



HAL
open science

Economic impacts of climate change

Luis Guillermo Becerra-Valbuena

► **To cite this version:**

Luis Guillermo Becerra-Valbuena. Economic impacts of climate change. Economics and Finance. Université Panthéon-Sorbonne - Paris I, 2021. English. NNT : 2021PA01E044 . tel-03539624

HAL Id: tel-03539624

<https://theses.hal.science/tel-03539624v1>

Submitted on 21 Jan 2022

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.



UNIVERSITÉ PARIS 1 PANTHÉON-SORBONNE, UFR Economie

PARIS SCHOOL OF ECONOMICS

Research Institute: Paris-Jourdan Sciences Économiques

PHD THESIS

for the degree of Doctor of Philosophy in Economic Sciences

Presented and defended publicly at the Paris School of Economics on July 8th, 2021 by:

LUIS GUILLERMO BECERRA-VALBUENA

Economic Impacts of Climate Change

Under the direction of:

- Katrin MILLOCK, Senior Research Fellow CNRS-Paris School of Economics

Jury Composition:

<i>Reviewers:</i>	Olivier CHANEL	Senior Research Fellow CNRS, Aix-Marseille School of Economics
	Olivier DESCHÊNES	Professor at University of California, Santa Barbara
<i>Examiners:</i>	Raja CHAKIR	Director of Research INRAE-AgroParisTech
	Lise CHAUVET	Professor at Université Paris 1 Panthéon-Sorbonne
	Jérémie GIGNOUX	Research fellow INRAE, Paris School of Economics



UNIVERSITÉ PARIS 1 PANTHÉON-SORBONNE, UFR d'Économie
ÉCOLE D'ÉCONOMIE DE PARIS
Laboratoire de rattachement: Paris-Jourdan Sciences Économiques

THÈSE

pour l'obtention du titre de Docteur en Sciences Économiques

Présentée et soutenue publiquement à l'École d'Économie de Paris le 8 juillet 2021 par:

LUIS GUILLERMO BECERRA-VALBUENA

Impacts Économiques du Changement Climatique

Sous la direction de:

- Katrin MILLOCK, Directrice de recherche CNRS, École d'Économie de Paris

Composition du jury:

<i>Rapporteurs:</i>	Olivier CHANEL	Directeur de recherche CNRS, École d'Économie d'Aix-Marseille
	Olivier DESCHÊNES	Professeur, University of California, Santa Barbara
<i>Examineurs:</i>	Raja CHAKIR	Directrice de Recherche INRAE, AgroParisTech
	Lise CHAUVET	Professeure, Université Paris 1 Panthéon-Sorbonne
	Jérémy GIGNOUX	Chargé de recherche INRAE, École d'Économie de Paris

Contents

List of Figures	iv
List of Tables	vii
Acknowledgements	viii
Introduction	1
Résumé	14
1 Gendered Migration Responses to Drought in Malawi	29
1.1 Introduction	31
1.2 Background: migration and marriage patterns in Malawi	35
1.3 Empirical strategy	38
1.4 Data	40
1.4.1 Migration	40
1.4.2 Drought measures	42
1.4.3 Socio-economic characteristics and other covariates	44
1.4.4 Summary statistics	44
1.5 Results	44
1.6 Robustness tests	53
1.6.1 Marital status after migration corroborates the stated reason for migration	54
1.6.2 The influence of drought in the destination district	56
1.6.3 Household interaction effects	56
1.6.4 Results from a multinomial model	57
1.6.5 The role of ethnicity	58
1.7 Descriptive evidence on outcomes following migration	58
1.7.1 School attendance after migration	58
1.7.2 Do female migrants have higher probability of being pregnant after migrating?	60
1.8 Conclusion	62
Appendices	64
Appendices	65
1A Temperature in Malawi and SPI maps	65
1B Marginal Effects by socioeconomic characteristics	67
1C Timeline surveys and rainy seasons in Malawi	68
1D Kernel density of migration by age	69
1E District Classification	70

