regress a

²²Pre-war migrants represent 28.4% of the whole sample of adults aged 20-65 in the 1999 Kosovo DSHS, and 17.1% of the entire sample of adults in the 2000 Kosovo LSMS.

variety of baseline characteristics for economic performance in 1991 on the interaction of conflict intensity and distance to the Albanian border for females and males, respectively. The sample sizes are slightly smaller compared the main specification since this variable is measured only for individuals older than 20 years old in 1991. The dependent variable in column 1 is a dummy indicating whether the individual was working in 1991. The dependent variables in columns 2 to 6 are also dummies indicating whether the individual had a professional occupation, an administrative occupation, a clerical occupation, a service occupation or an agricultural occupation.

On the one hand, the correlation between labor-force participation in 1991 and the interaction of war casualty rate with distance to the Albanian border is negative and statistically significant for both women and men, indicating that before the war economic prosperity was lower in municipalities that were located further away from Albania and that suffered more from war casualties. This result seems to be driven mostly by agricultural occupations, which accounted for 60% of employment in 1991. Similarly, female labor-force participation before the war is also lower in municipalities that received more bombings and were also located further away from Albania. On the other hand, the correlation between male labor-force participation in 1991 and the interaction of bombings with distance to the Albanian border is also negative but statistically insignificant. Overall, these results seem to suggest that the war in Kosovo was more intense in the less prosperous regions of the province.

As a third check, I test whether pre-war ethnicity is correlated to the interaction of conflict intensity and distance to the Albanian border. Tables 3.8 and 3.9 regress a dummy for being Albanian, Serbian or other ethnic group in 1991 on the interaction of conflict intensity and distance to the Albanian border for females and males, respectively. The results from these tables indicate that municipalities with higher conflict intensity and located further away from Albania were more likely to have a higher proportion of Albanians before the war. In other words, municipalities with a higher proportion of Serbs before the war were less likely to suffer from war casualties and bombings. These findings corroborate the historical evidence mentioned earlier in this Section.

Even though historical references suggest that in Kosovo there was no targeting of individuals and regions based on the local economic differences, as a result of these tests, I will report IV estimates after controlling for labor-force participation in 1991 and proportion of Albanians

in 1991 at the municipality level. Even though, I control for these additional variables it is important to acknowledge that the exclusion restriction might still be violated on a number of other dimensions.²³

3.5 Results

The results presented in this section are divided between the initial impacts of conflict displacement in 1999 when individuals had just returned to their homes (i.e. short-run impacts) and the post-displacement impacts in 2000 after individuals had already returned home for approximately one year (i.e. medium-run impacts). All the results are estimated separately for female and male due to substantial gender differences in education and the labor market in Kosovo. The instrumental variables approach will estimate the impact of displacement on various outcomes for those individuals that were induced by the conflict and their residence to be forcibly displaced from their homes. The first sub-section presents the impact of displacement on labor market outcomes for women and men aged 20 to 65 years old. While, the second sub-section presents the impact of displacement on schooling enrolment rates for boys and girls aged 6 to 19 years old.

3.5.1 Conflict Displacement and Labor Market Outcomes

Labor market outcomes are measured as dummy variables indicating whether an adult aged 20 to 65 years old is *employed*, *unemployed* or *inactive*. For cases when the individual declared being employed, I also measure employment with dummy variables indicating: work-off-farm, work-on-farm, work for somebody else, work for family and self-employed. Only for the 2000 Kosovo LSMS, weekly hours are used as an additional labor market outcome.

Table 3.10 and Table 3.11 present the OLS and IV results of the effect of displacement on female's and male's labor market outcomes, respectively. The OLS estimates in Table 3.10 indicate that the effect of displacement on employment for Kosovar women is negative and statistically significant in 1999, implying a fall in the probability to work by 1.8 percentage

²³Unfortunately, I cannot perform similar tests for the school enrollment outcomes due to data availability. Basically, it is practically impossible to have pre-war data on education for children who were in primary or even in secondary school just after the 1998/99 Kosovo war. In spite of this, I do not find any reason hard to believe that educational outcomes, such as enrollment rates or schooling performance, might have had any implication on conflict intensity in Kosovo.

points relative to stayers. This result seems to be driven by those women who are employed by a non-family member. The effect of displacement on female employment is still negative in 2000 but not statistically significant. Displacement does not seem to have an effect on female unemployment nor inactivity.

The OLS estimates in Table 3.11 indicate that displacement increases Kosovar men's inactivity by 2.4 percentage points in 1999. Additionally, displacement is associated with a negative effect on employment and a positive effect on unemployment in both years, but these effects are not statistically significant. Although these OLS estimates seem to imply adverse consequences on labor market outcomes for women and men, they should be taken with caution because they could well be biased.

Therefore, I turn next on second-stage estimates that rely on the interaction between conflict intensity and distance to the Albanian border as exogenous variation in displacement status. After correcting for the potential selection bias, I find that displacement increases the likelihood of a Kosovar woman to be inactive in 1999 by 24 percentage points. However, I also find that in medium-term Kosovar women are on average 7.5 percentage points more likely to be working off-farm compared to stayers (Table 3.10). In addition, the IV estimates for women do not indicate an effect on female unemployment and hours worked just after the conflict and neither one year later.

The IV estimates on the effect of displacement on Kosovar men's labor market outcomes (Table 3.11), which use as instrument ($WCR * DA$) seem to be the most robust across years. Displacement is associated to a large fall in Kosovar Albanian men's ability to be self-employed just after the war. More specifically, displaced Kosovar men are 17 percentage points less likely to work on their own account compared to those that did not move due to the war. There is also suggestive evidence of a negative impact on general employment just after the conflict and also one year later. The magnitude of the effect on employment in both years indicates that the negative effect is decreasing overtime. This negative effect of displacement is quite large and seems to be driven mostly by men who work in the agricultural sector. All these results are robust to using wild bootstrap standard errors.²⁴ In particular, displacement decreases the

²⁴Tables 3.A.3, 3.A.4 and 3.A.5 in the Appendix are a copy of Tables 3.12, 3.10 and 3.11, respectively, but report Quasi-F test statistics and the P-values computed using the wild bootstrap standard errors proposed by [Cameron, Gelbach and Miller \(2008\)](#) and [Cameron and Miller \(2015\)](#). In general, the statistical inference is not affected by the few clusters issue.

likelihood of Kosovar men to be working on-farm by 14 percentage points in 2000.²⁵ Using $(WCR * DA)$ as instrument, I find no effect on unemployment and inactivity.

However, the IV estimates that use $(B * DA)$ as instrument suggest that Kosovar displaced men are also more likely to work off-farm one year after the conflict. In particular, the effect of displacement on men's work-off-farm is large and positive, associated to an increase of 23 percentage points. When analyzing this result by type of occupation, I find that the positive effect is mostly driven by Kosovar men working in the construction and public administration sectors.²⁶ The IV estimates for Kosovar men in 2000 also imply a statistically significant increase in inactivity by 22 percentage points, with no effect on unemployment and hours worked.

3.5.2 Conflict Displacement and School Enrollment Outcomes

Armed conflict is generally expected to adversely affect school enrollment and educational attainment. Basically, the ability of children to attend school may be negatively affected by direct youth enrollment in the military, limited mobility or school destruction among other reasons. In particular, recent research suggests that exposure to civil conflict has adverse effects on the enrollment and completion of schooling (e.g. [Swee, 2015](#); [Chamarbagwala and Morán, 2011](#); [Shemyakina, 2011](#); [Leon, 2012](#); [Akresh and De Walque, 2008](#); [Akbulut-Yuksel, 2014](#); [Merrouche, 2011](#); [Valente, 2013](#)). Moreover, the schooling of girls is often affected more to worsening economic conditions than that of boys. However, the expected result that school enrollment is disrupted in conflict areas may not be well founded in the particular case of Kosovo especially due to the presence of post-war aid through refugee camps.

In this section, I analyze the impact of conflict displacement on post-war school enrollment outcomes.²⁷ Table 3.12 presents the OLS and IV estimates of the impact of displacement on enrolment rates. *Enrolment* is measured as a dummy variable that takes the value of one if

²⁵See Table 3.A.6 in Appendix for an estimation of the effect of displacement on employment by type of occupation. The IV estimates by type of occupation indicate that the negative effect on employment is driven mostly by Kosovar men employed in the agricultural sector.

²⁶See Table 3.A.7 in Appendix for an estimation of the effect of displacement on work-off-farm by type of occupation.

²⁷Two alternative identification strategies are used in this paper to increase confidence in the reliability of the education estimates: the first relies on the interaction-based instrument (conflict intensity \times distance to the Albanian border) as an exogenous source of variation in the displacement decision, and the second relies on a difference-in-difference estimation which uses variation in conflict displacement exposure across birth cohorts and geographic areas (municipalities). The difference-in-difference estimation can be found in the Section 3.B of the Appendix.

the child is enrolled in school and zero otherwise. The OLS estimates suggest that there is no effect of conflict displacement on children's enrolment neither in 1999 nor in 2000.

After correcting for the potential selection bias in the displacement decision, the IV estimates indicate that the effect of displacement on female enrolment in 1999 is positive, quite large and statistically significant. More specifically, displacement increases enrolment of Kosovar girls in 1999 by 18 percentage points, on average. This positive effect is mostly driven by young girls enrolled in primary school (although this effect is lower in magnitude -12.9 percentage points-), as the effect of displacement on secondary school girls is not statistically significant. However, the positive effect on female enrolment seems to disappear one year later, as none of the IV estimates is statistically significant in 2000, both for primary and secondary school girls. Also, I find no effect of displacement on enrolment for Kosovar boys, in general.

Overall, after controlling for endogeneity, young Kosovar girl's enrollment rates respond stronger to forced displacement and high-conflict activity than boys during the post-war period. In Section 3.6, I examine some plausible channels through which household's schooling decisions may have been influenced as a consequence of the forced displacement.

3.6 Mechanisms

3.6.1 Channels on Labor Outcomes

First, focusing on labor market outcomes, the regression analysis implies that displacement is associated to a significant and large decrease in men's employment in the agricultural sector and men's self-employment (which is in general also related to work in the farm). Women are also more likely to drop out of the labor-force. However, the results also indicate that, one year after the end of the war, displaced Kosovar men are also more likely to be working off-farm (i.e. construction and public administration sectors). There are two plausible channels behind these results: first, loss of assets, land and livestock in an agrarian skill-based economy and second, loss of social networks in an informal labor market.

In the immediate aftermath of the conflict, the population of Kosovo faced a complex situation where their livelihoods were radically altered: infrastructure and housing were damaged or destroyed; crops had failed and large amounts of agricultural land were left under-utilised or abandoned ([Douarin, Litchfield and Sabates-Wheeler, 2012](#)). [Westley and](#)

[Mikhalev \(2002\)](#) describe how the war and conflict displacement created constraints across the economy where many households were unable to cultivate land on a commercial basis due to loss of equipment and livestock, damage to irrigation systems as well as limited access to their land due to security reasons, including the presence of landmines and cluster bombs. Therefore, displacement might have made very difficult post-war employment in the agriculture sector for those individuals that had to abandon their lands and livestock.²⁸ However, despite the lack of work opportunities in the agriculture sector, displaced Albanian Kosovar men might have turned to wage labour as a post-displacement measure, especially in the construction and public administration sectors as the demand for labor in these two sectors increased significantly after the war. Indeed, [Douarin, Litchfield and Sabates-Wheeler \(2012\)](#) find that one of the most successful post-war livelihood strategies of Kosovar Albanians was associated with access to non-farm income sources such as non-farm businesses and remittances.

In order to test this plausible channel, Panel A of Table 3.13 presents the effect of household displacement on assets, measured as land ownership, land size, livestock and number of livestock. First, in line with previous studies I find suggestive evidence that returnees have fewer assets after the end of the war than those who stayed during the conflict, especially in terms of both land and livestock ownership. Moreover, using the 2000 Kosovo LSMS database, I find that in the medium-term conflict displacement has a negative and statistically significant effect on the number of livestock, which reinforces the first proposed mechanism behind the labor market outcomes.

Moreover, another plausible mechanism behind these results might be the loss of informal networks, such as separation from family members, relatives, friends and communities ([Kondylis, 2010](#)). Several studies in the literature on migration suggests that networks are a key entry point to informal labor markets in an informal economy. For instance, [Edin, Fredriksson and Åslund \(2003\)](#) finds that living in an enclave enhances the access to informal ethnic networks and improves immigrants access to employment by increasing the performance of refugee immigrants job-search. In addition, this channel may be linked to the literature that studies the role of social networks as adverse coping mechanisms in the management of violent shocks. Most of the conflicts take place in poor countries, where -in the absence of formal

²⁸[Bozzoli, Brueck and Muhumuza \(2016\)](#) and [Deininger \(2003\)](#) have also found that the probability to start non-farm activities is substantially reduced for households affected by war using data from the 20-year civil conflict in Northern Uganda.

insurance mechanisms- social networks provide support such as informal loans and transfers to mitigate various negative shocks (Foster and Rosenzweig, 2001; Fafchamps and Lund, 2003). Therefore, conflict displacement might have decreased access to informal networks for Kosovar Albanians since not everybody might have returned to the same pre-war residence. Also, taking into account the informal nature of the agriculture sector in Kosovo, the poorer access to informal networks might have further decreased the likelihood of displaced Kosovar men to find employment relative to stayers.

In order to test this channel, I exploit the Networks Module of the 2000 Kosovo LSMS to define access to informal networks. This section contains information on who would the individuals turn to in case of economic loss (i.e bad harvest, loss of employment), with the following categories: humanitarian group, relatives, neighbours, friends, community leaders, religious leaders, others etc. Using Pistaferri (1999), I define informal networks when the individual seeks employment through relatives, neighbours or friends. Panel B in Table 3.13 shows the IV estimates of the effect of household displacement on informal social networks. I find that displaced households are less likely to have access to informal networks compared to stayers. This channel might be closely linked with the increase in women's inactivity.

3.6.2 Channels on Education Outcomes

Second, focusing on education outcomes, the results found in this paper indicate that displacement in Kosovo had positive short-run effects on female's school enrolment, especially for those in primary level. One possible channel through which this effect might be operating is the refugee camp experience. It is interesting to note that the likelihood of children accessing education as refugees could either increase or decrease depending on the context. For instance, in conflict-affected countries, where virtually many children are out of school, refugee children, especially if they reside in refugee camps, are much more likely to increase their access to education compared to those who still stay in the the conflict-affected areas. However, for children leaving countries with fairly good access to schooling, it is likely that their ability to access education will decrease as a refugee (Ferris and Winthrop, 2010).

Between 1991 and the late 1990s the Albanian Kosovar population received education services in an informal system parallel to the official one. As schools and faculties in Albanian language were closed, most Kosovar Albanian students received classes outside school

facilities and often in private homes. During this period, the availability of educational inputs declined significantly, and teachers were unable to update their teaching skills and methodologies (Alva, Murrugarra and Paci, 2002; Cutts, 2000). Given this precarious pre-war situation, being displaced in a refugee centre might have increased access to education for Kosovar Albanian girls. Young female refugees, especially those who were in camps, might have had better access to basic education and better conditions than the IDPs and the stayer girls. The 1999 UNHCR Global Report seems to confirm this idea:

“The Ministry of Education in Albania and Macedonia organised summer schools for refugee children to make up for the schooling lost in the winter and spring 1998/99. UNHCR and UNICEF assisted by contributing to the cost of printing school books for 150,000 refugee children of primary school age. Many also received new furniture and supplies”

In order to test this, Table 3.14 presents the effect of displacement in a refugee camp on enrollment outcomes in 1999 for Kosovar females using the instrumental variables technique. The IV estimates indicate that displaced Kosovar girls residing in a refugee camp are more likely to be enrolled in school after returning in Kosovo compared to those that stayed in Kosovo and also to those that were internally displaced or residing in host families. This effect is driven mostly by girls enrolled in primary level, as the effect of being displaced in a refugee camp for girls enrolled in secondary school is not statistically significant. These results suggest that since primary schooling is considered to have higher priority in refugee centres compared to secondary schooling, the refugee camp’s conditions might have been more beneficial for younger girls compared to older ones (teenagers).

3.7 Conclusion

This paper contributes to the literature on the impacts of conflict displacement in developing countries. More specifically, this study analyzes the impact of forced displacement on children’s schooling and adult’s labor market outcomes in the context of the post-war Kosovo. During the Kosovo war and especially during the NATO air campaign, more than a million of individuals of all ethnicities were displaced, which represented around 70% of Kosovo’s pre-war population.

Using a combination of household survey data and municipality level data on conflict intensity, I exploit the interaction between spatial variation in conflict intensity and distance to the Albanian border as a source of exogenous variation in the displacement decision. As the

targeting of individuals and regions in Kosovo was not based on pre-war economic differences, it is possible to argue that the severity of the conflict, measured through war casualties and NATO bombing days per municipality, is not related to unobserved characteristics that may also affect post-war economic outcomes.

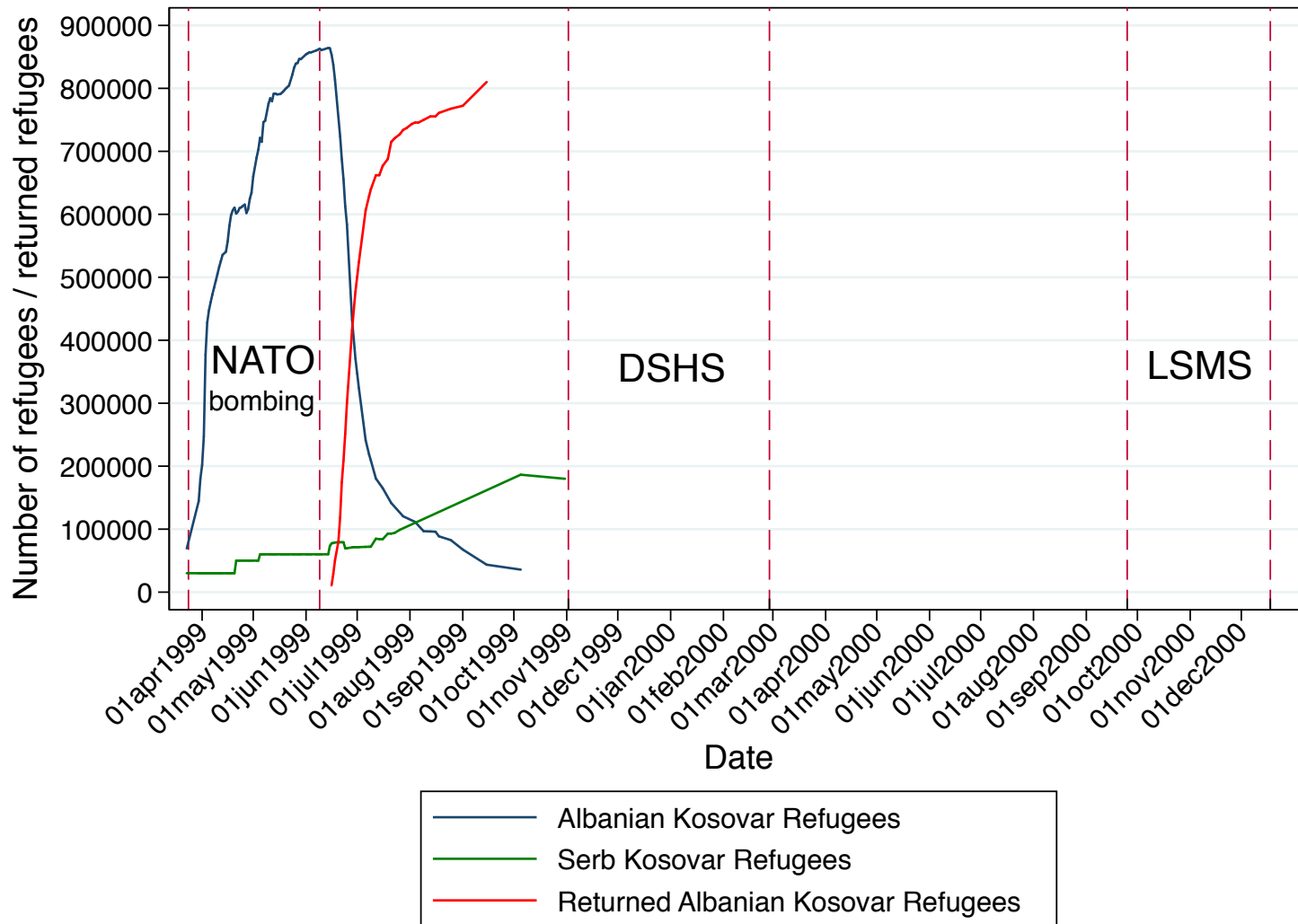
The regression analysis implies some positive but also negative impacts of displacement on labor market and education outcomes. In particular, in terms of education outcomes, the results found in this paper indicate that displacement in Kosovo had positive short-run effects on female's school enrolment, especially for those in primary level. However, there is no evidence of changes in school enrollment for Kosovar displaced boys. One possible channel behind these results could be the experience of refugee camps, in the sense that conditions in the refugee camps might have provided better conditions and access to education to young Kosovar girls compared to the pre-war access which was characterized by the "parallel" education system.