1F	Important events in the EA by survey-year	71
1G	Characteristics of migrants and Balance Tables	72
1H	Additional Robustness checks	74
1I	Droughts in the destination district	76
1J	Household interaction effects	78
1K	A multinomial estimation	79
1L	Measurement error in the variable of droughts	81
2	Climatic shocks, air quality, and health at birth in Bogotá	84
2.1	Introduction	86
2.2	The importance of health, air pollution and ENSO in Bogotá	90
2.3	Data and stylized facts	92
2.3.1	Descriptive statistics	96
2.4	Method and empirical strategy	99
2.4.1	Identification of health effects	99
2.4.2	Potential threats to identification	105
2.5	Results and robustness checks	108
2.5.1	Total ENSO effect on health	108
2.5.2	The Instrumental Variable approach: how ENSO affects health	112
2.5.3	Spatial variation and residential sorting	123
2.5.4	Restricting the sample to the South-West of the city, aggregating by localities and using a bigger buffer	126
2.5.5	Direction of the OLS bias	127
2.5.6	Heterogeneous effects, additional health outcomes and air pollution index	128
2.6	Conclusions	131
	Appendices	133
	Appendices	134
2A	Appendix Summary Traffic Policies - Wind Direction	134
2B	Identification of the effects on Weather and Pollution Hourly Analysis	135
2B.1	ONI Index	135
2B.2	Separating the effects of NIÑO and NIÑA	136
2B.3	Results of ONI-NIÑO-NIÑA on weather - Hourly Analysis	137
2B.4	Results of ONI-NIÑO-NIÑA on pollution - Hourly Analysis	139
2B.5	Results: ONI-NIÑO-NIÑA on weather	143
2B.6	Results: ONI-NIÑO-NIÑA on pollution	144
2C	Attrition in the sample	146
2D	Wind-direction over year	147
2E	Reduced form results IV on health	150
2F	IV 2 STAGE: moving average 7 days-maximum	153
2G	Variation in the spatial correlation over time and residential sorting	156
2H	Restricting sample to South-West of the city	160
2I	Aggregating the sample by localities of the city	165
2J	Changing buffer to 4km	170
2K	OLS estimations of pollution on Health - NO instruments	173
2L	Heterogeneity: mothers with tertiary education	174
2M	Effects on additional outcomes	177
2N	Index of pollution - Principal Component Analysis	179

3	Droughts and Agricultural Adaptation to Climate Change	180
3.1	Introduction	182
3.2	Data and Descriptive Statistics	186
3.2.1	ELCA 2010-2013-2016	186
3.2.2	Weather variables - The Standard Precipitation Index (SPI)	195
3.3	A theoretical framework for analysing adaptation of rural households in Colombia	198
3.4	Effects of climate and weather deviations on agricultural productivity	203
3.4.1	Identification Strategy for climate and yearly weather deviations	204
3.4.2	Results of climate and yearly weather deviations	206
3.4.3	Identification strategy for nonlinear effects of temperature	211
3.4.4	Results for nonlinear effects of temperature	213
3.5	Adaptation responses to drought impacts	218
3.5.1	Identification strategy using the DID approach and the SPI measure of droughts	219
3.5.2	Effects of drought on agricultural productivity in the short and medium term	221
3.5.3	Effects of consecutive droughts on agricultural productivity	224
3.5.4	How do households adapt?	226
3.5.5	Heterogeneous effects and alternative strategies	232
3.6	Conclusion	235
	Appendices	239
	Appendices	240
3A	Average temperature - municipalities of the sample	240
3B	Maps of SPI drought measures by municipality	242
3C	ELCA sample	244
3D	Test of climate and weather deviations with temperature squared	245
3E	Identification strategy and results using the SPI	246
3F	Balance tables Drought vs No Drought 2010	249
3G	Effects on migration and heterogeneity with credit	252
	Bibliography	265
	Summary	266
	Résumé	267