In addition, in terms of labor market outcomes, the regression analysis implies that displacement is associated to a significant and large decrease in men's employment in the agricultural sector and their capacity to work on their own account. I also find that displaced Kosovar women are more likely to drop out of the labor force. In addition, households that were displaced have significantly fewer assets, land and livestock ownership in an agrarian skill-based economy and also experienced loss of social networks in an informal labor market compared to not displaced households. However, shortly after the return home, the results also indicate that displaced Kosovar men and women are more likely to be working off-farm, especially in the construction and public administration sectors, which indicates a relatively quick recovery.

It is clear that by 2000 Kosovar displaced people were unable to completely recover from the conflict. Even though I find some suggestive evidence of a post-conflict reconstruction effort, the results found in this paper imply that there is still a role for the international community and the local government to develop and support these livelihood activities in a post-conflict context through early interventions.

Chapter 3 - Figures and Tables

Figure 3.1: Total Cumulative Kosovar Refugee, Returned Kosovar Refugee, Serbian Refugee Populations and Survey Timelines, 23 (March-October 1999)



Source: UNHCR

Figure 3.2: War Casualty Rate- Number of Casualties per 1,000 Inhabitants across Municipalities

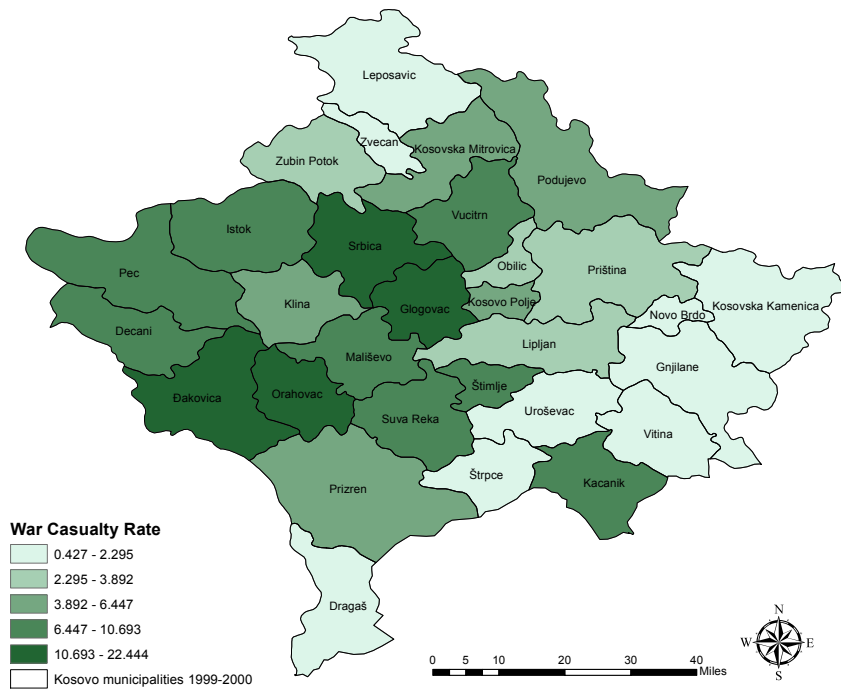
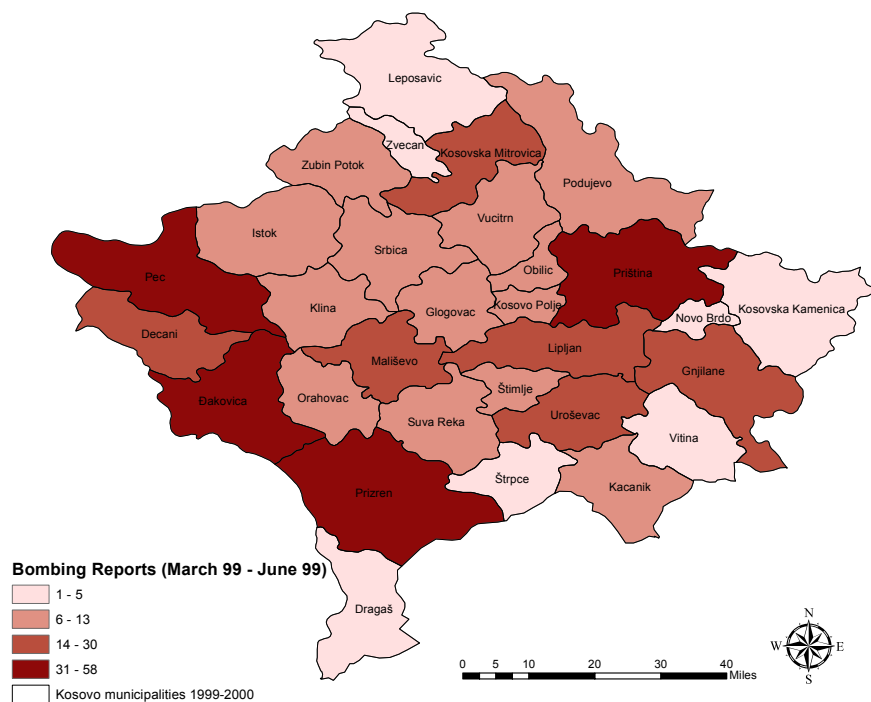


Figure 3.3: Bombing Intensity- NATO bombing Days across Municipalities



Notes: The war casualty rate is computed as a proportion of the 1991 population at the municipality level.
Sources: Kosovo Memory Book (1998-2000) and Human Rights Group for Data Analysis (HRDAG)

Table 3.1: *Descriptive Statistics of Displacement by Gender and Age Group - Kosovo (1999-2000)*

Samples	1999 Kosovo DSHS		2000 Kosovo LSMS	
	Female	Male	Female	Male
Children [6 - 19 years old]				
Displaced	0.652	0.654	0.734	0.742
Move out of Kosovo	0.412	0.382	0.447	0.418
IDPs	0.223	0.224	0.279	0.282
Refugee center	0.214	0.220	—	—
Returnees*	0.851	0.854	0.961	0.951
Observations	5,919	6,385	2,616	2,716
Adults [20 - 65 years old]				
Displaced	0.652	0.654	0.734	0.742
Move out of Kosovo	0.412	0.382	0.447	0.417
IDPs	0.192	0.192	0.240	0.236
Refugee center	0.193	0.181	—	—
Returnees*	0.872	0.871	0.965	0.962
Observations	10,751	9,554	4,639	4,317

Notes: *The returnees are calculated only for the displaced individuals, therefore the number of observations is lower.

Table 3.2: *Descriptive Statistics of Displacement Status by Ethnic Group - Kosovo (1999-2000)*

Ethnicity	1999 Kosovo DSHS			2000 Kosovo LSMS		
	All %	Displaced %	Not Displaced %	All %	Displaced %	Not Displaced %
Albanian	86.14	97.66	67.97	84.83	96.33	58.95
Croat	0.03	0.03	0.03	0.16	0.09	0.31
Bosniak	2.89	1.66	4.83	1.79	0.79	4.02
Roma	1.07	0.14	2.54	1.36	0.93	2.33
Serb	8.69	0.20	22.09	10.75	1.65	31.21
Turk	0.93	0.28	1.95	1.06	0.15	3.10
Montenegrians	—	—	—	0.04	0.04	0.04
Others	0.25	0.03	0.59	0.03	0.03	0.04
Total (Observations)	40,757	24,952	15,805	16,749	11,594	5,155

Table 3.3: *First-stage Results - Children Aged (6-19 Years Old) - Kosovo (1999-2000)*

	1999 Kosovo DSHS				2000 Kosovo LSMS			
Children [6-19] Years Old	Dependent variable: Displaced							
	Female		Male		Female		Male	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Distance to Alb. Border	-0.0035 (0.0024) [0.0023]	0.0052 (0.0021)** [0.0025]*	-0.0033 (0.0025) [0.0021]	0.0046 (0.0020)** [0.0024]*	-0.0039 (0.0016)** [0.0024]	-0.0022 (0.0013)* [0.0019]	-0.0036 (0.0017)** [0.0024]	-0.0016 (0.0013) [0.0020]
Bombings		0.0153 (0.0294) [0.0251]		0.0132 (0.0323) [0.0246]		-0.0543 (0.0159)*** [0.0208]**		-0.0525 (0.0160)*** [0.0218]**
War Casualty Rate	-0.0670 (0.0221)*** [0.0257]**		-0.0622 (0.0217)*** [0.0242]**		-0.0547 (0.0171)*** [0.0218]**		-0.0511 (0.0192)*** [0.0207]**	
Excluded Instruments:								
Bombings x Dist. Alb.		-0.0002 (0.0003) [0.0002]		-0.0002 (0.0003) [0.0002]		0.0005 (0.0002)*** [0.0002]***		0.0005 (0.0002)*** [0.0002]**
WCR x Dist. Alb.	0.0008 (0.0002)*** [0.0003]**		0.0007 (0.0002)*** [0.0003]**		0.0007 (0.0002)*** [0.0003]**		0.0007 (0.0002)*** [0.0003]**	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,919	5,919	6,385	6,385	2,616	2,616	2,716	2,716
Mean dependent var.	0.651	0.651	0.653	0.653	0.734	0.734	0.741	0.741
F-stat Excluded Instruments	11.64/7.67	0.51/0.70	10.84/7.45	0.27/0.48	13.38/6.50	12.59/7.67	10.18/6.55	10.15/5.18
Number of clusters	55/27	55/27	55/27	55/27	200/29	200/29	203/29	203/29
Partial R-Squared	0.067	0.004	0.060	0.002	0.041	0.040	0.044	0.033

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Controls include age, ethnicity (albanian), dummies for mothers and fathers educational attainment (medium and high), number of male and female adults in a household aged 20–65, number of children aged 0, 1-5, 6-10, 11-15, and 16-19, distance to school, dummy for rural location, municipality labor-force participation in 1991 and proportion of Albanians in 1991. LSMS sample weights are used in all the regressions. Data source: 1999 Kosovo DSHS and 2000 Kosovo LSMS

Table 3.4: First-stage Results - Adults (20-65 Years Old) - Kosovo (1999-2000)

	1999 Kosovo DSHS				2000 Kosovo LSMS			
Adults [20-65] Years Old	Dependent variable: Displaced							
	Female		Male		Female		Male	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Distance to Alb. Border	-0.0021 (0.0022) [0.0020]	0.0038 (0.0019)** [0.0023]	-0.0018 (0.0021) [0.0019]	0.0035 (0.0018)* [0.0022]	-0.0028 (0.0016)* [0.0026]	-0.0017 (0.0012) [0.0021]	-0.0037 (0.0016)** [0.0025]	-0.0014 (0.0011) [0.0021]
Bombings		0.0027 (0.0274) [0.0249]		0.0005 (0.0257) [0.0239]		-0.0527 (0.0142)*** [0.0219]**		-0.0505 (0.0148)*** [0.0214]**
War Casualty Rate	-0.0530 (0.0197)*** [0.0218]**		-0.0493 (0.0188)** [0.0201]**		-0.0439 (0.0153)*** [0.0218]*		-0.0588 (0.0141)*** [0.0220]**	
Excluded Instruments:								
Bombings x Dist. Alb.		-0.0000 (0.0002) [0.0002]		-0.0000 (0.0002) [0.0002]		0.0005 (0.0001)*** [0.0002]**		0.0005 (0.0002)*** [0.0002]**
WCR * Dist. Alb.	0.0007 (0.0002)*** [0.0003]**		0.0006 (0.0002)*** [0.0002]**		0.0006 (0.0002)*** [0.0003]*		0.0008 (0.0002)*** [0.0003]**	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,751	10,751	9,454	9,454	4,639	4,639	4,317	4,317
Mean dependent var.	0.604	0.604	0.575	0.575	0.687	0.687	0.653	0.653
F-stat Excluded Instruments	9.14/6.40	0.01/0.01	7.39/5.78	0.00/0.00	10.88/3.98	11.89/5.08	18.46/6.15	9.15/4.71
Number of clusters	55/27	55/27	55/27	55/27	207/29	207/29	206/29	206/29
Partial R-squared	0.051	0.001	0.040	0.000	0.034	0.035	0.043	0.030

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Controls include age, ethnicity (albanian), dummies for marital status, dummies for parental educational attainment (low and medium), number of male and female adults in a household aged 20-65, number of dependent members aged 0, 1-5, 6-10, 11-15, and 16-19, dummy for rural location, municipality labor-force participation in 1991 and proportion of Albanians in 1991. LSMS sample weights are used in all the regressions. Data source: 1999 Kosovo DSHS and 2000 Kosovo LSMS.

Table 3.5: *Falsification Test on the First-stage Results - Adults Aged (20-65 Years Old) - Kosovo (1999-2000)*

	1999 Kosovo DSHS				2000 Kosovo LSMS			
Adults [20-65] Years Old	Dependent variable: Migration Status (Non-conflict)							
	Female		Male		Female		Male	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Distance to Alb. Border	0.0003 (0.0019) [0.0028]	-0.0018 (0.0018) [0.0028]	-0.0001 (0.0013) [0.0018]	-0.0011 (0.0011) [0.0017]	-0.0001 (0.0009) [0.0020]	-0.0000 (0.0011) [0.0021]	-0.0002 (0.0004) [0.0006]	-0.0003 (0.0006) [0.0009]
Bombings in the municip. of birth		-0.0346 (0.0203)* [0.0297]		-0.0287 (0.0141)** [0.0172]		-0.0091 (0.0125) [0.0230]		-0.0072 (0.0099) [0.0116]
WCR in the municip. of birth	-0.0033 (0.0225) [0.0314]		-0.0088 (0.0152) [0.0208]		-0.0092 (0.0109) [0.0223]		-0.0042 (0.0077) [0.0101]	
Excluded Instruments:								
Bombings x Distance to Alb. Border		0.0003 (0.0002) [0.0003]		0.0003 (0.0002) [0.0002]		0.0001 (0.0002) [0.0003]		0.00004 (0.0001) [0.0001]
WCR x Distance to Alb. Border	0.0001 (0.0002) [0.0004]		0.0002 (0.0002) [0.0003]		0.0002 (0.0001) [0.0003]		0.0001 (0.0001) [0.0001]	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,327	10,327	9,200	9,200	4,481	4,481	4,251	4,251
Mean dependent var.	0.352	0.352	0.207	0.207	0.240	0.240	0.098	0.098
F-stat Excluded Instruments	0.19/0.08	2.82	1.37/0.50	3.67/2.35	1.55/0.37	0.23/0.10	0.55/0.44	0.10/0.11
Number of clusters	55/27	55/27	55/27	55/27	201/29	201/29	201/29	201/29
Partial R-squared	0.001	0.020	0.004	0.015	0.002	0.001	0.001	0.0001

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. The sample includes women aged 20 to 65 in 1991. Controls include age, dummies for marital status, dummies for educational attainment (low and medium), and dummy for rural residence. LSMS sample weights are used in all the regressions. Data source: 2000 Kosovo LSMS.

Table 3.6: *Pre-war Economic Performance and Conflict Incidence (Women) - Kosovo (1999-2000)*

Sample: Female [20-65 y.o. in 1991]	Pre-war economic performance 1991					
	LFP (1)	professional (2)	administrative (3)	clerical (4)	services (5)	agricultural (6)
WCR x Dist. Alb.	-0.0003 (0.0001)** [0.0001]**	-0.00001 (0.00001) [0.00002]	-0.000005 (0.000004) [0.000004]	-0.00002 (0.00002) [0.00003]	0.0001 (0.00005)** [0.00009]	-0.0004 (0.0001)** [0.0001]**
WCR	0.0211 (0.0120)* [0.0141]	0.0018 (0.0015) [0.0016]	0.0004 (0.0004) [0.0005]	0.0021 (0.0020) [0.0025]	-0.0067 (0.0039)* [0.0064]	0.0235 (0.0107)** [0.0132]*
Dist. Alb.	0.0050 (0.0010)** [0.0009]**	0.0001 (0.0001) [0.0001]	0.00001 (0.00002) [0.00003]	0.0003 (0.0002)** [0.0002]	0.0009 (0.0004)** [0.0007]	0.0037 (0.0008)** [0.0008]**
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,210	3,210	3,210	3,210	3,210	3,210
Mean dep. var	0.357	0.055	0.002	0.038	0.046	0.213
Number of clusters	206/29	206/29	206/29	206/29	206/29	206/29
R-squared	0.1572	0.2672	0.0196	0.1190	0.0524	0.1358
Bombings x Dist. Alb.	-0.0002 (0.0001)** [0.0001]**	-0.00002 (0.00002) [0.00002]	-0.000002 (0.000003) [0.000004]	-0.000005 (0.00002) [0.00002]	-0.00009 (0.00004)* [0.00009]	-0.0001 (0.0001) [0.0001]
Bombings	0.0188 (0.0103)* [0.0124]	0.0029 (0.0029) [0.0032]	0.0005 (0.0003) [0.0004]	-0.0002 (0.0028) [0.0028]	0.0090 (0.0051)* [0.0093]	0.0065 (0.0094) [0.0159]
Dist. Alb.	0.0054 (0.0007)** [0.0008]**	0.0002 (0.0001) [0.0002]	-0.00001 (0.00002) [0.00002]	0.0002 (0.0001)* [0.0002]	0.0019 (0.0005)** [0.0012]	0.0032 (0.0007)** [0.0012]**
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,210	3,210	3,210	3,210	3,210	3,210
Mean dep. var	0.357	0.055	0.002	0.038	0.046	0.213
Number of clusters	206/29	206/29	206/29	206/29	206/29	206/29
R-squared	0.1625	0.2676	0.0225	0.1192	0.0550	0.1311

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. The sample includes men aged 20 to 65 in 1991. Controls include age, dummies for marital status, dummies for educational attainment (low and medium), and dummy for rural residence. LSMS sample weights are used in all the regressions. Data source: 2000 Kosovo LSMS.

Table 3.7: Pre-war Economic Performance and Conflict Incidence (Men) - Kosovo (1999-2000)

Sample: Male [20-65 y.o. in 1991]	Pre-war economic performance 1991					
	LFP (1)	professional (2)	administrative (3)	clerical (4)	services (5)	agricultural (6)
WCR x Dist. Alb.	-0.0003 (0.0001)*** [0.0001]***	-0.00005 (0.00004) [0.00003]	0.000003 (0.00002) [0.00002]	-0.0000003 (0.00003) [0.00002]	0.00001 (0.00003) [0.00004]	-0.0002 (0.0001)*** [0.0001]***
WCR	0.0152 (0.0049)*** [0.0059]**	0.0046 (0.0034) [0.0030]	-0.0013 (0.0019) [0.0019]	-0.0011 (0.0027) [0.0022]	-0.0009 (0.0036) [0.0038]	0.0140 (0.0059)** [0.0054]**
Dist. Alb.	0.0008 (0.0003)** [0.0004]**	0.0001 (0.0003) [0.0003]	-0.000005 (0.0002) [0.0002]	-0.000003 (0.0003) [0.0002]	0.000004 (0.0003) [0.0004]	0.0007 (0.0004)* [0.0004]*
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,983	2,983	2,983	2,983	2,983	2,983
Mean dep. var	0.829	0.127	0.036	0.093	0.078	0.492
Number of clusters	206/29	206/29	206/29	206/29	206/29	206/29
R-squared	0.1285	0.2054	0.0504	0.0543	0.0127	0.1831
Bombings x Dist. Alb	-0.00005 (0.00004) [0.00006]	0.00001 (0.00003) [0.00003]	0.000007 (0.00001) [0.00001]	-0.000008 (0.00003) [0.00002]	-0.00001 (0.00003) [0.00003]	-0.00004 (0.00005) [0.00005]
Bombings	0.0028 (0.0047) [0.0063]	0.0001 (0.0040) [0.0036]	0.0008 (0.0021) [0.0018]	0.0003 (0.0031) [0.0026]	0.0023 (0.0036) [0.0034]	-0.0007 (0.0058) [0.0066]
Dist. Alb.	0.0001 (0.0004) [0.0006]	-0.0004 (0.0003) [0.0004]	-0.0000 (0.0002) [0.0002]	0.00002 (0.0002) [0.0002]	0.0001 (0.0003) [0.0003]	0.0003 (0.0005) [0.0005]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,983	2,983	2,983	2,983	2,983	2,983
Mean dep. var	0.829	0.127	0.036	0.093	0.078	0.492
Number of clusters	206/29	206/29	206/29	206/29	206/29	206/29
R-squared	0.1156	0.2057	0.0525	0.0545	0.0129	0.1812

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. The sample includes males aged 20 to 65 in 1991. Controls include age, ethnicity (albanian), dummies for marital status, dummies for parental educational attainment (low and medium), number of male and female adults in a household aged 20-65, number of dependent members aged 0, 1-5, 6-10, 11-15, and 16-19 and dummy for rural. LSMS sample weights are used in all the regressions. Data source: 1999 Kosovo DSHS and 2000 Kosovo LSMS.