List of Figures

1.1	Average Temperature in Malawi	65
1.2	SPI-CHIRPS by district (scale 12-24)	66
1.3	Average marginal effects by different categories Migrating for Marriage	67
1.4	Time-line of Surveys and Rainy seasons. LSMS-ISA for MALAWI	68
1.5	Additional graphs-Kernel density of migration by age (Main reasons to migrate: 1 for work, 3 for marriage, 4 for family reasons)	69
2.1	ONI Index-ENSO events and Frequency	93
2.2	Localities and Monitoring Stations in Bogotá	95
2.3	Weather variables-24hour schedule	95
2.4	Map to assign weather and pollution to each health center	100
2.5	ENSO and Pollutants effect on Low Birth Weight	111
2.6	Wind Direction (WD) and Wind Speed (WS) on Pollutants	112
2.7	ONI-NIÑO-NIÑA on weather-24hours schedule	138
2.8	ONI-NIÑO-NIÑA on Pollutants ($PM_{10} - PM_{2.5}$)-24hours schedule	140
2.9	ONI-NIÑO-NIÑA on Pollutants ($NO_x - CO$)-24hour schedule	141
2.10	Type of attrition by month and source Percentage and Total	146
2.11	Categories of wind direction and wind speed per year	147
2.12	Categories of wind direction and wind speed per quarter	148
2.13	Map of wind direction (WD) and wind speed (WS) in the city	149
3.1	Coefficients estimated for the temperature bins - Average agricultural growing season	215
3.2	Average Temperature Colombia - ELCA municipalities	240
3.3	Droughts by year in the ELCA municipalities	240
3.4	Histogram Temperature Agricultural Growing Season	241
3.5	Time-line of ELCA surveys and El Niño phenomenon.	244

List of Tables

1.1	Reasons to migrate for the different type of migrants	36
1.2	Household head works in agricultural activities in the destination (by type of migrant)	37
1.3	Percentage of agricultural income in the household by person in destination (by type of migrant)	37
1.4	Total Number of Migrants by survey	41
1.5	Between and Within District Migration	46
1.6	Between and Within District Migration - Women	48
1.7	Between and Within District Migration - Women migrating for marriage (reason 3)	49
1.8	Between and Within District Migration (split sample) Women migrating for marriage (reason 3)	50
1.9	Between District Migration - Women and men migrating for work (reason 1)	51
1.10	Between District Migration - Women and men migrating for marriage (reason 3)	53
1.11	Between and Within District Migration (split sample) Women migrating for marriage (reason 3) and work (reason 1)	54
1.12	Marital status in destination- By category	55
1.13	Marital status in destination- By reason to migrate	55
1.14	DHS Malawi	58
1.15	Drop-out school in destination- By category	59
1.16	Drop-out school in destination - By reason to migrate	59
1.17	Drop-out school in destination- Years ago to migrate	60
1.18	Have a pregnancy in the last 24 months in destination - By category	61
1.19	Have a pregnancy in the last 24 months in destination - By reason to migrate	61
1.20	Have a pregnancy in the last 24 months in destination- Years ago to migrate	62
1.21	District Classification for the LSMS-ISA	70
1.22	Most important events that happened in the EA-by survey	71
1.23	Percentage of transfers in the household by person in destination (by type of migrant)	72
1.24	Percentage of no agricultural income (no agricultural wage and self-employment) in the household by person in destination (by type of migrant)	72
1.25	Balance Tables for droughts at origin-district using SPI measure	73
1.26	Within District Migration - Women and Men migrating for work (reason 1)	74
1.27	Within District Migration - Women and men migrating for marriage (reason 3)	75
1.28	$Drought_o$ and $Drought_d$ together. Women migrating for marriage (reason 3) Between District	76
1.29	Adding $dist_d \times year$ fixed effects. Between and Within District Migration Women migrating for marriage (reason 3)	77

1.30	Including control for whether household head and partner migrated together Between and Within District Migration - Women migrating for marriage (reason 3)	78
1.31	Multinomial Logit: No migrants and Between-Within Migrants - Women migrating for marriage (reason 3)	79
2.1	Descriptive Statistics for children (health outcomes and pollution)	97
2.2	Descriptive statistics for socio-economic variables, weather, ENSO events and other controls	98
2.3	REDUCED FORM: ENSO EFFECTS ON HEALTH	109
2.4	IV FIRST STAGE: IV ON SO_2 PM_{10} $PM_{2.5}$ (average)	115
2.5	IV FIRST STAGE: IV ON CO NO_X (peak afternoon)	116
2.6	IV SECOND STAGE: SO_2 (average) ON HEALTH	118
2.7	IV SECOND STAGE: PM_{10} - $PM_{2.5}$ (average) ON HEALTH	120
2.8	IV SECOND STAGE: CO - NO_X (peak afternoon) ON HEALTH	122
2.9	Traffic Policies	134
2.10	Wind Direction (Categories)	134
2.11	OLS Estimations on Weather Variables (24 hours schedule)	143
2.12	OLS Estimations on Pollutants (PM_{10} - $PM_{2.5}$) (24 hours schedule)	144
2.13	OLS Estimations on Pollutants (NO_x - CO) (24 hours schedule)	145
2.14	REDUCED FORM: IV ON HEALTH SO_2 (average)	150
2.15	REDUCED FORM: IV ON HEALTH PM_{10} $PM_{2.5}$ (average)	151
2.16	REDUCED FORM: IV ON HEALTH CO - NO_X (peak afternoon)	152
2.17	IV SECOND STAGE: SO_2 (Ave. MA7) ON HEALTH	153
2.18	IV SECOND STAGE: PM_{10} - $PM_{2.5}$ (Ave. MA7) ON HEALTH	154
2.19	IV SECOND STAGE: CO - NO_X (Max.) ON HEALTH	155
2.20	Moran Index for main pollutants and weather variables (Yearly)	156
2.21	OLS Estimations of NIÑO-NIÑA and Instrumental Variables (Wind direction and Wind Speed) on socioeconomic characteristics	157
2.22	Socioeconomic characteristics from the EMB - Migration	158
2.23	Socioeconomic characteristics from the EMB	159
2.24	IV SECOND STAGE: SO_2 (average) ON HEALTH	160
2.25	IV SECOND STAGE: PM_{10} - $PM_{2.5}$ (average) ON HEALTH	161
2.26	IV SECOND STAGE: CO - NO_X (peak afternoon) ON HEALTH	162
2.27	IV FIRST STAGE: IV ON SO_2 PM_{10} $PM_{2.5}$ (average)	163
2.28	IV FIRST STAGE: IV ON CO NO_X (peak afternoon)	164
2.29	IV SECOND STAGE: SO_2 (average) ON HEALTH	165
2.30	IV SECOND STAGE: PM_{10} - $PM_{2.5}$ (average) ON HEALTH	166
2.31	IV SECOND STAGE: CO - NO_X (peak afternoon) ON HEALTH	167
2.32	IV FIRST STAGE: IV ON SO_2 PM_{10} $PM_{2.5}$ (average)	168
2.33	IV FIRST STAGE: IV ON CO NO_X (peak afternoon)	169
2.34	IV SECOND STAGE: SO_2 (average) ON HEALTH	170
2.35	IV SECOND STAGE: PM_{10} - $PM_{2.5}$ (average) ON HEALTH	171
2.36	IV SECOND STAGE: CO - NO_X (peak afternoon) ON HEALTH	172
2.37	OLS: SO_2 - PM_{10} (average) CO - NO_X (peak afternoon) ON HEALTH	173
2.38	IV SECOND STAGE: SO_2 (average) ON HEALTH	174
2.39	IV SECOND STAGE: PM_{10} $PM_{2.5}$ (average) ON HEALTH	175
2.40	IV SECOND STAGE: CO NO_X (peak afternoon) ON HEALTH	176
2.41	IV SECOND STAGE: SO_2 - PM_{10} (average) ON HEALTH	177