Table 3.8: *Pre-War Ethnicity and Conflict Incidence (Women) - Kosovo (1999-2000)*

Sample: Female [20-65 y.o. in 1991]	Pre-war ethnicity 1991		
	Albanian (1)	Serbian (2)	Other ethnicity (3)
WCR x Dist. Alb.	0.0002 (0.0002) [0.0003]	-0.0004 (0.0001)*** [0.0002]*	0.0002 (0.0001)*** [0.0001]**
WCR	-0.0009 (0.0127) [0.0227]	0.0248 (0.0096)** [0.0176]	-0.0239 (0.0069)*** [0.0086]**
Dist. Alb.	-0.0011 (0.0013) [0.0024]	0.0042 (0.0010)*** [0.0018]**	-0.0031 (0.0007)*** [0.0010]***
Controls	Yes	Yes	Yes
Observations	3,540	3,540	3,540
Mean dep. var	0.849	0.096	0.053
Number of clusters	206/29	206/29	206/29
R-squared	0.1198	0.2127	0.1046
Bombings x Dist. Alb	0.0004 (0.0001)*** [0.0002]	-0.0002 (0.0001)** [0.0002]	-0.0001 (0.0001) [0.0001]
Bombings	-0.0350 (0.0147)** [0.0218]	0.0239 (0.0111)** [0.0164]	0.0111 (0.0085) [0.0114]
Dist. Alb.	-0.0027 (0.0011)** [0.0022]	0.0034 (0.0009)*** [0.0018]*	-0.0007 (0.0007) [0.0011]
Controls	Yes	Yes	Yes
Observations	3,540	3,540	3,540
Mean dep. var	0.849	0.096	0.053
Number of clusters	206/29	206/29	206/29
R-squared	0.0986	0.1885	0.0791

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. The sample includes women aged 20 to 65 in 1991. Controls include age, dummies for marital status, dummies for educational attainment (low and medium), and dummy for rural residence. LSMS sample weights are used in all the regressions. Data source: 2000 Kosovo LSMS.

Table 3.9: *Pre-War Ethnicity and Conflict Incidence (Men) - Kosovo (1999-2000)*

Sample: Male [20-65 y.o. in 1991]	Pre-war ethnicity 1991		
	Albanian (1)	Serbian (2)	Other ethnicity (3)
WCR x Dist. Alb.	0.0002 (0.0002) [0.0003]	-0.0004 (0.0001)*** [0.0002]*	0.0002 (0.0001)*** [0.0001]*
WCR	-0.0010 (0.0136) [0.0236]	0.0263 (0.0099)*** [0.0180]	-0.0253 (0.0079)*** [0.0105]**
Dist. Alb.	-0.0012 [0.0024] (0.0014)	0.0044 [0.0019]** (0.0010)***	-0.0032 [0.0010]*** (0.0008)***
Controls	Yes	Yes	Yes
Observations	3,293	3,293	3,293
Mean dep. var	0.842	0.104	0.052
Number of clusters	206/29	206/29	206/29
R-squared	0.0980	0.1720	0.1373
Bombings x Dist. Alb	0.0003 (0.0001)** [0.0002]	-0.0002 (0.0001)** [0.0002]	-0.0001 (0.0001) [0.0001]
Bombings	-0.0319 (0.0154)** [0.0219]	0.0260 (0.0120)** [0.0167]	0.0059 (0.0089) [0.0118]
Dist. Alb.	-0.0024 (0.0012)* [0.0023]	0.0034 (0.0009)*** [0.0018]*	-0.0011 (0.0007) [0.0012]
Controls	Yes	Yes	Yes
Observations	3,293	3,293	3,293
Mean dep. var	0.842	0.104	0.052
Number of clusters	206/29	206/29	206/29
R-squared	0.0709	0.1467	0.1045

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. The sample includes men aged 20 to 65 in 1991. Controls include age, dummies for marital status, dummies for educational attainment (low and medium), and dummy for rural residence. LSMS sample weights are used in all the regressions. Data source: 2000 Kosovo LSMS.

Table 3.10: *The Effect of Conflict Displacement on Women's Labor Market Outcomes - Kosovo (1999-2000)*

Female [20-65]	1999 Kosovo DSHS				2000 Kosovo LSMS				
			OLS	IV (WCR* Dist.Alb)			OLS	IV (WCR* Dist.Alb)	IV (Bombs* Dist.Alb)
	Obs.	Mean	(1)	(2)	Obs.	Mean	(3)	(4)	(5)
Employed	10,751	0.150	-0.018 (0.008)** [0.010]*	-0.056 (0.083) [0.096]	4,639	0.290	-0.015 (0.026) [0.030]	0.192 (0.197) [0.163]	0.023 (0.154) [0.098]
Work off-farm	—	—	—	—	4,639	0.078	-0.009 (0.012) [0.016]	-0.029 (0.035) [0.045]	0.075 (0.050) [0.044]*
Work on-farm	—	—	—	—	4,639	0.189	-0.010 (0.024) [0.029]	0.298 (0.215) [0.207]	-0.041 (0.166) [0.108]
Work for someone	10,751	0.135	-0.014 (0.009) [0.009]	-0.083 (0.082) [0.097]	—	—	—	—	—
Work for family	10,751	0.010	-0.005 (0.004) [0.004]	0.028 (0.029) [0.030]	—	—	—	—	—
Self-Employed	10,751	0.005	0.001 (0.002) [0.002]	-0.0004 (0.008) [0.008]	4,639	0.019	0.001 (0.006) [0.005]	-0.058 (0.039) [0.038]	0.003 (0.028) [0.024]
Unemployed	10,751	0.173	-0.014 (0.022) [0.024]	-0.184 (0.125) [0.1113]	4,639	0.070	-0.001 (0.011) [0.011]	-0.018 (0.061) [0.054]	-0.040 (0.058) [0.063]
Inactive	10,751	0.676	0.033 (0.024) [0.025]	0.239 (0.122)** [0.106]**	4,639	0.638	0.016 (0.026) [0.028]	0.174 (0.179) [0.154]	0.017 (0.127) [0.110]
Hours worked (o)	—	—	—	—	4,639	9.57	-0.596 (1.124) [1.217]	7.666 (9.205) [8.877]	10.378 (7.983) [6.580]
Hours worked	—	—	—	—	1,303	34.08	-0.561 (1.787) [1.909]	-7.459 (12.097) [16.816]	12.139 (10.923) [10.331]
Controls			Yes	Yes			Yes	Yes	Yes

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Controls include: conflict intensity (casualties/bomings), distance to the Albanian border, age, ethnicity (albanian), dummies for marital status, dummies for educational attainment (low and medium), number of male and female adults in a household aged 20-65, number of dependent members aged 0, 1-5, 6-10, 11-15, and 16-19, dummy for rural location, municipality labor-force participation in 1991 and proportion of Albanians in 1991. LSMS sample weights are used in all the regressions. Data source: 1999 Kosovo DSHS and 2000 Kosovo LSMS.

Table 3.11: The Effect of Conflict Displacement on Men's Labor Market Outcomes - Kosovo (1999-2000)

Male [20-65]	1999 Kosovo DSHS				2000 Kosovo LSMS				
			OLS	IV			OLS	IV	IV
	Obs.	Mean	(1)	(WCR* Dist.Alb) (2)	Obs.	Mean	(3)	(WCR* Dist.Alb) (4)	(Bombs* Dist.Alb) (5)
Employed	9,454	0.474	-0.016 (0.029) [0.035]	-0.272 (0.176) [0.181]	4,317	0.642	-0.017 (0.021) [0.020]	-0.104 (0.101) [0.123]	-0.121 (0.160) [0.146]
Work off-farm	—	—	—	—	4,317	0.284	-0.032 (0.022) [0.024]	0.064 (0.086) [0.104]	0.234 (0.116)** [0.114]**
Work on-farm	—	—	—	—	4,317	0.248	-0.027 (0.019) [0.021]	-0.140 (0.128) [0.085]*	-0.090 (0.153) [0.174]
Work for someone	9,454	0.365	0.028 (0.025) [0.028]	-0.003 (0.145) [0.171]	—	—	—	—	—
Work for family	9,454	0.025	-0.016 (0.012) [0.014]	-0.095 (0.079) [0.066]	—	—	—	—	—
Self-Employed	9,454	0.084	-0.028 (0.019) [0.019]	-0.174 (0.113) [0.065]**	4,317	0.146	0.017 (0.021) [0.018]	0.030 (0.074) [0.088]	-0.060 (0.125) [0.160]
Unemployed	9,454	0.422	-0.008 (0.025) [0.030]	0.201 (0.172) [0.199]	4,317	0.113	0.013 (0.013) [0.016]	0.086 (0.064) [0.072]	-0.106 (0.127) [0.127]
Inactive	9,454	0.103	0.024 (0.009)** [0.010]**	0.071 (0.059) [0.050]	4,317	0.244	0.004 (0.020) [0.015]	0.018 (0.085) [0.093]	0.227 (0.122)* [0.152]
Hours worked (o)	—	—	—	—	4,317	27.74	-1.104 (1.373) [1.389]	-9.328 (5.711) [5.361]	-8.496 (8.943) [9.934]
Hours worked	—	—	—	—	2,706	44.26	-0.175 (1.269) [1.093]	3.336 (5.431) [4.643]	-8.778 (9.371) [11.659]
Controls			Yes	Yes			Yes	Yes	Yes

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Controls include: conflict intensity (casualties/bomings), distance to the Albanian border, age, ethnicity (albanian), dummies for marital status, dummies for educational attainment (low and medium), number of male and female adults in a household aged 20-65, number of dependent members aged 0, 1-5, 6-10, 11-15, and 16-19, dummy for rural location, municipality labor-force participation in 1991 and proportion of Albanians in 1991. LSMS sample weights are used in all the regressions. Data source: 1999 Kosovo DSHS and 2000 Kosovo LSMS.

Table 3.12: *The Effect of Conflict Displacement on Children's Enrollment - Kosovo (1999-2000)*

Children [6-19] Years Old	1999 Kosovo DSHS				2000 Kosovo LSMS				
	Obs.	Mean	OLS	IV	Obs.	Mean	OLS	IV	IV
			(1)	(WCR* Dist. Alb) (2)			(3)	(WCR* Dist. Alb) (4)	(Bombs* Dist. Alb) (5)
Sample: FEMALE									
Enrolled in school (6-19 years old)	5,919	0.749	0.014 (0.020) [0.023]	0.188 (0.065)*** [0.052]***	2,616	0.820	0.008 (0.027) [0.025]	0.038 (0.116) [0.145]	-0.004 (0.125) [0.124]
Enrolled in school (6-14 years old)	3,709	0.864	0.023 (0.018) [0.019]	0.129 (0.041)*** [0.036]***	1,650	0.920	0.029 (0.024) [0.022]	-0.101 (0.105) [0.106]	0.019 (0.157) [0.155]
Enrolled in school (15-19 years old)	2,210	0.556	-0.008 (0.032) [0.031]	0.118 (0.137) [0.106]	966	0.635	-0.031 (0.045) [0.038]	0.297 (0.275) [0.273]	0.076 (0.195) [0.176]
Sample: MALE									
Enrolled in school (6-19 years old)	6,385	0.804	0.006 (0.011) [0.010]	0.028 (0.044) [0.041]	2,716	0.820	0.008 (0.021) [0.016]	0.039 (0.099) [0.067]	0.076 (0.124) [0.106]
Enrolled in school (6-14 years old)	4,121	0.869	-0.004 (0.010) [0.012]	0.001 (0.037) [0.038]	1,762	0.920	0.011 (0.021) [0.022]	-0.086 (0.072) [0.057]	-0.060 (0.122) [0.073]
Enrolled in school (15-19 years old)	2,264	0.687	0.006 (0.025) [0.029]	0.168 (0.120) [0.118]	954	0.635	0.004 (0.039) [0.030]	0.183 (0.184) [0.128]	0.359 (0.227) [0.234]
Controls			Yes	Yes			Yes	Yes	Yes

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Controls include: conflict intensity (casualties/bomings), distance to the Albanian border, age, ethnicity (albanian), dummies for mothers and fathers educational attainment (medium and high), number of male and female adults in a household aged 20-65, number of children aged 0, 1-5, 6-10, 11-15, and 16-19, distance to school, dummy for rural location, municipality labor-force participation in 1991 and proportion of Albanians in 1991. LSMS sample weights are used in all the regressions. Data source: 1999 Kosovo DSHS and 2000 Kosovo LSMS

Table 3.13: Channels of Conflict Displacement on Labor Market Outcomes - Kosovo (1999-2000)

Households	1999 Kosovo DSHS			2000 Kosovo LSMS			
	Obs.	Mean	IV (WCR* Dist. Alb) (1)	Obs.	Mean	IV (WCR* Dist. Alb) (2)	IV (Bombs* Dist. Alb) (3)
Panel A: Assets							
Land Owned	7,190	0.431	-0.353 (0.323) [0.190]*	2,807	0.595	0.058 (0.275) [0.292]	0.397 (0.245) [0.287]
Land Size (in ha)	7,190	0.663	-0.236 (0.660) [0.454]	2,807	1.118	0.678 (0.727) [0.702]	-0.261 (0.959) [0.734]
Livestock	7,190	0.305	-0.294 (0.225) [0.178]*	2,807	0.413	-0.063 (0.135) [0.115]	-0.084 (0.145) [0.112]
Number of Livestock	—	—	—	2,807	6.161	-13.623 (4.900)*** [5.866]**	0.999 (5.503) [6.904]
Panel B: Social Networks							
Informal Networks	—	—	—	2,807	0.620	-0.581 (0.262)** [0.449]	0.249 (0.230) [0.320]
First stage				<i>Dependent variable: Displaced</i>			
Bombing x Dist. Alb							0.0004 (0.0001)*** (0.0001)**
WCR x Dist. Alb			0.0006 (0.0002)*** [0.0003]**			0.0006 (0.0002)*** (0.0003)**	
F-stat excl. Instrum.			10.00/6.17			12.48/5.66	11.43/5.41
Number of clusters			55/27			206/29	206/29
Partial R-squared			0.050			0.037	0.031
Observations			7,190			2,807	2,807
Controls			Yes			Yes	Yes

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Controls include: conflict intensity (casualties/bomings), distance to the Albanian border, age, ethnicity (albanian), dummies for marital status, dummies for educational attainment (low and medium), number of male and female adults in a household aged 20-65, number of dependent members aged 0, 1-5, 6-10, 11-15, and 16-19, dummy for rural location, municipality labor-force participation in 1991 and proportion of Albanians in 1991. LSMS sample weights are used in all the regressions. Data source: 1999 Kosovo DSHS and 2000 Kosovo LSMS.

Table 3.14: *Channels of Conflict Displacement on Education Outcomes - (IV estimation) - Kosovo (1999)*

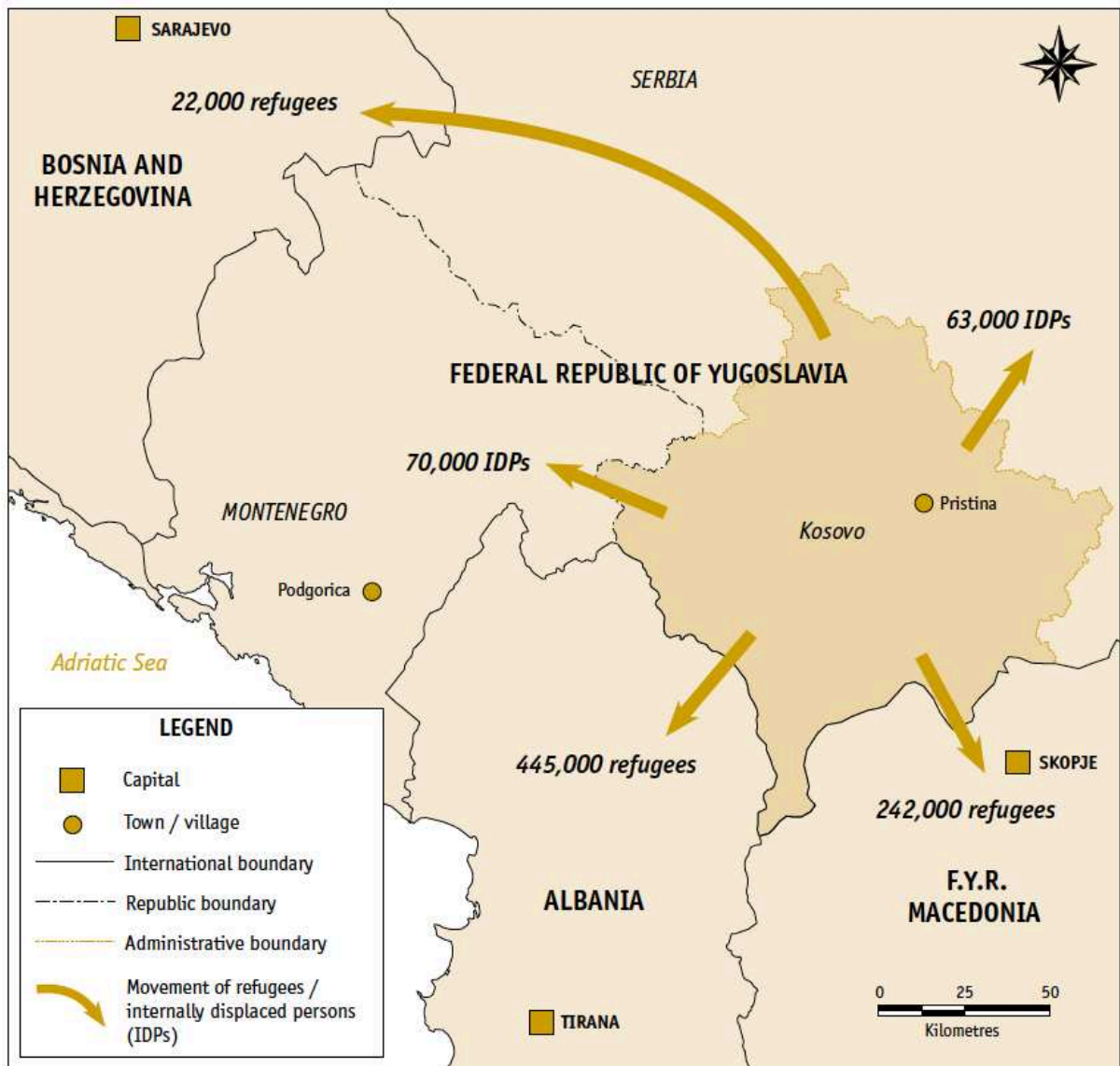
	All [6-19] y.o.		Primary Level [6-14] y.o.		Secondary Level [15-19] y.o.	
Second-stage	Dependent variable: Enrollment					
	(1)	(2)	(3)	(4)	(5)	(6)
1999 Kosovo DSHS						
Sample: FEMALE						
Displaced * Out of Kosovo	0.198 (0.080)** [0.049]***		0.129 (0.050)** [0.040]***		0.135 (0.163) [0.111]	
Displaced * Refugee Camp		0.337 (0.149)** [0.093]***		0.222 (0.106)** [0.095]**		0.223 (0.264) [0.159]
Mean Dep. Var.	0.749	0.749	0.864	0.864	0.556	0.556
First-stage	Dependent variable: Displaced * Out of Kosovo / Displaced * Refugee Camp					
	(1)	(2)	(3)	(4)	(5)	(6)
WCR x Dist. Alb.	0.0007 (0.0002)*** [0.0002]**	0.0004 (0.0001)** [0.0001]**	0.0007 (0.0002)*** [0.0003]**	0.0004 (0.0001)** [0.0001]**	0.0007 (0.0003)** [0.00023]**	0.0004 (0.0001)** [0.0002]**
Controls	Yes	Yes	Yes	Yes	Yes	Yes
F-stat excl. Instrum.	8.30/6.72	6.70/5.00	8.97/7.36	6.25/4.87	6.39/4.84	6.75/4.75
Number of clusters	55/27	55/27	55/27	55/27	55/27	55/27
Partial R-squared	0.051	0.025	0.054	0.026	0.045	0.023
Observations	5,919	5,919	3,709	3,709	2,210	2,210

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Controls include: conflict intensity (casualties/bomings), distance to the Albanian border, age, ethnicity (albanian), dummies for mothers and fathers educational attainment (medium and high), number of male and female adults in a household aged 20–65, number of children aged 0, 1-5, 6-10, 11-15, and 16-19, distance to school, dummy for rural location, municipality labor-force participation in 1991 and proportion of Albanians in 1991. LSMS sample weights are used in all the regressions. Data source: 1999 Kosovo DSHS

Appendix

3.A Supplementary Figures and Tables

Figure 3.A.1: *Displaced populations from Kosovo in neighbouring countries/territories, mid-June 1999*



Source: UNHCR (2000)

Figure 3.A.2: Total Cumulative refugee population in Montenegro, Albania, FYR Macedonia, Bosnia and Herzegovina and HEP, (March-October 1999)