2.42	IV SECOND STAGE: $CO-NO_X$ (peak afternoon) ON HEALTH	178
2.43	IV SECOND STAGE: PCA-A and PCA-B ON HEALTH	179
3.1	Outcome variables rural households - Descriptive Statistics	190
3.2	Controls and weather variables rural households - Descriptive Statistics	191
3.3	Values of gross agricultural productivity by municipality	193
3.4	Climate and Yearly weather deviations	210
3.5	Climate and Yearly weather deviations- Maximum monthly Temperature- Agricultural Growing Season	211
3.6	GDD and HDD for temperature	217
3.7	GDD and HDD for temperature- Maximum monthly Temperature- Agricultural Growing Season	218
3.8	DID for 2010-2013-2016 using Drought in 2010 on main agricultural outcomes	223
3.9	DID for 2010-2013-2016 using the consecutive Drought 2010 and 2013 on main agricultural outcomes	225
3.10	DID for Drought in 2010 by type on investment	229
3.11	DID for Drought in 2010 by type of land use	231
3.12	SPI variables by ELCA year - Average	241
3.13	ELCA sample by municipality and year	244
3.14	Climate and Yearly weather deviations	245
3.15	Impact of SPI on main agricultural outcomes	248
3.16	Socioeconomic variables 2010 - Balance tables	249
3.17	Agricultural variables in 2010 - Drought vs No Drought in 2010	250
3.18	DID for 2010-2013-2016 for type of movement in the municipality	252
3.19	DID for Drought in 2010 on main agricultural outcomes - Heterogeneity having credit	253

Acknowledgements

Throughout the writing of this dissertation I have received a great deal of support and assistance.

I would first like to thank my supervisor, Professor Katrin Millock, whose expertise was invaluable through the whole thesis. Your insightful feedback pushed me to sharpen my thinking and brought my work to a higher level; also helping me to refine the research questions, and to improve the quality of my research. I am not less but very grateful for your unconditional support and the trust deposited on me to pursue my PhD and finish the thesis. In spite of not having a full financial support at the beginning of my studies, she helped me to get, year by year, the financial assistance and complete this thesis without struggling in financial terms. My words are short compared to the whole support I have received during these years from her. I still remember the first day I met her at the Maison des Sciences Économiques (MSE), and telling to my friends that I would be delighted to be supervised by her. With no doubt, I think she is the right choice as a wonderful supervisor. I would like to thank her for all her patience, for her fruitful feedback and sorry for the many mistakes in my written English. I hope to have improved it a lot thanks to her.

I would also like to express my immense gratitude to all the jury members of the thesis. Olivier Chanel and Olivier Deschênes, as reviewers, for your constructive comments and encouragement to make my research much better, during the defense and before, when we met at Aix-Marseille and in Santa Barbara-UCSB. And Raja Chakir, Lisa Chauvet and Jérémie Gignoux, as examiners, for being part of my jury and taking your time to read carefully my work. I feel very fortunate for having the opportunity to have such a high quality jury. Your generous and insightful comments enabled me, without any doubt, to improve my dissertation, have insightful papers and think further

about my research.

I have not more than gratitude with Liam and with Kenneth. I really enjoyed and learn a lot working with you during my master thesis and in our project in Benin. I hope we can continue working together in new projects. Liam, thank you so much for the useful comments, and advice during these years. I will never forget the trust you put on me and thanks to that, I could continue my studies and finance the beginning of my PhD. Also to Sylvie, for accepting me in the PhD program despite of not having financial support for three years and to Véronique, Pauline M. and Jean Christophe and Jean-Philippe for your help at PSE and at the Doctoral School ED465. Also to the mentor's program CITE-CIUP (Cité Universitaire Paris) and the people of the Eiffel Scholarship. A special thanks to Thomas Picard, whose support to renew the scholarship CITE-CIUP, made me possible to start my PhD and this thesis. A feeling of gratefulness I also have with the people of the group of Regulation and Environment at PSE and the Institut Convergences Migrations, for the useful feedback during the meetings and conferences.