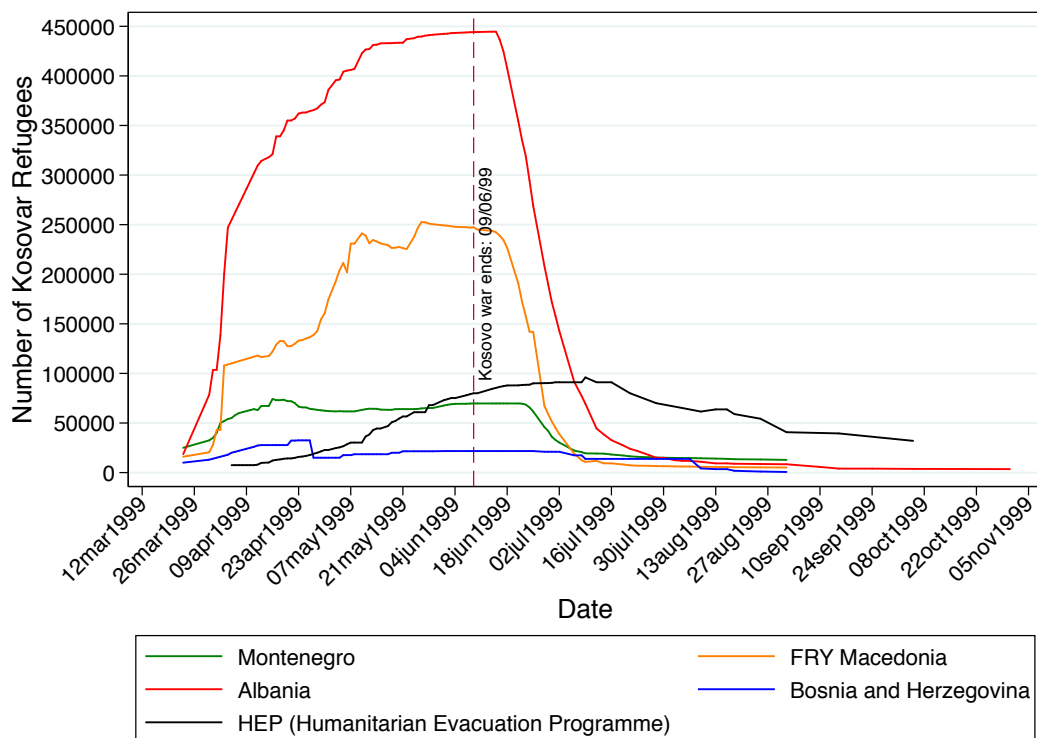
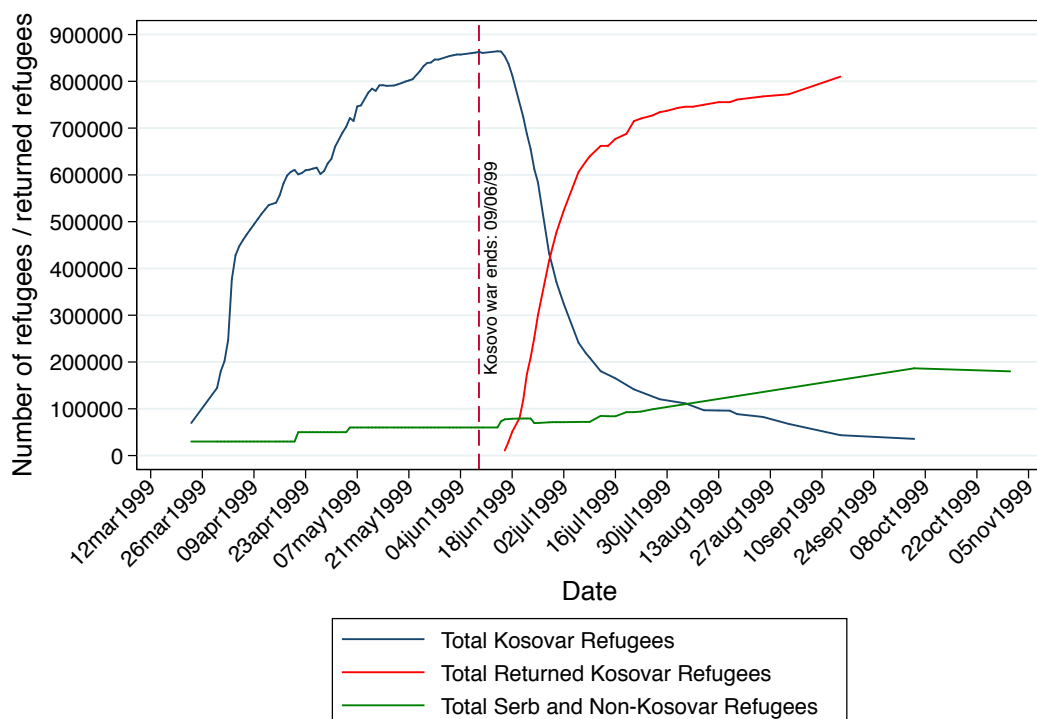


Figure 3.A.3: Total Cumulative Albanian Kosovar Refugee, Returned Albanian Kosovar Refugee and Serbian Refugee Populations (March-October 1999)



Source: UNHCR

Figure 3.A.4: *Daily Returned Refugees (June-October 1999) - UNHCR Estimates*

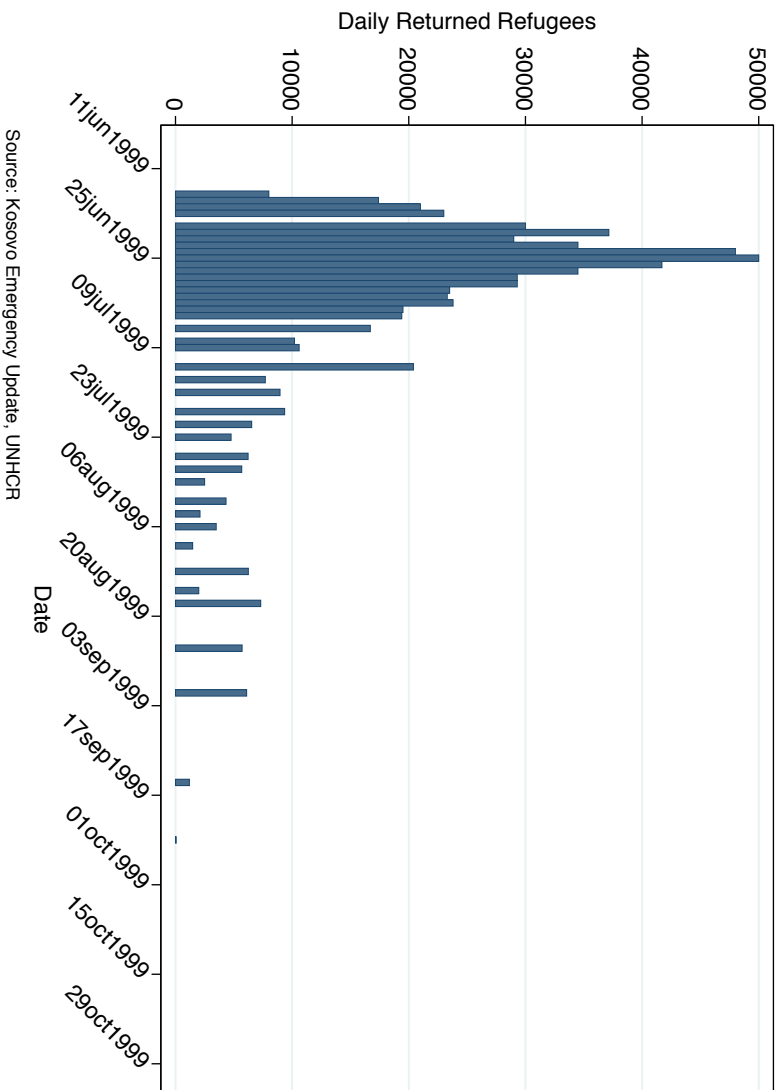


Figure 3.A.5: *Month of First and Last Displacement - 1999 DSHS*

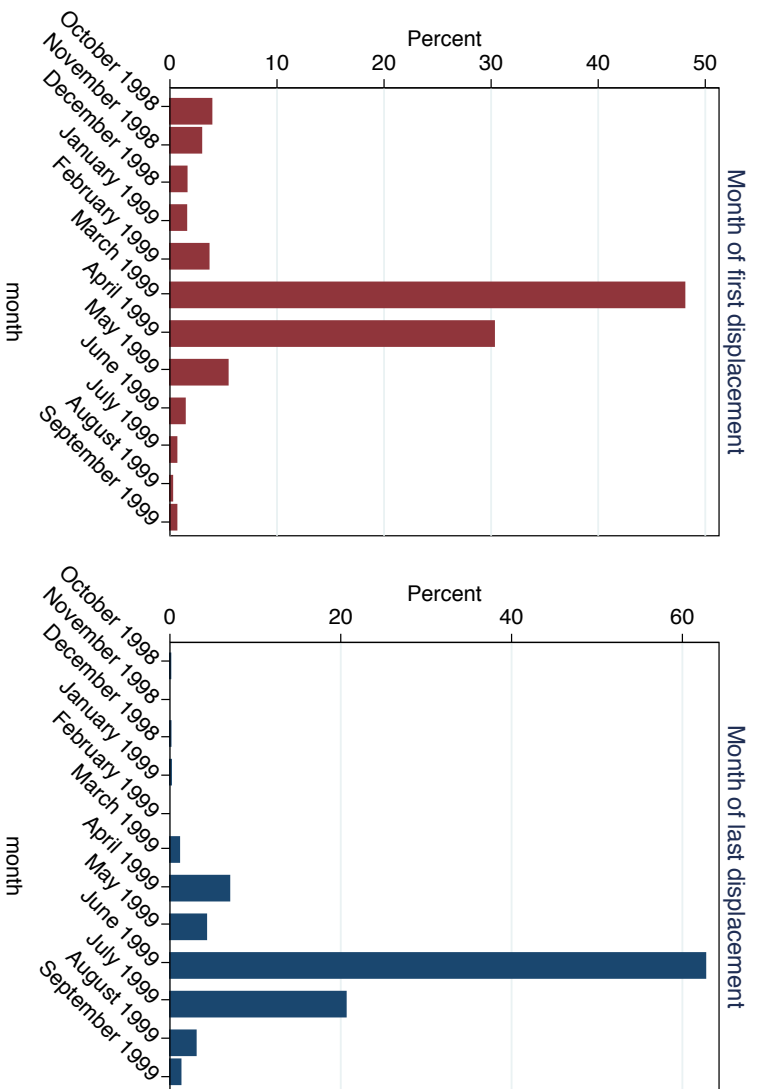


Figure 3.A.6: *Proportion of Displaced Individuals at the Municipality Level - 1999 Kosovo DSHS*

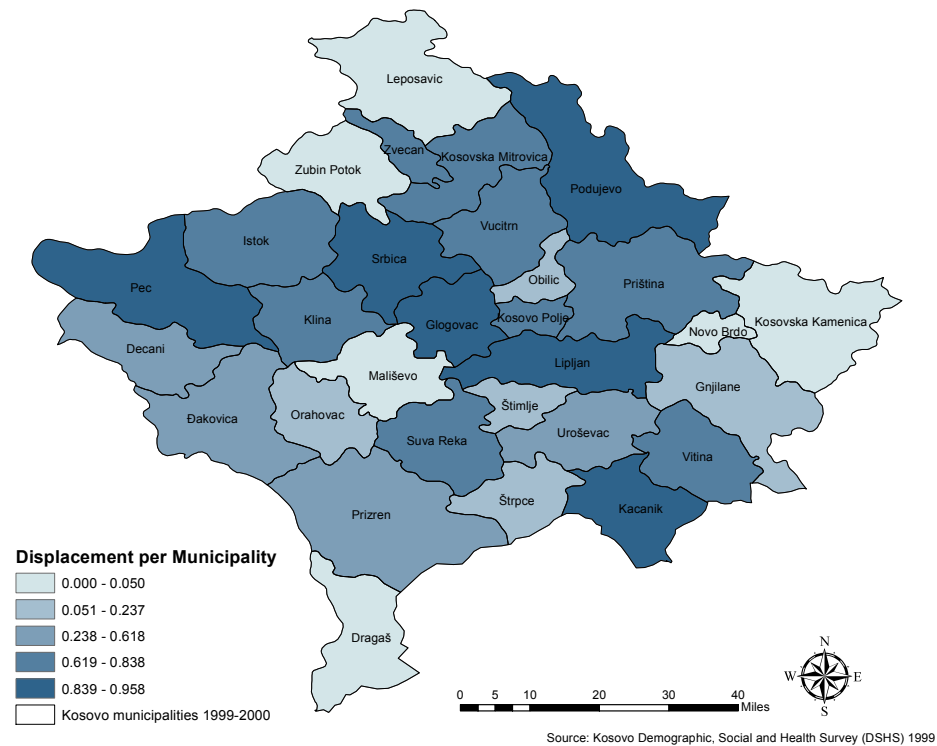
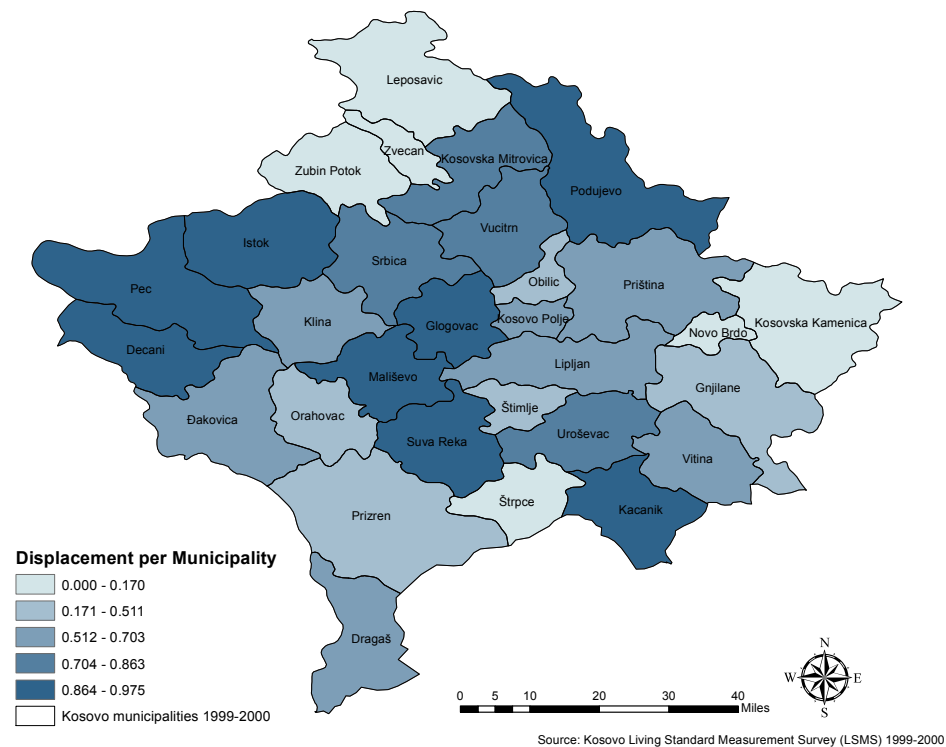


Figure 3.A.7: *Proportion of Displaced Individuals at the Municipality Level - 2000 Kosovo LSMS*



Notes: Each local proportion of displaced individuals is computed as the average displaced population at the municipality level in each database. This proportion ranges from 0-95% in the 1999 DSHS and from 0-97% in the 2000 LSMS.

Figure 3.A.8: Ethnic Majorities across Municipalities in 1991

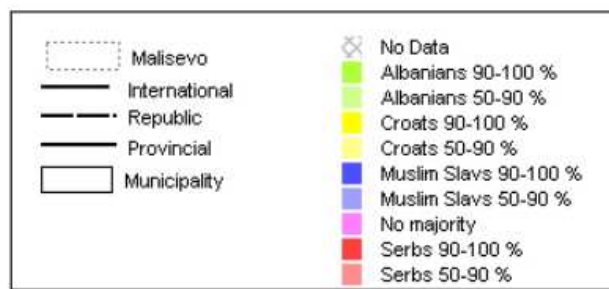


Figure 3.A.9: Labor Market Activity Status by Gender and Age Group - 1999 DSHS and 2000 LSMS

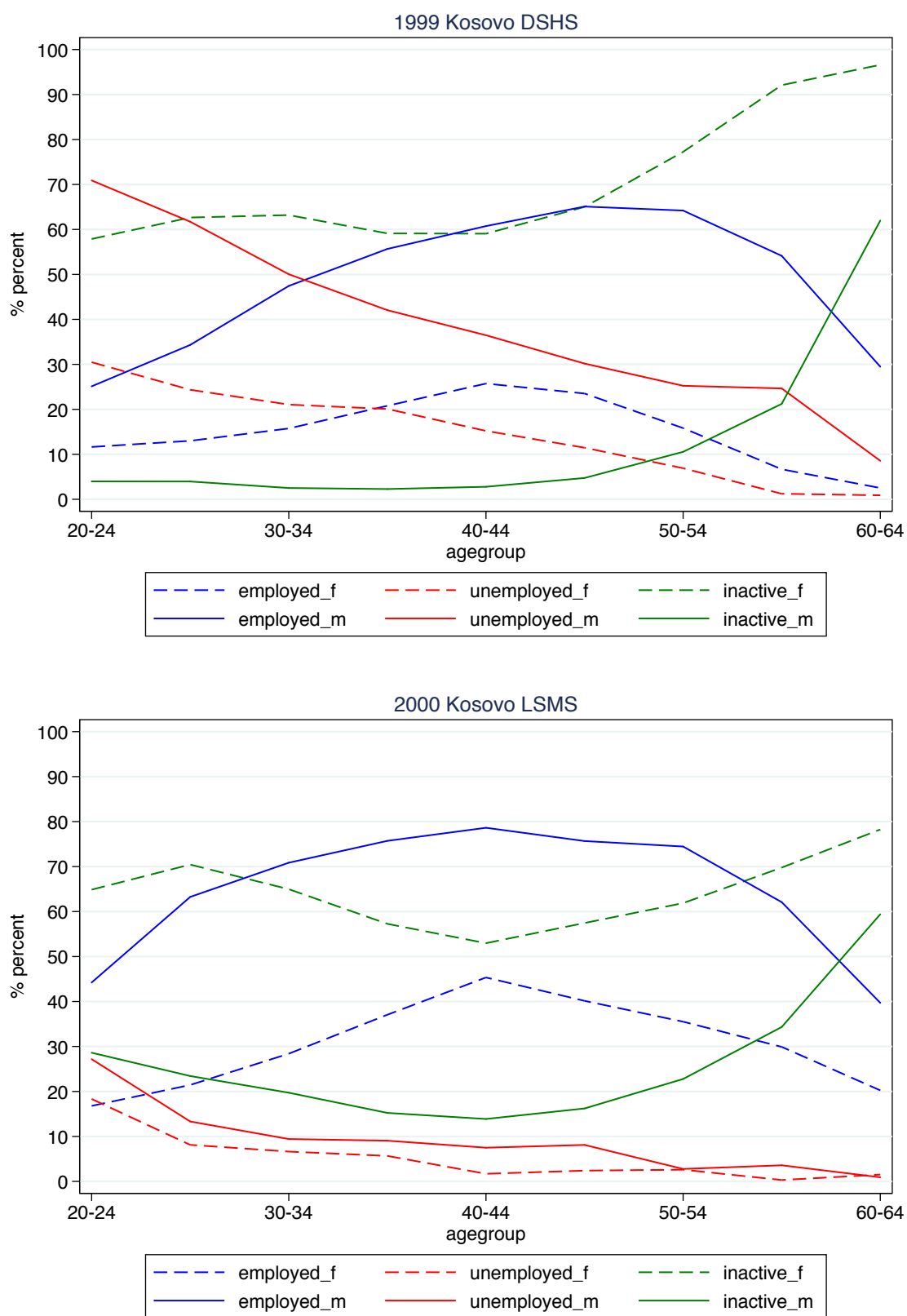


Figure 3.A.10: Enrollment Rates by Gender and Age - 1999 DSHS and 2000 LSMS

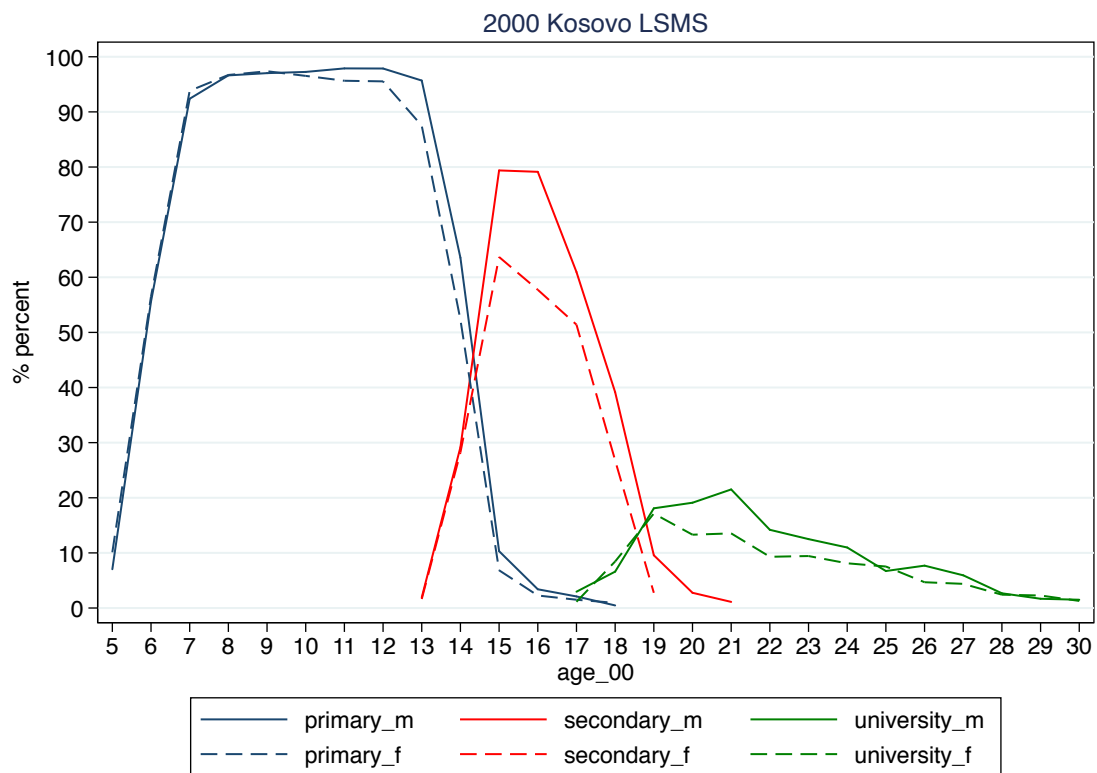
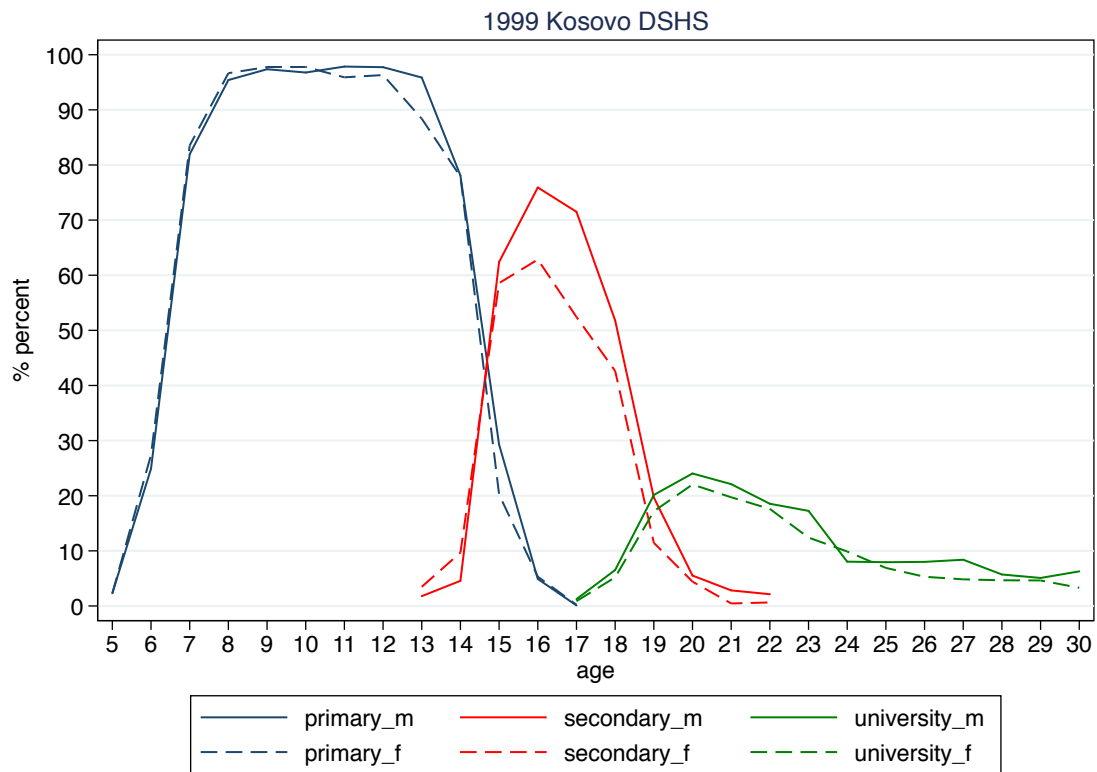
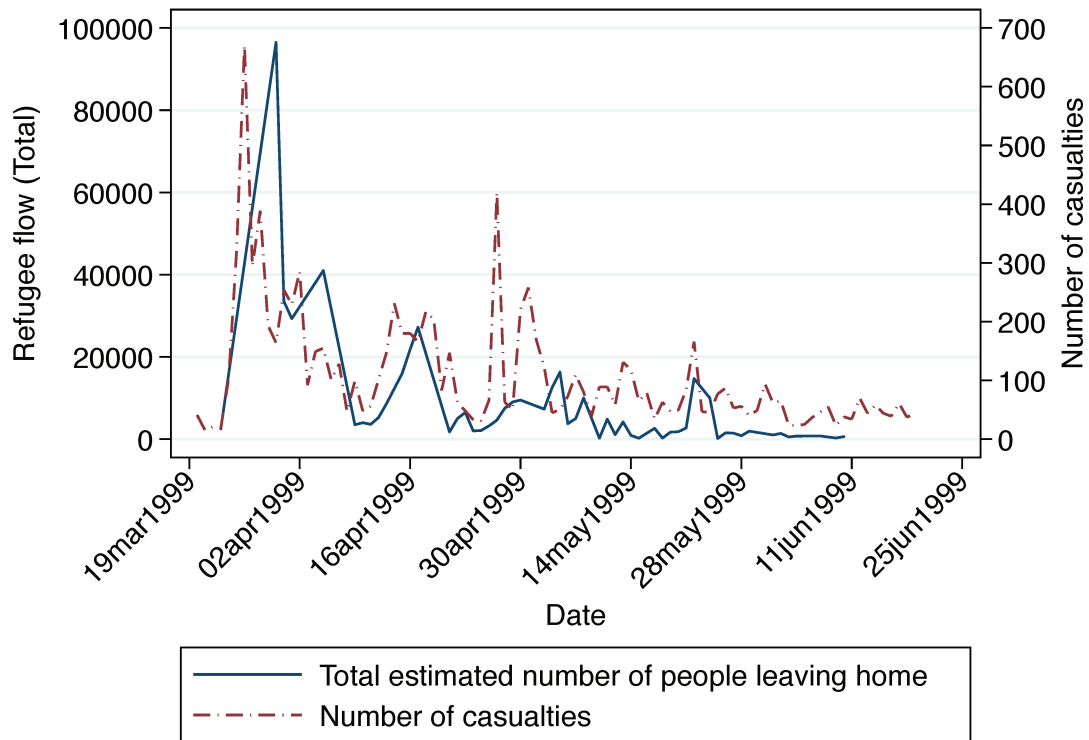
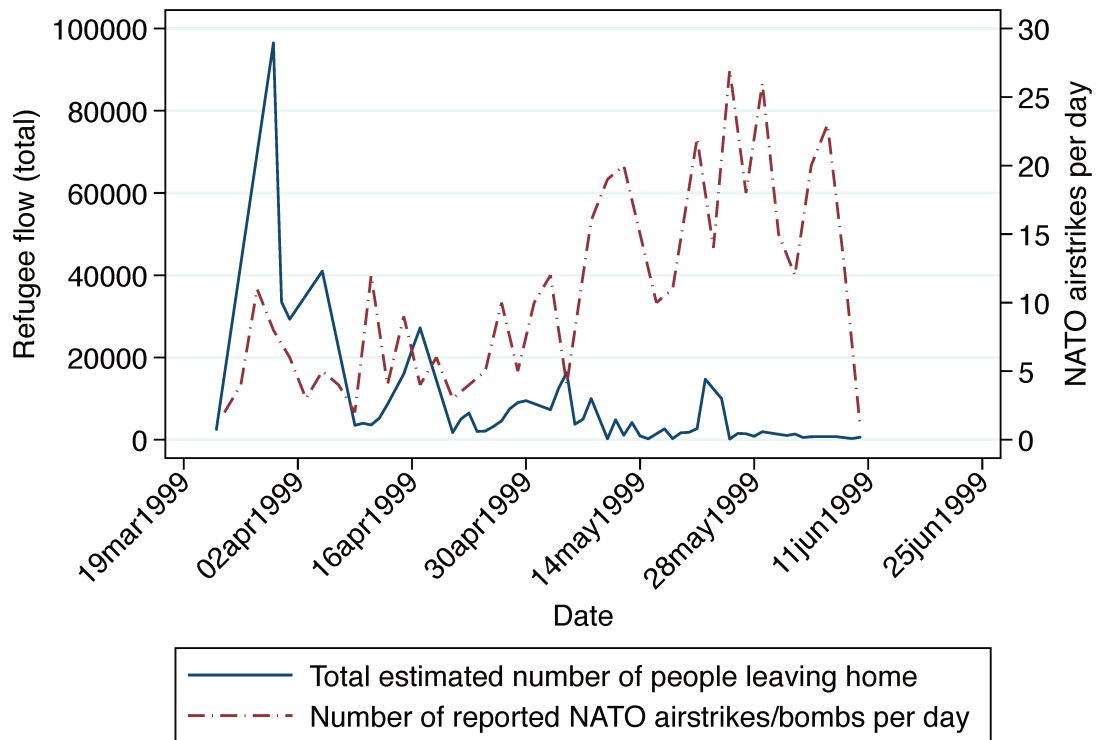


Figure 3.A.11: *Estimated Daily Refugee Displacement and Casualties (March - June 1999)*



Source: 1998-2000 KMB and UNHCR

Figure 3.A.12: *Estimated Daily Refugee Displacement and NATO Bombings (March - June 1999)*



Source: HRDAG and UNHCR

Table 3.A.1: *Compliers - Instrument (WCR x Distance to the Albanian Border) -*

<i>Dependent variable: Displaced</i>							
	Baseline	Less LFP in 1991	More	Less Agricultural in 1991	More	Less Albanians in 1991	More
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
WCR x Dist. Alb	0.0007 (0.0002)*** [0.0003]**	-0.0000 (0.0002) [0.0001]	0.0018 (0.0002)*** [0.0004]***	0.0000 (0.0002) [0.0002]	0.0016 (0.0002)*** [0.0004]***	0.0011 (0.0002)*** [0.0003]***	0.0007 (0.0002)*** [0.0004]*
WCR	-0.0500 (0.0138)*** [0.0201]**	-0.0026 (0.0153) [0.0128]	-0.1150 (0.0154)*** [0.0263]***	-0.0032 (0.0168) [0.0141]	-0.1137 (0.0178)*** [0.0299]***	-0.0720 (0.0120)*** [0.0143]***	-0.0551 (0.0177)*** [0.0299]*
Dist. Alb	-0.0032 (0.0015)** [0.0022]	0.0026 (0.0016) [0.0013]*	-0.0064 (0.0014)*** [0.0023]**	0.0021 (0.0016)*** [0.0013]	-0.0077 (0.0014)*** [0.0025]***	-0.0026 (0.0011)** [0.0010]**	-0.0054 (0.0017)*** [0.0033]
Observations	8,956	4,572	4,384	4,605	4,351	4,746	4,158
R-squared	0.2573	0.2684	0.3446	0.2591	0.3359	0.3749	0.0879
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% Controls include: conflict intensity (casualties/bomings), distance to the Albanian border, age, dummies for parental educational attainment (low and medium), number of male and female adults in a household aged 20-65, number of dependent members aged 0, 1-5, 6-10, 11-15, and 16-19 and dummy for rural location. Column 1 examines the full sample (pooled male and female). Municipalities that had a lower labor force participation in 1991 than the median municipality are examined in column 2, whereas municipalities that had a higher labor force participation in 1991 as compared to the median are examined in column 3. Municipalities that had a lower percentage of their population working in agriculture in 1991 than the median municipality are examined in column 4, whereas municipalities that had a higher percentage of their population working in agriculture as compared to the median municipality are examined in column 5. Municipalities in which a lower percentage of the population spoke Albanian in 1991 than in the median municipality are examined in column 6, whereas municipalities in which a higher than average percentage of the population spoke Albanian are examined in column 7. Data source: 2000 Kosovo LSMS.

Table 3.A.2: *Compliers - Instrument (Bombings x Distance to the Albanian Border) -*

<i>Dependent variable: Displaced</i>							
	Baseline	Less LFP in 1991	More	Less Agricultural in 1991	More	Less Albanians in 1991	More
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Bombings x Dist. Alb	0.0005 (0.0001)*** [0.0002]**	0.0003 (0.0002)** [0.0001]**	0.0006 (0.0002)** [0.0003]*	0.0001 (0.0002) [0.0001]	0.0009 (0.0003)*** [0.0003]**	0.0005 (0.0001)*** [0.0001]***	0.0004 (0.0002)** [0.0003]
Bombings	-0.0578 (0.0145)*** [0.0220]**	-0.0284 (0.0176) [0.0093]***	-0.0726 (0.0240)*** [0.0302]**	-0.0046 (0.0195) [0.0161]	-0.0781 (0.0241)*** [0.0272]**	-0.0573 (0.0132)*** [0.0142]**	-0.0451 (0.0205)** [0.0326]
Dist. Alb	-0.0025 (0.0011)** [0.0021]	0.0005 (0.0010) [0.0009]	-0.0021 (0.0019) [0.0030]	0.0017 (0.0010)* [0.0010]	-0.0065 (0.0016)*** [0.0022]**	-0.0013 (0.0011) [0.0011]	-0.0028 (0.0014)** [0.0028]
Observations	8,956	4,572	4,384	4,605	4,351	4,746	4,158
R-squared	0.2516	0.2743	0.2864	0.2634	0.3192	0.3573	0.0629
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1% Controls include conflict intensity (casualties/bomings), distance to the Albanian border, age, dummies for parental educational attainment (low and medium), number of male and female adults in a household aged 20-65, number of dependent members aged 0, 1-5, 6-10, 11-15, and 16-19 and dummy for rural location. Column 1 examines the full sample (pooled male and female). Municipalities that had a lower labor force participation in 1991 than the median municipality are examined in column 2, whereas municipalities that had a higher labor force participation in 1991 as compared to the median are examined in column 3. Municipalities that had a lower percentage of their population working in agriculture in 1991 than the median municipality are examined in column 4, whereas municipalities that had a higher percentage of their population working in agriculture as compared to the median municipality are examined in column 5. Municipalities in which a lower percentage of the population spoke Albanian in 1991 than in the median municipality are examined in column 6, whereas municipalities in which a higher than average percentage of the population spoke Albanian are examined in column 7. Data source: 2000 Kosovo LSMS.

Table 3.A.3: The Effect of Conflict Displacement on Children's Enrollment - (Wild Bootstrap Inference)

Children [6-19] Years Old	1999 Kosovo DSHS				2000 Kosovo LSMS				
			OLS	IV (WCR* Dist. Alb)			OLS	IV (WCR* Dist. Alb)	IV (Bombs* Dist. Alb)
	Obs.	Mean	(1)	(2)	Obs.	Mean	(3)	(4)	(5)
Sample: FEMALE									
Enrolled in school (6-19 years old)	5,919	0.749	0.014	0.188**	2,616	0.820	0.008	0.038	-0.004
<i>quasi-F</i>			(0.392)	(13.747)			(0.095)	(0.070)	(0.349)
<i>P-val</i>			0.785	0.0250			0.771	0.839	0.592
Enrolled in school (6-14 years old)	3,709	0.864	0.023	0.129**	1,650	0.920	0.029	-0.101	0.019
<i>quasi-F</i>			(1.462)	(13.231)			(1.805)	(0.946)	(2.407)
<i>P-val</i>			0.365	0.035			0.182	0.371	0.193
Enrolled in school (15-19 years old)	2,210	0.556	-0.008	0.118	966	0.635	-0.031	0.297	0.076
<i>quasi-F</i>			(0.065)	(1.296)			(0.693)	(1.247)	(2.165)
<i>P-val</i>			0.852	0.480			0.440	0.417	0.126
Sample: MALE									
Enrolled in school (6-19 years old)	6,385	0.804	0.006	0.028	2,716	0.820	0.008	0.039	0.076
<i>quasi-F</i>			(0.424)	(0.494)			(0.270)	(0.349)	(0.546)
<i>P-val</i>			0.527	0.556			0.612	0.592	0.511
Enrolled in school (6-14 years old)	4,121	0.869	-0.004	0.001	1,762	0.920	0.011	-0.086	-0.060
<i>quasi-F</i>			(0.099)	(0.001)			(0.248)	(2.407)	(0.722)
<i>P-val</i>			0.824	0.977			0.636	0.193	0.349
Enrolled in school (15-19 years old)	2,264	0.687	0.006	0.168	954	0.635	0.004	0.183	0.359
<i>quasi-F</i>			(0.048)	(3.915)			(0.013)	(2.165)	(2.487)
<i>P-val</i>			0.872	0.251			0.911	0.126	0.150
Controls			Yes	Yes			Yes	Yes	Yes

Notes: Standard errors are clustered at the municipality level. The *Quasi-F* statistic shown in parenthesis is the test statistic computed using wild bootstrap with clustered standard errors. *P-val* indicates the wild bootstrap P-value from Cameron and Miller (2015). Wild bootstrap P-values are obtained with the post-estimation command *boottest* by Roodman (2015), using Rademacher weights, assuming the null hypothesis and setting replications to 1000. *** P-val <0.01, ** P-val <0.05, * P-val <0.1. Controls include: conflict intensity (casualties/bomings), distance to the Albanian border, age, ethnicity (albanian), dummies for marital status, dummies for parental educational attainment (low and medium), number of male and female adults in a household aged 20-65, number of dependent members aged 0, 1-5, 6-10, 11-15, and 16-19, dummy for rural location, municipality labor-force participation in 1991 and proportion of Albanians in 1991. LSMS sample weights are used in all the regressions. Data source: 1999 Kosovo DSHS and 2000 Kosovo LSMS.

Table 3.A.4: *The Effect of Conflict Displacement on Women's Labor Market Outcomes - Kosovo (Wild Bootstrap Inference)*

Female [20-65]	1999 Kosovo DSHS				2000 Kosovo LSMS				
			OLS	IV			OLS	IV	IV
				(WCR* Dist.Alb)				(WCR* Dist.Alb)	(Bombs* Dist.Alb)
	Obs.	Mean	(1)	(2)	Obs.	Mean	(3)	(4)	(5)
Employed	10,751	0.150	-0.018*	-0.056	4,639	0.290	-0.015	0.192	0.023
<i>quasi-F</i>			(3.619)	(0.344)			(0.264)	(1.437)	(0.057)
<i>P-val</i>			0.077	0.741			0.608	0.156	0.832
Work off-farm	—	—	—	—	4,639	0.078	-0.009	-0.029	0.075*
<i>quasi-F</i>							(0.325)	(0.426)	(2,954)
<i>P-val</i>							0.771	0.487	0.101
Work on-farm	—	—	—	—	4,639	0.189	-0.010	0.298	-0.041
<i>quasi-F</i>							(0.124)	(2.160)	(0.149)
<i>P-val</i>							0.740	0.198	0.758
Work for someone	10,751	0.135	-0.014	-0.083	—	—	—	—	—
<i>quasi-F</i>			(2.461)	(0.758)					
<i>P-val</i>			0.146	0.563					
Work for family	10,751	0.010	-0.005	0.028	—	—	—	—	—
<i>quasi-F</i>			(1.897)	(0.891)					
<i>P-val</i>			0.205	0.452					
Self-Employed	10,751	0.005	0.001	-0.0004	4,639	0.019	0.001	-0.058	0.003
<i>quasi-F</i>			(0.174)	(0.000)			(0.074)	(2.382)	(0.015)
<i>P-val</i>			0.686	0.995			0.791	0.173	0.907
Unemployed	10,751	0.173	-0.014	-0.184	4,639	0.070	-0.001	-0.018	-0.040
<i>quasi-F</i>			(0.359)	(2.753)			(0.003)	(0.117)	(0.417)
<i>P-val</i>			0.564	0.211			0.949	0.732	0.550
Inactive	10,751	0.676	0.033	0.239**	4,639	0.638	0.016	0.174	0.017
<i>quasi-F</i>			(1.699)	(5.327)			(0.334)	(1.328)	(0.024)
<i>P-val</i>			0.243	0.043			0.586	0.167	0.902
Hours worked (0)	—	—	—	—	4,639	9.57	-0.596	7.666	10.378
<i>quasi-F</i>							(0.240)	(0.775)	(2.588)
<i>P-val</i>							0.639	0.409	0.090
Hours worked	—	—	—	—	1,303	34.08	-0.561	-7.459	12.139
<i>quasi-F</i>							(0.086)	(0.173)	(1.454)
<i>P-val</i>							0.758	0.805	0.177
Controls			Yes	Yes			Yes	Yes	Yes

Notes: Standard errors are clustered at the municipality level. The *Quasi-F* statistic shown in parenthesis is the test statistic computed using wild bootstrap with clustered standard errors. *P-val* indicates the wild bootstrap P-value from Cameron and Miller (2015). Wild bootstrap P-values are obtained with the post-estimation command *boottest* by Roodman (2015), using Rademacher weights, assuming the null hypothesis and setting replications to 1000. *** P-val < 0.01, ** P-val < 0.05, * P-val < 0.1. Controls include: conflict intensity (casualties/bomings), distance to the Albanian border, age, ethnicity (albanian), dummies for marital status, dummies for parental educational attainment (low and medium), number of male and female adults in a household aged 20-65, number of dependent members aged 0, 1-5, 6-10, 11-15, and 16-19, dummy for rural location, municipality labor-force participation in 1991 and proportion of Albanians in 1991. LSMS sample weights are used in all the regressions. Data source: 1999 Kosovo DSHS and 2000 Kosovo LSMS.

Table 3.A.5: *The Effect of Conflict Displacement on Men's Labor Market Outcomes - Kosovo (Wild Bootstrap Inference)*

Male [20-65]	1999 Kosovo DSHS				2000 Kosovo LSMS				
			OLS	IV (WCR* Dist.Alb)			OLS	IV (WCR* Dist.Alb)	IV (Bombs* Dist.Alb)
	Obs.	Mean	(1)	(2)	Obs.	Mean	(3)	(4)	(5)
Employed	9,454	0.474	-0.016 (0.214)	-0.272 (2.345)	4,317	0.642	-0.017 (0.717)	-0.104 (0.750)	-0.121 (0.719)
<i>quasi-F</i>			0.672	0.223			0.437	0.529	0.419
<i>P-val</i>									
Work off-farm	—	—	—	—	4,317	0.284	-0.032 (1.773)	0.064 (0.396)	0.234* (4.429)
<i>quasi-F</i>							0.238	0.599	0.090
<i>P-val</i>									
Work on-farm	—	—	—	—	4,317	0.248	-0.027 (1.620)	-0.140* (2.685)	-0.090 (0.274)
<i>quasi-F</i>							0.223	0.096	0.635
<i>P-val</i>									
Work for someone	9,454	0.365	0.028 (1.023)	-0.003 (0.000)	—	—	—	—	—
<i>quasi-F</i>			0.373	0.990					
<i>P-val</i>									
Work for family	9,454	0.025	-0.016 (1.303)	-0.095 (2.152)	—	—	—	—	—
<i>quasi-F</i>			0.424	0.364					
<i>P-val</i>									
Self-Employed	9,454	0.084	-0.028 (2.317)	-0.174** (7.550)	4,317	0.146	0.017 (0.960)	0.030 (0.121)	-0.060 (0.145)
<i>quasi-F</i>			0.201	0.067			0.365	0.750	0.771
<i>P-val</i>									
Unemployed	9,454	0.422	-0.008 (0.066)	0.201 (1.064)	4,317	0.113	0.013 (0.700)	0.086 (1.516)	-0.106 (0.724)
<i>quasi-F</i>			0.796	0.447			0.443	0.367	0.429
<i>P-val</i>									
Inactive	9,454	0.103	0.024 (5.922)	0.071 (2.119)	4,317	0.244	0.004 (0.059)	0.018 (0.038)	0.227 (2.328)
<i>quasi-F</i>			0.021	0.242			0.807	0.872	0.143
<i>P-val</i>									
Hours worked (0)	—	—	—	—	4,317	27.74	-1.104 (0.962)	-9.328 (0.054)	-8.496 (0.918)
<i>quasi-F</i>							0.377	0.832	0.413
<i>P-val</i>									
Hours worked	—	—	—	—	2,706	44.26	-0.175 (0.025)	3.336 (0.538)	-8.778 (0.591)
<i>quasi-F</i>							0.868	0.512	0.410
<i>P-val</i>									
Controls			Yes	Yes			Yes	Yes	Yes

Notes: Standard errors are clustered at the municipality level. The *Quasi-F* statistic shown in parenthesis is the test statistic computed using wild bootstrap with clustered standard errors. *P-val* indicates the wild bootstrap P-value from Cameron and Miller (2015). Wild bootstrap P-values are obtained with the post-estimation command *boottest* by Roodman (2015), using Rademacher weights, assuming the null hypothesis and setting replications to 1000. *** P-val < 0.01, ** P-val < 0.05, * P-val < 0.1. Controls include: conflict intensity (casualties/bomings), distance to the Albanian border, age, ethnicity (albanian), dummies for marital status, dummies for parental educational attainment (low and medium), number of male and female adults in a household aged 20-65, number of dependent members aged 0, 1-5, 6-10, 11-15, and 16-19, dummy for rural location, municipality labor-force participation in 1991 and proportion of Albanians in 1991. LSMS sample weights are used in all the regressions. Data source: 1999 Kosovo DSHS and 2000 Kosovo LSMS.

Table 3.A.6: The Effect of Displacement on Labor Market Outcomes by Occupation Type

Men [20-65] Years Old	1999 Kosovo DSHS			2000 Kosovo LSMS			
	Obs.	Mean	IV (WCR* Dist. Alb) (1)	Obs.	Mean	IV (WCR* Dist. Alb) (2)	IV (Bombs* Dist. Alb) (3)
Agricultural/Livestock	9,455	0.064	-0.450 (0.150)*** [0.121]***	4,317	0.208	-0.133 (0.111) [0.087]	-0.227 (0.132)** [0.162]
Industry and Mining	9,455	0.076	-0.105 (0.092) [0.122]	4,317	0.009	0.004 (0.019) [0.017]	0.025 (0.024) [0.018]
Construction	9,455	0.050	0.042 (0.060) [0.072]	4,317	0.093	0.063 (0.053) [0.051]	0.081 (0.084) [0.064]
Artisanry	9,455	0.041	0.092 (0.053)* [0.049]	4,317	0.012	0.012 (0.017) [0.012]	0.014 (0.022) [0.016]
Electricity and Water Supply	9,455	0.033	0.008 (0.017) [0.021]	4,317	0.025	0.005 (0.020) [0.023]	0.013 (0.032) [0.033]
Education	9,455	0.053	0.010 (0.028) [0.026]	4,317	0.044	-0.109 (0.041)*** [0.055]**	-0.025 (0.043) [0.041]
Public Administration	9,455	0.060	-0.001 (0.038) [0.037]	4,317	0.057	-0.007 (0.044) [0.059]	0.086 (0.071) [0.053]*
Trade	9,455	0.077	0.061 (0.047) [0.041]	4,317	0.072	0.002 (0.034) [0.031]	-0.023 (0.058) [0.043]
Transport	—	—	— (0.174)	4,317	0.036	-0.003 (0.030) [0.023]	-0.003 (0.043) [0.033]
Finance and Banking	9,455	0.010	-0.012 (0.009) [0.009]	4,317	0.010	0.014 (0.012) [0.011]	-0.001 (0.016) [0.017]
Health	9,455	0.024	0.036 (0.026) [0.024]	4,317	0.011	0.015 (0.015) [0.018]	-0.004 (0.018) [0.020]
Housekeeping	—	—	—	4,317	0.001	0.0001 (0.002) [0.002]	0.004 (0.007) [0.004]
Turism	9,455	0.011	0.005 (0.008) [0.008]	—	—	—	—
Controls			Yes			Yes	Yes

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. Controls include: conflict intensity (casualties/bomings), distance to the Albanian border, age, ethnicity (albanian), dummies for marital status, dummies for parental educational attainment (low and medium), number of male and female adults in a household aged 20-65, number of dependent members aged 0, 1-5, 6-10, 11-15, and 16-19, dummy for rural location, municipality labor-force participation in 1991 and proportion of Albanians in 1991. LSMS sample weights are used in all the regressions. Data source: 1999 Kosovo DSHS and 2000 Kosovo LSMS.

Table 3.A.7: *The Effect of Displacement on Labor Market Outcomes by Occupation Type (Only Work off-farm)*

Men [20-65] Years Old	2000 Kosovo LSMS			
	Obs.	Mean	(WCR* Dist. Alb) (1)	(Bombs* Dist. Alb) (2)
Agricultural/Livestock	4,317	0.008	0.004 (0.017)	0.004 (0.015)
Industry and Mining	4,317	0.006	-0.017 (0.016)	0.014 (0.015)
Construction	4,317	0.068	-0.004 (0.015)	0.111* (0.064)
Artisanry	4,317	0.006	0.060 (0.040)	-0.009 (0.018)
Electricity and Water Supply	4,317	0.016	-0.004 (0.015)	0.015 (0.025)
Education	4,317	0.043	-0.078* (0.042)	-0.013 (0.036)
Public Administration	4,317	0.037	0.046 (0.035)	0.084* (0.045)
Trade	4,317	0.025	0.023 (0.024)	0.029 (0.023)
Transport	4,317	0.022	-0.039 (0.024)	-0.021 (0.025)
Finance and Banking	4,317	0.009	0.026 (0.016)	0.009 (0.016)
Health	4,317	0.010	0.016 (0.012)	-0.005 (0.015)
Housekeeping	4,317	0.0009	0.001 (0.001)	0.009 (0.006)
Controls			Yes	Yes

Notes: Standard errors clustered at the village of residence are in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Controls include conflict intensity (casualties/bomings), distance to the Albanian border, dummies for marital status, ethnicity (albanian), dummies for educational attainment (medium and high), number of male and female adults in a household aged 20–65, number of dependent members aged 0, 1-5, 6-10, 11-15, and 16-18, individual age-group dummies (four years by four years from 23 to 62, and one for 63 to 65), rural location and dummy for land ownership. LSMS sample weights are used in all the regressions. Data source: 2000 Kosovo LSMS.

3.B Alternative Identification Strategy

3.B.1 Identification: Difference-in-Difference (DID) Strategy

Additionally, I also use a difference-in-difference (DID) empirical strategy. This strategy exploits two sources of variation in order to isolate the effect of forced displacement on schooling completion: the spatial variation in municipality displacement and the birth cohorts of children - which determines whether they were in primary or secondary school during the forced displacement from Kosovo. Using the cross-sectional data of the 1999 Kosovo DSHS and 2000 Kosovo LSMS, I compare primary and secondary schooling completion outcomes across birth cohorts.

This identification strategy uses the characteristics of the education system of Kosovo till the year 2000 which regulated 8 years of mandatory or primary level schooling between the ages 6-14, and 4 years of secondary schooling between the ages 15-18 (SOK, 2001).²⁹ Therefore, for each education level, I define as pre-displacement or unaffected cohorts those that completed primary/secondary schooling before the conflict displacement, that is, before September 1998. While, displacement or affected cohorts are those that should have completed the last year of primary/secondary schooling during or after the conflict displacement, that is, after June 1999.

Table 3.B.1 presents the composition of the samples used to identify cohort effects on schooling completion outcomes. The samples are restricted to observations of boys and girls, whose cohort characteristics of schooling completion are observed before or after the 1999 Kosovo displacement. Sample A concentrates on primary schooling completion. The unaffected cohort contains children born between 1979-1983, which ensures that the child is at least 15 years old before the start of the conflict displacement and has already finished primary school. In contrast, the affected cohort contains children born between 1984-1986, which ensures that the child is younger than 15 years old in 1999. This means that her primary schooling was interrupted by the conflict displacement. Similarly, Sample B presents the unaffected and affected cohorts for the secondary school completion. Children born between 1975-1979 belong to the pre-displacement cohorts, as they finished secondary school in peacetime. While, all children born between 1980-1982 have experienced conflict displacement during their schooling.

²⁹From 2002 school year, compulsory education in Kosovo was extended to 9 years and divided into 5 years of primary education and 4 years of lower secondary education.

Here, the identification of conflict displacement relies on the assumption that differences in schooling completion between affected and unaffected cohorts would have been equal across municipalities in the absence of forced displacement. For the DID estimation, the regression model is specified as follows:³⁰

$$Schooling_{ijkt} = \alpha(MDisp_j * A_t) + \lambda X'_{ijkt} + \gamma_j + \delta_k + c_t + \epsilon_{ijkt} \quad (3.5)$$

where $Schooling_{ijkt}$: is the schooling outcome variable for child i of birth cohort t residing in municipality j before displacement (and in k after displacement), $MDisp_j$ is proportion of the population that was displaced in municipality j , A_t is a dummy indicating whether child born in cohort t was affected by the 1999 displacement, X'_{ijkt} is set of control variables (ethnicity, parental schooling completion, household demographics and rural dummy), γ_j are pre-displacement municipality fixed-effects; δ_k are post-displacement municipality fixed-effects; c_t are birth cohort fixed-effects and ϵ_{ijkt} is the error term. The two sets of municipality fixed effects capture municipal heterogeneity in schooling conditions. The parameter of interest which estimates the effect of conflict displacement on schooling outcomes is thus α . Standard errors are clustered at the village and municipality level.

3.B.2 Endogeneity in Municipality Displacement

Even though the municipality fixed effects take care of endogeneity arising from unobserved pre- and post-war municipal conditions, the identification assumption could still be violated by self-selection due to unobserved individual heterogeneity such as ability, social networks, risk aversion etc.³¹ In order to control for this, I use conflict intensity -as measured by casualties and bombings at the municipality level- interacted with distance to Albanian border as an instrument for municipality displacement. The identifying assumption underlying this alternative estimation strategy is that the change in schooling outcomes over cohorts are similar across municipalities that suffered different conflict intensities and were located near of far

³⁰Similar difference-in-difference regression models have been used earlier in the literature to assess the effect of armed conflict on child schooling (e.g. [Swee, 2015](#); [Chamarbagwala and Morán, 2011](#); [Shemyakina, 2011](#); [Leon, 2012](#); [Akresh and De Walque, 2008](#); [Akbulut-Yuksel, 2014](#); [Merrouche, 2011](#); [Valente, 2013](#); [Pivovarova and Swee, 2015](#)). All these models have been influenced by earlier applications of cross-sectional difference-in-difference models by Esther Duflo in different research settings (see [Duflo \(2001, 2003\)](#).)

³¹This endogeneity issue can potentially be overcome by accounting for individual fixed-effect using panel data and assuming that the unobserved attribute is time-invariant. However, this is clearly infeasible since in this case the data are cross-sectional.

from the Albanian border, in the absence of forced displacement.

The first-stage specifications are thus described by the following equations:

$$(MDisp_j * A_t) = \beta_1(WCR_j * D_v * A_t) + \beta_2(WCR_j * D_v) + \lambda_1 X'_{ijkt} + \gamma_j + \delta_k + c_t + \epsilon_{ijkt} \quad (3.6)$$

$$(MDisp_j * A_t) = \beta_1(B_j * D_v * A_t) + \beta_2(B_j * D_v) + \lambda_1 X'_{ijkt} + \gamma_j + \delta_k + c_t + \epsilon_{ijkt} \quad (3.7)$$

More precisely, I predict the interaction of municipality displacement and affected cohort ($MDisp_j * A_t$) with interactions of municipality conflict intensity -measured as casualties (WCR_j) and bombings (B_j)- with distance to the Albanian border (D_v) and affected cohort (A_t). On the one hand, conflict intensity and distance to the Albanian border are good predictors for displacement. On the other hand, the interaction of these variables ($WCR_j * D_v$) with affected cohort (A_t) should be uncorrelated with unobserved individual characteristics in the schooling equation.

Tables 3.B.2 and 3.B.3 show the regression coefficients of the first-stage estimation for the primary and secondary school children, respectively. In the same line as the findings with the IV strategy, these results indicate that further away from the Albanian border, an increase in conflict intensity leads to higher municipality displacement for the affected cohorts. Both specifications are statistically significant at conventional levels. In particular, the interaction between war casualty rate and distance to the Albanian border seems to be a stronger predictor for municipality displacement in both databases. In other words, affected children located further away from the Albanian border and in municipalities with more casualties are more likely to experience forced displacement.

Similarly, the second specification which uses the interaction between bombing intensity and distance to the Albanian border as instrument also shows that further way from the Albanian border, an increase in bombing intensity leads to higher municipality displacement. However, the F-statistic of the excluded instruments is below 10 in most of the cases, which makes this instrument less strong and subject to bias.

3.B.3 Displacement and Schooling Completion Outcomes

This Section presents the estimation results following the difference-in-difference strategy outlined previously. I estimate the impact of conflict displacement on schooling outcomes of Kosovar children after the 1999 Kosovo war. I first present the impact of conflict displacement

on female schooling completion in Table 3.B.4 and then, in Table 3.B.5, I present the effect of conflict displacement on male schooling completion. The primary (secondary) level sample comprises children aged 13-20 (17-24) in 1999. The average primary schooling completion is quite equal across genders, for boys around 87% and for girls 85%. The average secondary schooling completion for boys drops to around 60-65%, while for girls the drop is much more pronounced being around 45-50%, suggesting clear inequality in secondary school attainment across genders.

For each schooling outcome, I run three sets of difference-in-difference regressions - first, OLS; second, IV using ($WCR * DistAlb$) as instrument and third, IV using ($Bombings * DistAlb$) as instrument-. Each cell in Tables 3.B.4 and 3.B.5 presents the coefficients of interest, the interaction of belonging to the affected cohort and living in a municipality with high displacement. Since statistical inference when using difference-in-difference models is vulnerable to serial correlation that possibly produces a downward bias, I apply standard errors clustered at the village and municipality level.

By and large, I find no evidence of municipality displacement effects on neither primary nor secondary schooling completion for boys and girls in general. These results suggest that there are no significant patterns of cohort-specific displacement intensity effects. In particular, only the specification that uses ($Bombings * DistAlb$) as instrument shows a negative and statistically significant effect of displacement on children's secondary schooling completion in 1999. However, this result is subject to bias due to the weakness of this instrument in the first-stage.

Table 3.B.1: *Identification of Displacement Affected Cohorts and Unaffected Cohorts by Schooling Completion Outcomes*

	Sample A		Sample B	
	<i>Primary Schooling Completion</i>		<i>Secondary Schooling Completion</i>	
	Pre-Displacement (Unaffected)	Displacement (Affected)	Pre-Displacement (Unaffected)	Displacement (Affected)
1999 Kosovo DSHS				
Birth cohort	1979 - 1983	[1984 - 1985]	[1975 - 1979]	[1980 - 1981]
Age	>15	[14 - 15]	>19	[18 - 19]
2000 Kosovo LSMS				
Birth cohort	[1979 - 1983]	[1984 - 1986]	[1975 - 1979]	[1980 - 1982]
Age	>16	[14 - 16]	>20	[18 - 20]

Table 3.B.2: *First-stage Results for Primary School Completion Outcome - (Difference-in-Difference Estimation)*

	1999 Kosovo DSHS				2000 Kosovo LSMS			
Cohorts: [14-21] Years Old	Dependent variable: Affected cohort (aged 14-15 in 1999) * Municipality Displacement							
	Female		Male		Female		Male	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
WCR * Dist.Alb.	-0.0002 (0.0001)** [0.0001]**		-0.0002 (0.0001)** [0.0001]**		-0.0002 (0.0001)** [0.0001]**		-0.0002 (0.0001)*** [0.0001]**	
Affected cohort * WCR * Dist.Alb	0.0005 (0.0001)*** [0.0002]***		0.0005 (0.0001)*** [0.0001]***		0.0004 (0.00005)*** [0.0001]***		0.0004 (0.00006)*** [0.0001]***	
Bombings * Dist.Alb.		0.00001 (0.00003) [0.00002]		-0.00004 (0.00007) [0.00007]		-0.0001 (0.0001) [0.0000]		0.0001 (0.0001) [0.0001]
Affected cohort * Bombings * * Dist.Alb.		0.0001 (0.00004)*** [0.00004]***		0.0002 (0.0001)*** [0.0001]***		0.0002 (0.0001)*** [0.0001]*		0.0002 (0.0001)*** [0.0001]**
Observations	2,314	2,314	2,458	2,458	1,196	1,196	1,292	1,292
Number of clusters	55/27	55/27	55/27	55/27	190/29	190/29	198/29	198/29
F-stat Excluded Instruments	15.96/9.40	7.20/9.45	15.35/7.93	16.40/13.03	30.27/5.89	6.20/1.94	25.54/6.63	4.61/5.96
Partial R-squared	0.358	0.107	0.410	0.140	0.107	0.125	0.313	0.105
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-displacement municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Post-displacement municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets.* significant at 10%; ** significant at 5%; *** significant at 1%. The samples in both surveys contain individuals aged 14 and above in 1999 for the primary schooling completion and individuals 18 and above in 1999 for the secondary schooling completion. Individual controls include ethnicity (=1 if Albanian), parental secondary schooling completion, number of siblings and dummy for rural residence. Data source: 1999 Kosovo DSHS and 2000 Kosovo LSMS.

Table 3.B.3: First-stage Results for Secondary School Completion Outcome - (Difference-in-Difference Estimation)

Cohorts: [18-25] Years Old	1999 Kosovo DSHS				2000 Kosovo LSMS			
	Dependent variable: Affected cohort (aged 18-19 in 1999) * Municipality Displacement							
	Female		Male		Female		Male	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
WCR * Dist.Alb.	-0.0001 (0.00007)** [0.00006]**		-0.0001 (0.0001) [0.0001]		-0.0001 (0.00008)* [0.00008]*		-0.0001 (0.00008)** [0.00007]**	
Affected cohort * WCR * * Dist.Alb.	0.0005 (0.0001)** [0.0002]**		0.0005 (0.0001)** [0.0001]**		0.0004 (0.0001)** [0.0001]**		0.0005 (0.0001)** [0.0001]**	
Bombings * Dist.Alb.		-0.0001 (0.0001) [0.0001]*		0.0001 (0.0001) [0.0001]		-0.0001 (0.0001) [0.0001]		-0.00003 (0.00008) [0.00006]
Affected cohort * Bombings * * Dist.Alb.		0.0002 (0.0001)** [0.0001]**		0.0001 (0.0001)** [0.0001]**		0.0001 (0.00004)** [0.0001]*		0.0002 (0.0001)** [0.0001]*
Observations	1,687	1,687	1,919	1,919	799	799	1,113	1,113
Number of clusters	55/27	55/27	55/27	55/27	172/29	172/29	192/29	192/29
F-stat Excluded Instruments	12.68/12.68	4.50/4.19	12.33/41.29	2.94/2.80	31.67/6.61	6.37/2.76	30.80/6.80	6.00/1.83
Partial R-squared	0.392	0.120	0.410	0.091	0.358	0.082	0.426	0.124
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-displacement municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Post-displacement municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets.* significant at 10%; ** significant at 5%; *** significant at 1%. The samples in both surveys contain individuals aged 14 and above in 1999 for the primary schooling completion and individuals 18 and above in 1999 for the secondary schooling completion. Individual controls include ethnicity (=1 if Albanian), parental secondary schooling completion, number of siblings and dummy for rural residence. Data source: 1999 Kosovo DSHS and 2000 Kosovo LSMS.

Table 3.B.4: *The Effect of Municipality Displacement on Female's Schooling Attainment - Kosovo (1999-2000)*

Dependent Variable:	Completed Primary School			Completed Secondary School		
	DID	DID-IV (WCR* Dist.Alb)	DID-IV (Bombs* Dist.Alb)	DID	DID-IV (WCR* Dist.Alb)	DID-IV (Bombs* Dist.Alb)
	(1)	(2)	(3)	(4)	(5)	(6)
Sample: FEMALE						
1999 Kosovo DSHS						
Affected cohort (14-15 in 1999)	0.002	0.043	0.047			
* Municipality Displacement	(0.057)	(0.069)	(0.122)			
	[0.042]	[0.052]	[0.126]			
Affected cohort (18-19 in 1999)				0.066	0.126	-0.474
* Municipality Displacement				(0.062)	(0.101)	(0.159)***
				[0.067]	[0.086]	[0.158]***
Mean of dep. variable	0.840	0.840	0.840	0.476	0.476	0.476
Observations	2,314	2,314	2,314	1,687	1,687	1,687
Number of clusters	55/27	55/27	55/27	55/27	55/27	55/27
R-squared	0.308	0.304	0.294	0.325	0.289	0.323
Sample: FEMALE						
2000 Kosovo LSMS						
Affected cohort (14-15 in 1999)	-0.010	-0.029	0.134			
* Municipality Displacement	(0.058)	(0.165)	(0.228)			
	[0.049]	[0.062]	[0.113]			
Affected cohort (18-19 in 1999)				0.085	0.085	-0.108
* Municipality Displacement				(0.083)	(0.126)	(0.193)
				[0.078]	[0.075]	[0.167]
Mean of dep. variable	0.851	0.851	0.851	0.414	0.414	0.414
Observations	1,196	1,196	1,196	799	799	799
Number of clusters	190/29	190/29	190/29	172/29	172/29	172/29
R-squared	0.225	0.224	0.225	0.329	0.331	0.341
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Pre-war municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Post-war municipality FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. The samples in both surveys contain individuals aged 14 and above in 1999 for the primary schooling completion and individuals 18 and above in 1999 for the secondary schooling completion. Individual controls include ethnicity (=1 if Albanian), parental secondary schooling completion, number of siblings and dummy for rural residence. Schooling attainment is a binary indicator. Data source: 1999 Kosovo DSHS and 2000 Kosovo LSMS.

Table 3.B.5: *The Effect of Municipality Displacement on Male's Schooling Attainment - Kosovo*

Dependent Variable:	Completed Primary School			Completed Secondary School		
	DID	DID-IV (WCR* Dist.Alb)	DID-IV (Bombs* Dist.Alb)	DID	DID-IV (WCR* Dist.Alb)	DID-IV (Bombs* Dist.Alb)
	(1)	(2)	(3)	(4)	(5)	(6)
Sample: MALE 1999 Kosovo DSHS						
Affected cohort (14-15 in 1999)	-0.012	-0.078	0.146			
* Municipality Displacement	(0.058)	(0.066)	(0.114)			
	[0.053]	[0.077]	[0.106]			
Affected cohort (18-19 in 1999)				-0.013	-0.015	-0.255
* Municipality Displacement				(0.060)	(0.069)	(0.103)**
				[0.052]	[0.064]	[0.093]***
Mean of dep. variable	0.861	0.861	0.861	0.645	0.645	0.645
Observations	2,458	2,458	2,458	1,919	1,919	1,919
Number of clusters	55/27	55/27	55/27	55/27	55/27	55/27
R-squared	0.370	0.365	0.371	0.232	0.229	0.236
Sample: MALE 2000 Kosovo LSMS						
Affected cohort (14-15 in 1999)	0.001	-0.218	0.207			
* Municipality Displacement	(0.067)	(0.150)	(0.172)			
	[0.071]	[0.166]	[0.189]			
Affected cohort (18-19 in 1999)				-0.048	0.026	-0.085
* Municipality Displacement				(0.080)	(0.135)	(0.200)
				[0.098]	[0.126]	[0.141]
Mean of dep. variable	0.887	0.887	0.887	0.594	0.594	0.594
Observations	1,292	1,292	1,292	1,113	1,113	1,113
Number of clusters	198/29	198/29	198/29	192/29	192/29	192/29
R-squared	0.343	0.316	0.335	0.227	0.219	0.221
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Birth cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Pre-war municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Post-war municipality FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors clustered at the village of residence are in parentheses. Standard errors clustered at the municipality level are in brackets. * significant at 10%; ** significant at 5%; *** significant at 1%. The samples in both surveys contain individuals aged 14 and above in 1999 for the primary schooling completion and individuals 18 and above in 1999 for the secondary schooling completion. Individual controls include ethnicity (=1 if Albanian), parental secondary schooling completion, number of siblings and dummy for rural residence. Schooling attainment is a binary indicator. Data source: 1999 Kosovo DSHS and 2000 Kosovo LSMS.

Bibliography

- Aaronson, Daniel, Rajeev H Dehejia, Andrew Jordan, Cristian Pop-Eleches, Cyrus Samii, and Karl Schulze.** 2017. "The Effect of Fertility on Mothers' Labor Supply over the Last Two Centuries." Mimeo.
- Abdel-Rahim, Asha, Dany Jaimovich, and Aleksi Ylönen.** 2015. "Forced displacement and behavioral change: an empirical study of returnee households in the Nuba Mountains." *Defence and Peace economics*, 1–31.
- Agüero, Jorge.** 2013. "Causal estimates of the intangible costs of violence against women in Latin America and the Caribbean." Inter American Development Bank Working Paper No. IDB-WP 414.
- Agüero, Jorge M, and Mindy S Marks.** 2008. "Motherhood and female labor force participation: evidence from infertility shocks." *The American Economic Review*, 98(2): 500–504.
- Agüero, Jorge M, and Mindy S Marks.** 2011. "Motherhood and female labor supply in the developing world evidence from infertility shocks." *Journal of Human Resources*, 46(4): 800–826.
- Aizer, Anna.** 2010. "The gender wage gap and domestic violence." *The American Economic Review*, 100(4): 1847.
- Aizer, Anna.** 2011. "Poverty, violence, and health the impact of domestic violence during pregnancy on newborn health." *Journal of Human Resources*, 46(3): 518–538.
- Akbulut-Yuksel, Mevlude.** 2014. "Children of War The Long-Run Effects of Large-Scale Physical Destruction and Warfare on Children." *Journal of Human Resources*, 49(3): 634–662.
- Akresh, Richard, and Damien De Walque.** 2008. "Armed conflict and schooling: Evidence from the 1994 Rwandan genocide." *World Bank Publications*, 3516.
- Alva, Soumya, Edmundo Murrugarra, and Pierella Paci.** 2002. "The hidden costs of ethnic conflict: decomposing trends in educational outcomes of young Kosovars." *World Bank Policy Research Working Paper*, No. 2880.
- Amaral, Sofia, Prakash Nishith, and Sonia Bhalotra.** 2018. "Gender, Crime and Punishment: Evidence from Women Police Stations in India." *MIMEO*, 104: 123–135.
- Angelucci, Manuela.** 2008. "Love on the rocks: Domestic violence and alcohol abuse in rural Mexico." *The BE Journal of Economic Analysis & Policy: Contributions to Economic Analysis and Policy*, 8(1): 43.

- Angrist, J.D., and W.N. Evans.** 1998. "Children and their parents' labor supply: Evidence from exogenous variation in family size." *American Economic Review*, 80(3): 450–477.
- Angrist, Joshua D, and William N Evans.** 1996. "Children and their parents' labor supply: Evidence from exogenous variation in family size." National Bureau of Economic Research.
- Arias, María Alejandra, Ana María Ibáñez, and Pablo Querubin.** 2014. "The desire to return during civil war: Evidence for internally displaced populations in Colombia." *Peace Economics, Peace Science and Public Policy*, 20(1): 209–233.
- Arnold, Fred, and Liu Zhaoxiang.** 1986. "Sex preference, fertility, and family planning in China." *Population and Development Review*, 12: 221–246.
- Attanasio, Orazio, and Valerie Lechene.** 2002. "Tests of income pooling in household decisions." *Review of economic dynamics*, 5(4): 720–748.
- Bailey, Martha J.** 2013. "Fifty years of family planning: new evidence on the long-run effects of increasing access to contraception." National Bureau of Economic Research.
- Baldry, Anna C.** 2003. "Bullying in schools and exposure to domestic violence." *Child abuse & neglect*, 27(7): 713–732.
- Ball, Patrick, Wendy Betts, Fritz Scheuren, Jana Dudukovich, and Jana Asher.** 2002. "Killings and Refugee Flow in Kosovo March-June 1999." *Report to the Intl Criminal Tribunal for the former Yugoslavia*.
- Banerjee, Abhijit, Raghavendra Chattopadhyay, Esther Duflo, Daniel Keniston, and Nina Singh.** 2012. "Improving police performance in Rajasthan, India: Experimental evidence on incentives, managerial autonomy and training." National Bureau of Economic Research.
- Bauer, Thomas K, Sebastian Braun, and Michael Kvasnicka.** 2013. "The Economic Integration of Forced Migrants: Evidence for Post-War Germany." *The Economic Journal*, 123(571): 998–1024.
- Becker, Gary S, and H Gregg Lewis.** 1973. "On the Interaction between the Quantity and Quality of Children." *Journal of Political Economy*, 81(2): 279–288.
- Becker, Gary S, and Nigel Tomes.** 1976. "Child endowments and the quantity and quality of children." *Journal of Political Economy*, 84(4): 143–162.
- Bellows, John, and Edward Miguel.** 2009. "War and local collective action in Sierra Leone." *Journal of Public Economics*, 93(11): 1144–1157.
- Bisbee, James, Rajeev Dehejia, Cristian Pop-Eleches, and Cyrus Samii.** 2015. "Local instruments, global extrapolation: external validity of the labor supply-fertility local average treatment effect." National Bureau of Economic Research.
- Black, David S, Steve Sussman, and Jennifer B Unger.** 2010. "A further look at the intergenerational transmission of violence: Witnessing interparental violence in emerging adulthood." *Journal of Interpersonal Violence*, 25(6): 1022–1042.

- Blattman, Christopher, and Jeannie Annan.** 2010. "The consequences of child soldiering." *The Review of Economics and Statistics*, 92(4): 882–898.
- Blau, David M, and Philip K Robins.** 1998. "Child-care costs and family labor supply." *The Review of Economics and Statistics*, 70(3): 374–381.
- Bloch, Francis, Vijayendra Rao, and Sonalde Desai.** 2004. "Wedding celebrations as conspicuous consumption signaling social status in rural India." *Journal of Human Resources*, 39(3): 675–695.
- Bloom, David E, David Canning, Günther Fink, and Jocelyn E Finlay.** 2009. "Fertility, female labor force participation, and the demographic dividend." *Journal of Economic Growth*, 14(2): 79–101.
- Blunch, Niels-Hugo, Sudharshan Canagarajah, Dhushyanth Raju, et al.** 2001. "The informal sector revisited: A synthesis across space and time." *World Bank Social Protection Discussion Papers*, 119.
- Bobonis, Gustavo J, Melissa González-Brenes, and Roberto Castro.** 2013. "Public transfers and domestic violence: The roles of private information and spousal control." *American Economic Journal: Economic Policy*, 5(1): 179–205.
- Boca, Daniela Del, Silvia Pasqua, and Chiara Pronzato.** 2005. "Fertility and employment in Italy, France, and the UK." *Labour*, 19(1): 51–77.
- Boesten, Jelke.** 2012. "The state and violence against women in Peru: intersecting inequalities and patriarchal rule." *Social Politics: International Studies in Gender, State & Society*, 19(3): 361–382.
- Bott, Sarah, Alessandra Guedes, Mary Goodwin, and Jennifer Adams Mendoza.** 2012. "Violence Against Women in Latin America and the Caribbean: A comparative analysis of population-base." *Pan American Health Organization*.
- Bozzoli, Carlos, Tilman Brueck, and Tony Muhumuzi.** 2016. "Activity Choices of Internally Displaced Persons And Returnees: Quantitative Survey Evidence from Post-War Northern Uganda." *Bulletin of Economic Research*, 68(4): 329–247.
- Brakman, Steven, Harry Garretsen, and Marc Schramm.** 2004. "The strategic bombing of German cities during World War II and its impact on city growth." *Journal of Economic Geography*, 4(2): 201–218.
- Brassiolo, Pablo.** 2016. "Domestic violence and divorce law: When divorce threats become credible." *Journal of Labor Economics*, 34(2): 443–477.
- Bronars, Stephen G, and Jeff Grogger.** 1994. "The economic consequences of unwed motherhood: Using twin births as a natural experiment." *The American Economic Review*, 84(5): 1141–1156.
- Bruijns, Sanne.** 2014. "Children and Their Parents' Labor Supply: Evidence from 2012 Dutch Data." Working Paper, University of Amsterdam.

- Brunborg, Helge.** 2002. "Report on the Size and Ethnic Composition of the Population of Kosovo." *Case of Slodovan Milosevic (IT-02-54), ICTY, La Haye*.
- Burgert, Clara R, Josh Colston, Thea Roy, and Blake Zachary.** 2013. "Geographic displacement procedure and georeferenced data release policy for the Demographic and Health Surveys." Calverton Maryland ICF International.
- Bütikofer, Aline.** 2010. "Sibling Sex Composition and Cost of Children." Mimeo.
- Cain, Glen G, and Martin D Dooley.** 1976. "Estimation of a model of labor supply, fertility, and wages of married women." *Journal of Political Economy*, 84(4, Part 2): 179–199.
- Cameron, A Colin, and Douglas L Miller.** 2015. "A practitioner's guide to cluster-robust inference." *Journal of Human Resources*, 50(2): 317–372.
- Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller.** 2008. "Bootstrap-based improvements for inference with clustered errors." *The Review of Economics and Statistics*, 90(3): 414–427.
- Campbell, Jacquelyn C.** 2002. "Health Consequences of Intimate Partner Violence." *The Lancet*, 359(9314): 1331–1336.
- Carlson, Bonnie E.** 2000. "Children exposed to intimate partner violence research findings and implications for intervention." *Trauma, Violence, & Abuse*, 1(4): 321–342.
- Carrell, Scott E, and Mark L Hoekstra.** 2010. "Externalities in the classroom: How children exposed to domestic violence affect everyone's kids." *American Economic Journal: Applied Economics*, 2(1): 211–228.
- Chamarbagwala, Rubiana, and Hilcías E Morán.** 2011. "The human capital consequences of civil war: Evidence from Guatemala." *Journal of Development Economics*, 94(1): 41–61.
- Chen, Siyan, Norman V Loayza, and Marta Reynal-Querol.** 2008. "The aftermath of civil war." *The World Bank Economic Review*, 22(1): 63–85.
- Chun, Hyunbae, and Jeungil Oh.** 2002. "An instrumental variable estimate of the effect of fertility on the labour force participation of married women." *Applied Economics Letters*, 9(10): 631–634.
- Clarke, Damian.** 2017. "Children and their Parents: A Review of Fertility and Causality." *Journal of Economic Surveys (Forthcoming)*.
- Compton, Janice, and Robert A Pollak.** 2014. "Family proximity, childcare, and women's labor force attachment." *Journal of Urban Economics*, 79: 72–90.
- Connelly, Rachel.** 1992. "The effect of child care costs on married women's labor force participation." *The Review of Economics and Statistics*, 74(1): 83–90.
- Cortes, Kalena E.** 2004. "Are refugees different from economic immigrants? Some empirical evidence on the heterogeneity of immigrant groups in the United States." *The Review of Economics and Statistics*, 86(2): 465–480.

- Crawford, Nicholas, John Cosgrave, Simone Haysom, and Nadine Walicki.** 2015. "Protracted Displacement: Uncertain paths to self-reliance in exile." London: Overseas Development Institute.
- Cristia, Julian P.** 2008. "The effect of a first child on female labor supply evidence from women seeking fertility services." *Journal of Human Resources*, 43(3): 487–510.
- Cruces, Guillermo, and Sebastian Galiani.** 2007. "Fertility and female labor supply in Latin America: New causal evidence." *Labour Economics*, 14(3): 565–573.
- Currie, Cheryl L.** 2006. "Animal cruelty by children exposed to domestic violence." *Child abuse & neglect*, 30(4): 425–435.
- Cutts, Mark.** 2000. *The state of the world's refugees, 2000: Fifty years of humanitarian action*. Geneva: UNHCR, Oxford University Press.
- Czaika, Mathias, and Krisztina Kis-Katos.** 2009. "Civil conflict and displacement: village-level determinants of forced migration in Aceh." *Journal of Peace Research*, 46(3): 399–418.
- Danaj, Ermira.** 2014. "Family in Albania as a primary solidarity network." in "Legacy and Change: Albanian transformation from multidisciplinary perspectives", edited by Robert Pichler, LIT Verlag.
- Da Rocha, Jose Maria, and Luisa Fuster.** 2006. "Why are fertility rates and female employment ratios positively correlated across OECD countries?" *International Economic Review*, 47(4): 1187–1222.
- Davis, Donald, and David Weinstein.** 2002. "Bones, bombs and break points: The geography of economic activity." *The American Economic Review*, 95(2): 1269–1289.
- De Chaisemartin, Clement.** 2017. "Tolerating defiance? Local average treatment effects without monotonicity." *Quantitative Economics*, 8(2): 367–396.
- Dehejia, Rajeev, Cristian Pop-Eleches, and Cyrus Samii.** 2015. "From local to global: External validity in a fertility natural experiment." National Bureau of Economic Research.
- Deininger, Klaus.** 2003. "Causes and consequences of civil strife: micro-level evidence from Uganda." *Oxford Economic Papers*, 55(4): 579–606.
- Dell, Melissa.** 2012. "Path dependence in development: Evidence from the Mexican Revolution." *Harvard University, mimeograph*.
- Douarin, Elodie, Julie Litchfield, and Rachel Sabates-Wheeler.** 2012. "War legacies, welfare and livelihoods in rural areas: the case of Kosovo." *Journal of Agricultural Economics*, 63(1): 201–222.
- Duflo, Esther.** 2001. "Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment." *American Economic Review*, 91(4): 795–813.
- Duflo, Esther.** 2003. "Grandmothers and granddaughters: old-age pensions and intrahousehold allocation in South Africa." *The World Bank Economic Review*, 17(1): 1–25.

- Duflo, Esther.** 2012. "Women empowerment and economic development." *Journal of Economic Literature*, 50(4): 1051–1079.
- Durand, Julia Garcia, Lilia Blima Schraiber, Ivan França-Junior, and Claudia Barros.** 2011. "Impact of exposure to intimate partner violence on children's behavior." *Revista de Saude Publica*, 45(2): 355–364.
- Ebenstein, Avraham.** 2009. "When is the local average treatment close to the average? Evidence from fertility and labor supply." *Journal of Human Resources*, 44(4): 955–975.
- Eder, Christoph.** 2014. "Displacement and education of the next generation: evidence from Bosnia and Herzegovina." *IZA Journal of Labor & Development*, 3(1): 12.
- Edin, Per-Anders, Peter Fredriksson, and Olof Åslund.** 2003. "Ethnic enclaves and the economic success of immigrants—Evidence from a natural experiment." *The Quarterly Journal of Economics*, 118(1): 329–357.
- Edleson, Jeffrey L.** 1999. "Children's witnessing of adult domestic violence." *Journal of Interpersonal Violence*, 14(8): 839–870.
- Emery, Clifton R.** 2011. "Controlling for selection effects in the relationship between child behavior problems and exposure to intimate partner violence." *Journal of Interpersonal Violence*, 26(8): 1541–1558.
- Engel, Stefanie, and Ana María Ibáñez.** 2007. "Displacement due to violence in Colombia: A household-level analysis." *Economic Development and Cultural Change*, 55(2): 335–365.
- Esarey, Justin.** 2015. "Using Interaction Terms as Instrumental Variables for Causal Identification: Does Corruption Harm Economic Development?" Working Paper. Rice University.
- Eswaran, Mukesh, and Nisha Malhotra.** 2011. "Domestic violence and women's autonomy in developing countries: theory and evidence." *Canadian Journal of Economics/Revue canadienne d'économique*, 44(4): 122–1263.
- Even, William E.** 1987. "Career interruptions following childbirth." *Journal of Labor Economics*, 5(2): 255–277.
- Fafchamps, Marcel, and Susan Lund.** 2003. "Risk-sharing networks in rural Philippines." *Journal of Development Economics*, 71(2): 261–287.
- Fantuzzo, John, Robert Boruch, Abdullahi Beriama, Marc Atkins, and Susan Marcus.** 1997. "Domestic violence and children: Prevalence and risk in five major US cities." *Journal of the American Academy of Child & Adolescent Psychiatry*, 36(1): 116–122.
- Farmer, Amy, and Jill Tiefenthaler.** 1996. "Domestic violence: the value of services as signals." *The American Economic Review*, 86(2): 274–279.
- Ferris, Elizabeth, and Rebecca Winthrop.** 2010. "Education and displacement: Assessing conditions for refugees and internally displaced persons affected by conflict." Background paper for the EFA Global Monitoring Report 2011. The hidden crisis: Armed conflict and education.

- Fiala, Nathan.** 2015. "Economic consequences of forced displacement." *The Journal of Development Studies*, 51(10): 1275–1293.
- Flake, Dallan F, and Renata Forste.** 2006. "Fighting families: family characteristics associated with domestic violence in five Latin American countries." *Journal of Family Violence*, 21(1): 19.
- Fleisher, Belton M, and George F Rhodes.** 1979. "Fertility, women's wage rates, and labor supply." *The American Economic Review*, 69(1): 14–24.
- Foster, Andrew D, and Mark R Rosenzweig.** 2001. "Imperfect commitment, altruism, and the family: Evidence from transfer behavior in low-income rural areas." *The Review of Economics and Statistics*, 83(3): 389–407.
- Franco, Rocío, and María González.** 2009. "Las mujeres en la justicia comunitaria: víctimas, sujetos y actores." *Serie Justicia Comunitaria en los Andes: Perú y Ecuador*, 3.
- Fregoso, Rosa-Linda, and Cynthia Bejarano.** 2009. *Terrorizing Women: Femicide in the Americas*. Duke University Press.
- Friedemann-Sánchez, Greta, and Rodrigo Lovatón.** 2012. "Intimate partner violence in Colombia: Who is at risk?" *Social Forces*, 91(2): 663–688.
- Garcia-Moran, Eva, and Zoe Kuehn.** 2017. "With strings attached: Grandparent-provided child care and female labor market outcomes." *Review of Economic Dynamics*, 23(80-98).
- Garcia-Moreno, Claudia, Henrica AFM Jansen, Mary Ellsberg, Lori Heise, Charlotte H Watts, et al.** 2006. "Prevalence of intimate partner violence: findings from the WHO multi-country study on women's health and domestic violence." *The Lancet*, 368(9543): 1260–1269.
- Gelbach, Jonah B.** 2002. "Public schooling for young children and maternal labor supply." *The American Economic Review*, 92(1): 307–322.
- Gjonça, Arjan, Arnstein Aassve, and Letizia Mencarini.** 2008. "Albania: Trends and patterns, proximate determinants and policies of fertility change." *Demographic Research*, 19(11): 261–292.
- Glewwe, Paul, and Karthik Muralidharan.** 2016. "Improving education outcomes in developing countries: Evidence, knowledge gaps, and policy implications." In *Handbook of the Economics of Education*. Vol. 5, 653–743. Elsevier.
- Gong, Xiaodong, and Arthur Van Soest.** 2002. "Family structure and female labor supply in Mexico City." *Journal of Human Resources*, 37(1): 163–191.
- Grant, Rebecca.** 1999. *The Kosovo campaign: aerospace power made it work*. Air Force Association.
- Gronau, Reuben.** 1973. "The intrafamily allocation of time: The value of the housewives' time." *The American Economic Review*, 63(4): 634–651.
- Gronau, Reuben.** 1988. "Sex-related wage differentials and women's interrupted labor careers—The chicken or the egg." *Journal of Labor Economics*, 6(3): 277–301.
- Heath, Rachel.** 2012. "Women's access to labor market opportunities, control of household resources, and domestic violence."

- Heckman, James.** 1974. "Shadow prices, market wages, and labor supply." *Econometrica: journal of the econometric society*, 42: 679–694.
- Heckman, James J, and Robert J Willis.** 1977. "A beta-logistic model for the analysis of sequential labor force participation by married women." *Journal of Political Economy*, 85(1): 27–58.
- Heinemann, Alessandra, and Dorte Verner.** 2006. "Crime and violence in development: A literature review of Latin America and the Caribbean."
- Hidrobo, Melissa, and Lia Fernald.** 2013. "Cash transfers and domestic violence." *Journal of Health Economics*, 32(1): 304–319.
- Holt, Stephanie, Helen Buckley, and Sadhbh Whelan.** 2008. "The impact of exposure to domestic violence on children and young people: A review of the literature." *Child abuse & neglect*, 32(8): 797–810.
- Hotz, V Joseph, Susan Williams McElroy, and Seth G Sanders.** 2005. "Teenage childbearing and its life cycle consequences exploiting a natural experiment." *Journal of Human Resources*, 40(3): 683–715.
- Huber, Martin.** 2012. *Statistical verification of a natural" natural experiment": Tests and sensitivity checks for the sibling sex ratio instrument.* School of Economics and Political Science, Department of Economics, Univ.
- Hupkau, Claudia, and Marion Leturcq.** 2016. "Fertility and Labor Supply: New Evidence from the UK." Working Paper, Centre for Economic Performance, London School of Economics.
- Huth-Bocks, Alissa C, Alytia A Levendosky, and Michael A Semel.** 2001. "The direct and indirect effects of domestic violence on young children's intellectual functioning." *Journal of Family Violence*, 16(3): 269–290.
- Iacopino, Vincent, Martina W Frank, Heidi M Bauer, Allen S Keller, Sheri L Fink, Doug Ford, Daniel J Pallin, and Ronald Waldman.** 2001. "A population-based assessment of human rights abuses committed against ethnic Albanian refugees from Kosovo." *American Journal of Public Health*, 91(12): 2013–2018.
- Ibáñez, Ana, and Andrés Moya.** 2006. "The impact of intra-state conflict on economic welfare and consumption smoothing: Empirical evidence for the displaced population in Colombia." Households in Conflict Network Working Paper No. 23). Brighton: HiCN.
- Ibáñez, Ana María, and Carlos Eduardo Vélez.** 2008. "Civil conflict and forced migration: The micro determinants and welfare losses of displacement in Colombia." *World Development*, 36(4): 659–676.
- ICTY, The International Criminal Tribunal for the former Yugoslavia.** 2000. "Final Report to the Prosecutor by the Committee Established to Review the NATO Bombing Campaign Against the Federal Republic of Yugoslavia."
- IMDC.** 2016. "Global Report on Internal Displacement." Geneva: Internal Displacement Monitoring Center.

- IMF, International Monetary Fund.** 2006. "Albania: Poverty Reduction Strategy Paper - Annual Progress Report." In "Albania: Poverty Reduction Strategy Paper - Annual Progress Report". USA.
- INEI.** 2001. *Encuesta demográfica y de salud familiar*. Instituto Nacional de Estadística and Informática (Peru) and Macro International. MEASURE/DHS+ (Programme).
- INEI.** 2015. *Encuesta demográfica y de salud familiar*. Instituto Nacional de Estadística and Informática (Peru) and Macro International. MEASURE/DHS+ (Programme).
- INEI.** 2017. *Peru: Brechas de Género 2017 - Avances hacia la igualdad de mujeres y hombres*. Instituto Nacional de Estadística e Informática (Peru) (Last available 11/03/2018 at https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1444/libro.pdf).
- INSTAT, Albanian Institute of Statistics.** 2004. "People and work in Albania - Work force, Employment and Unemployment." Tirana: INSTAT Albania.
- INSTAT, Albanian Institute of Statistics.** 2014. "Women and Men in Albania." Tirana: INSTAT Albania.
- Jacobsen, Joyce P, James Wishart Pearce, and Joshua L Rosenbloom.** 2001. "The effects of child-bearing on women's marital status: using twin births as a natural experiment." *Economics Letters*, 70(1): 133–138.
- Jayasinghe, Stephnie, Pushpa Jayawardena, and Hemamali Perera.** 2009. "Influence of intimate partner violence on behaviour, psychological status and school performance of children in Sri Lanka." *Journal of Family Studies*, 15(3): 274–283.
- Jubb, Nadine, Gloria Camacho, Almachiara D'Angelo, Katty Hernández, I Macassi León, LI Meléndez López, and G Yáñez De la Borda.** 2010. "Women's Police Stations in Latin America: An Entry Point for Stopping Violence and Gaining Access to Justice." *Ottawa: Centre for Planning and Social Studies and International Development Research Centre*.
- Justino, Patricia.** 2011. "The impact of armed civil conflict on household welfare and policy responses." *Securing Peace: State-Building and Economic Development in Post-Conflict Countries*, 19.
- Kiesel, L, K Piescher, and J Edleson.** 2011. "Direct and indirect child exposure to violence: Effects on academic performance and disability." Mimeo.
- Kimmel, Jean.** 1998. "Child care costs as a barrier to employment for single and married mothers." *The Review of Economics and Statistics*, 80(2): 287–299.
- King, Russell, and Julie Vullnetari.** 2006. "Orphan pensioners and migrating grandparents: the impact of mass migration on older people in rural Albania." *Ageing & Society*, 26(5): 783–816.
- King, Russell, Diana Mata-Codesal, and Julie Vullnetari.** 2013. "Migration, development, gender and the 'black box' of remittances: Comparative findings from Albania and Ecuador." *Comparative Migration Studies*, 1(1): 69–96.

- Koenen, Karestan C, Terrie E Moffitt, Avshalom Caspi, Alan Taylor, and Shaun Purcell.** 2003. "Domestic violence is associated with environmental suppression of IQ in young children." *Development and psychopathology*, 15(2): 297–311.
- Kondylis, Florence.** 2008. "Agricultural outputs and conflict displacement: Evidence from a policy intervention in Rwanda." *Economic Development and Cultural Change*, 57(1): 31–66.
- Kondylis, Florence.** 2010. "Conflict displacement and labor market outcomes in post-war Bosnia and Herzegovina." *Journal of Development Economics*, 93(2): 235–248.
- Krüger, Jule, and Patrick Ball.** 2014. "Evaluation of the database of the Kosovo memory book." Last accessed December 2017. https://www.hrdag.org/wp-content/uploads/2015/04/Evaluation_of_the_Database_KMB-2014.pdf.
- Kuvendi Popullor, Albania.** 1976. "The Constitution of the Socialist Republic of Albania." Article 14: 28–41.
- La Ferrara, Eliana, Alberto Chong, and Suzanne Duryea.** 2012. "Soap operas and fertility: Evidence from Brazil." *American Economic Journal: Applied Economics*, 4(4): 1–31.
- Landman, Andreas, Helke Seitz, and Susan Steiner.** 2017. "Intergenerational Coresidence and Female Labor Supply." ZBW-German Central Library for Economics, Leibniz Information Center for Economics.
- Lee, Daniel.** 2002. "Fertility and female labor supply in rural China." Mimeo.
- Lehrer, Evelyn L.** 1992. "The impact of children on married women's labor supply: Black-white differentials revisited." *Journal of Human Resources*, 27(3): 422–444.
- Leon, Gianmarco.** 2012. "Civil conflict and human capital accumulation the long-term effects of political violence in Perú." *Journal of Human Resources*, 47(4): 991–1022.
- Londoño, Juan Luis, Rodrigo Guerrero, JL Londoño, A Gaviria, and R Guerrero.** 2000. "Violencia en América Latina: epidemiología y costos." *Asalto al desarrollo. Violencia en América Latina*, 11–57.
- Lundberg, Shelly J, Robert A Pollak, and Terence J Wales.** 1997. "Do husbands and wives pool their resources? Evidence from the United Kingdom child benefit." *Journal of Human resources*, 463–480.
- Lundborg, Petter, Erik Plug, Astrid Würtz Rasmussen, et al.** 2017. "Can Women Have Children and a Career? IV Evidence from IVF Treatments." *American Economic Review*, 107(6): 1611–1637.
- Manser, Marilyn, and Murray Brown.** 1980. "Marriage and household decision-making: A bargaining analysis." *International Economic Review*, 32–44.
- Martin, Susan F.** 2016. "Rethinking Protection of Those Displaced by Humanitarian Crises." *American Economic Review: Papers & Proceedings*, 106(5): 446–450.
- Maurer-Fazio, Margaret, Rachel Connelly, Lan Chen, and Lixin Tang.** 2011. "Childcare, eldercare, and labor force participation of married women in urban China, 1982–2000." *Journal of Human Resources*, 46(2): 261–294.

- McElroy, Marjorie B, and Mary Jean Horney.** 1981. "Nash-bargained household decisions: Toward a generalization of the theory of demand." *International Economic Review*, 333–349.
- Merrouche, Ouarda.** 2011. "The long term educational cost of war: evidence from landmine contamination in Cambodia." *The Journal of Development Studies*, 47(3): 399–416.
- Miguel, Edward, and Gerard Roland.** 2011. "The long-run impact of bombing Vietnam." *Journal of Development Economics*, 96(1): 1–15.
- Miller, Amalia R, and Carmit Segal.** 2016. "Do Female Officers Improve Law Enforcement Quality? Effects on Crime Reporting and Domestic Violence."
- MIMDES, Ministerio de la Mujer y Desarrollo Social.** 2007. "¿Que son los Centros de Emergencia Mujer?"
- MIMDES, Ministerio de la Mujer y Desarrollo Social.** 2009. "Investigacion operativa: "Eficacia de la intervencion de los Centros Emergencia Mujer"."
- Mitchell, Rhoda.** 2013. "Domestic violence prevention through the Constructing Violence-free Masculinities programme: an experience from Peru." *Gender & Development*, 21(1): 97–109.
- Morrison, A, MB Orlando, and G Pizzolitto.** 2007. "The impact of intimate partner violence against women in Peru: estimates using matching techniques." [Unpublished] 2007. Presented at the Population Association of America 2007 Annual Meeting New York New York March 29-31 2007.
- Nakamura, Alice, and Masao Nakamura.** 1992. "The econometrics of female labor supply and children." *Econometric Reviews*, 11(1): 1–71.
- Nizalova, Olena Y, and Irina Murtazashvili.** 2016. "Exogenous treatment and endogenous factors: Vanishing of omitted variable bias on the interaction term." *Journal of Econometric Methods*, 5(1): 71–77.
- Nunn, Nathan, and Nancy Qian.** 2014. "US food aid and civil conflict." *American Economic Review*, 104(6): 1630–1666.
- OSCE, Kosovo Verification Mission.** 1999. "KOSOVO/KOSOVA: As Seen, As Told. An analysis of the human rights findings of the OSCE Kosovo Verification Mission, October 1998 to June 1999." Volume I and II.
- Oster, Emily, and Bryce Millett Steinberg.** 2013. "Do IT service centers promote school enrollment? Evidence from India." *Journal of Development Economics*, 104: 123–135.
- Peek-Asa, Corinne, Leah Maxwell, Ann Stromquist, Paul Whitten, Mary Ann Limbos, and James Merchant.** 2007. "Does parental physical violence reduce children's standardized test score performance?" *Annals of epidemiology*, 17(11): 847–853.
- Perez-Heydrich, Carolina, Joshua L Warren, Clara R Burgert, and Michael Emch.** 2013. "Guidelines on the Use of DHS GPS Data."
- Perova, Elizaveta, and Sarah Anne Reynolds.** 2017. "Women's police stations and intimate partner violence: Evidence from Brazil." *Social Science & Medicine*, 174(188-196).

- Pistaferri, Luigi.** 1999. "Informal networks in the Italian labor market." *Giornale degli economisti e annali di economia*, 353–375.
- Pivovarova, Margarita, and Eik Leong Swee.** 2015. "Quantifying the microeconomic effects of war using panel data: Evidence from Nepal." *World Development*, 66(308–321).
- Pollak, Robert A.** 2004. "An intergenerational model of domestic violence." *Journal of Population Economics*, 17(2): 311–329.
- Porter, Maria, and Elizabeth M King.** 2012. "Fertility and women's labor force participation in developing countries." Mimeo. Department of Economics, University of Oxford.
- Posadas, Josefina, and Marian Vidal-Fernández.** 2013. "Grandparents' childcare and female labor force participation." *IZA Journal of Labor Policy*, 2(1): 14.
- Priebe, Jan.** 2010. "Child Costs and the Causal Effect of Fertility on Female Labor Supply: An investigation for Indonesia 1993–2008." Courant Research Centre: Poverty, Equity and Growth-Discussion Papers.
- Rahman, Mosfequr, Md Aminul Hoque, and Satoru Makinoda.** 2011. "Intimate partner violence against women: Is women empowerment a reducing factor? A study from a national Bangladeshi sample." *Journal of Family Violence*, 26(5): 411–420.
- Revilla, Ana Teresa.** 1999. "La administración de justicia informal en el Perú." *Reforma judicial en América Latina: una tarea inconclusa*, Santa Fé de Bogotá: Corporación Excelencia en la Justicia.
- Roodman, David.** 2017. "BOOTTEST: Stata module to provide fast execution of the wild bootstrap with null imposed." *Statistical Software Components*.
- Rosenzweig, Mark R, and Kenneth I Wolpin.** 1980. "Testing the quantity-quality fertility model: The use of twins as a natural experiment." *Econometrica: Journal of the Econometric Society*, 48(1): 227–240.
- Rosenzweig, Mark R, and Kenneth I Wolpin.** 1985. "Specific experience, household structure, and intergenerational transfers: farm family land and labor arrangements in developing countries." *The Quarterly Journal of Economics*, 100(Supplement): 961–987.
- Rosenzweig, Mark R, and Kenneth I Wolpin.** 2000. "Natural" natural experiments" in economics." *Journal of Economic Literature*, 38(4): 827–874.
- Ruiz, Isabel, and Carlos Vargas-Silva.** 2013. "The Economics of Forced Migration." *The Journal of Development Studies*, 49(6): 772–784.
- Sarvimäki, Matti, Roope Uusitalo, and Markus Jantti.** 2009. "Long-term effects of forced migration."
- Sasaki, Masaru.** 2002. "The causal effect of family structure on labor force participation among Japanese married women." *Journal of Human Resources*, 37(2): 429–440.
- Schultz, T Paul.** 1978. "The influence of fertility on labor supply of married women: simultaneous equation estimates." JAI Press Greenwich Conn. United States 1978.

- Schultz, T Paul.** 2009. "How does family planning promote development? Evidence from a social experiment in Matlab, Bangladesh, 1977-1996." Yale University, Economic Growth Center, New Haven, Conn.
- Shemyakina, Olga.** 2011. "The effect of armed conflict on accumulation of schooling: Results from Tajikistan." *Journal of Development Economics*, 95(2): 186–200.
- Shen, Ke, Ping Yan, and Yi Zeng.** 2016. "Coresidence with elderly parents and female labor supply in China." *Demographic Research*, 35(645).
- SOK, Statistical Office of Kosovo.** 2001. "Statistics on Education in Kosovo." Statistical Office of Kosovo.
- Ssewanyana, Sarah, Stephen Younger, and Ibrahim Kasirye.** 2007. "Poverty under conflict: The case for northern Uganda."
- Stark, Oded.** 2004. "On the economics of refugee flows." *Review of Development Economics*, 8(2): 325–329.
- Stevenson, Betsey, and Justin Wolfers.** 2006. "Bargaining in the shadow of the law: Divorce laws and family distress." *The Quarterly Journal of Economics*, 121(1): 267–288.
- Stock, James H, Jonathan H Wright, and Motohiro Yogo.** 2002. "A survey of weak instruments and weak identification in generalized method of moments." *Journal of Business & Economic Statistics*, 20(4): 518–529.
- Swee, Eik Leong.** 2015. "On war intensity and schooling attainment: The case of Bosnia and Herzegovina." *European Journal of Political Economy*, 40: 158–172.
- Tarifa, Fatos.** 1994. "Disappearing from Politics: Social Change and Women in Albania." *Women in the Politics of Postcommunist Eastern Europe*. Armonk, NY: ME Sharpe, Inc, 133–152.
- Thomas, Duncan.** 1990. "Intra-household resource allocation: An inferential approach." *Journal of human resources*, 635–664.
- Thompson, Richard, and Lakecia A Whimper.** 2010. "Exposure to family violence and reading level of early adolescents." *Journal of Aggression, Maltreatment & Trauma*, 19(7): 721–733.
- UNFPA.** 2012. "Sex imbalances at birth in Albania." United Nations Development Program.
- UNHCR.** 2010. "Peru: Domestic violence, state protection and support services available." *Canada: Immigration and Refugee Board of Canada*. Available at: <http://www.refworld.org/docid/4e438b532.html> [Last accessed 3 May 2017].
- UNHCR, The UN Refugee Agency.** 2017. "Global Trends: Forced Displacement in 2016." UNHCR official webpage. Last accessed January 2018 <http://www.unhcr.org/5943e8a34.pdf>.
- Valente, Christine.** 2013. "Education and civil conflict in Nepal." *The World Bank Economic Review*, 28(2): 354–383.

- Verwimp, Philip, and Juan Carlos Muñoz-Mora.** 2018. "Returning Home after Civil War: Food Security and Nutrition among Burundian Households." *The Journal of Development Studies*, 6(1019-1040).
- Westley, Karen, and Vladimir Mikhalev.** 2002. *The use of participatory methods for livelihood assessment in situations of political instability: A case study from Kosovo*. Overseas Development Institute London.
- WHO.** 2013. *Global and regional estimates of violence against women: prevalence and health effects of intimate partner violence and non-partner sexual violence*. World Health Organization.
- Wolfe, David A, Claire V Crooks, Vivien Lee, Alexandra McIntyre-Smith, and Peter G Jaffe.** 2003. "The effects of children's exposure to domestic violence: A meta-analysis and critique." *Clinical child and family psychology review*, 6(3): 171–187.
- Wong, Rebeca, and Ruth E Levine.** 1992. "The effect of household structure on women's economic activity and fertility: evidence from recent mothers in urban Mexico." *Economic Development and Cultural Change*, 41(1): 89–102.
- Zickel, Raymond E, and Walter R Iwaskiw.** 1994. "Albania: a Country Study." Federal Research Division, Library of Congress.