A special thanks to all the PSE friends and colleagues at PSE and the Paris 1-Panthéon-Sorbonne. Jaime Montaña for his valuable feedback, and for showing me to appreciate more life despite the difficult times you had. I always admire you for that. Jaime Leyva for his countless encouragement and experiences. Mónica for being an unconditional friend and for all your help during so many years. Emmanuelle, Ilya, Irene, Héctor and Nhung, for all your support and countless talks about life, papers and the many drinks we had together. Those are valuable memories I will always remember. Florencia, Santiago, Juan, Christian, Dave, Max, Álvaro, Zeineb, Yaz, Paolo, Antton, Sara, Marion, Paul B., Oscar, Cristian F, Emmanuel, Cem, Caroline, Pauline C. and Yousseff. You made my time during the PhD more enjoyable at PSE and in Paris, and I feel very fortunate for meeting you and had so many great experiences across these years. Thanks Sofia for giving me a hand with my bad written French and for all the good moments. The good Martin for the so many years of enrichment, feedback and for your friendship. Thanks Juni for the countless talks about life and future during the thesis. And too all the guys of the Master PPD at PSE. You helped me to start this long journey and you took part of my heart for all the great moments we had. All of you provided me with stimulating discussions as well as happy distractions to rest my mind outside

of my research. Thanks also to those whom I might have forgotten and to those whom I have met briefly in the PSE community. This achievement could have not being finished without any of you.

I would like to acknowledge my colleagues from my exchange at the Los Andes University and at the UCSB-EMLAB and the Bren School of Environmental Science and Management. Although I could only stay for brief period of time, their wonderful collaboration and advice made possible the improvement of many parts of my thesis. Special thanks to Juliana H. and Jaime R. during my time at the UCSB and Miguel and Anghella at the Los Andes University. I am not more than grateful to professor Kyle Meng for hosting me at UCSB and to professor Jorge Bonilla for all his advice, that made my thesis a better product. I am delighted to work with Jorge, to whom I owe a good part of my second chapter.

A brief acknowledge to my friends from the alma mater in Colombia, UNAL, JJ, Deison, Maira, Natalia, Anderson, Cristian, Mario, Pedro, Norma, Liceth, Julian M. and those I met at the IDB and DC, Ana V, Nelson, Cristian S. Jaime U., Sebastián M., Daniel A., Cristina C., Fran, only to mention a few. A special thanks to Rosa for her friendship and all her help during the times of financial struggle I had in Paris, when I started this thesis. Also to Diego R, Polo, Elber, Cristian and Carmen, and all the guys who gave me support during these years. Thanks for your countless experiences and moments and for your encouragement to finish this thesis.

I am not more than thankful with my parents and my sister for their wise counsel and sympathetic ear, specially to my mother to whom I dedicate this thesis. Warmest thanks to Lupe and Gérard who received me in their home as another family member and for helping me to improve my French and immerse in the French culture. You have became my family in France. A special thanks to Nastya, my darling. Thanks for supporting me during the last years of my thesis and for always being there when I needed a voice to keep going. In spite of not understanding the things of what I was doing, your unconditional support and truly feelings was all I needed. Part of this is our achievement.

Introduction

This dissertation gives evidence of different economic impacts of climate variability and climatic phenomena in low and middle income countries. Climate change affects many aspects of human life, not only because it distorts weather patterns, but because it increases the frequency of extreme events and the intensity of phenomena such as the El Niño Southern Oscillation (ENSO events or El Niño-La Niña events),¹ which might induce individuals to migrate, affect agricultural outcomes or alter individuals' health, among several other consequences. With respect to migration, a large literature documents how populations affected by drought or other adverse climatic shocks may turn to migration as a means to cope (Berlemann and Steinhardt, 2017; Beine and Jeusette, 2018; Cattaneo et al., 2019; Hoffmann et al., 2020). Another strand of the literature has studied the negative impacts of climate change and extreme temperature on newborns and mortality (Deschênes et al., 2009; Barreca and Schaller, 2020; Barreca et al., 2015). In terms of agriculture, D'Agostino and Schlenker (2016) consider that climate change increases the volatility of food production, which can reverse the gains in average yields obtained with the Green Revolution. Just recently, Ortiz-Bobea et al. (2021) found that anthropogenic climate change reduced global agricultural total factor productivity (TFP) by about 21% since 1961, equivalent to losing the last 7 years of productivity growth. The damaging effect of climate change on TFP has not been not the same across regions, affecting more severely warmer regions such as Africa and Latin America and the Caribbean (a reduction of around 26% to 34%). In this sense, the dissertation gathers evidence on three particular economic impacts of climate change: first, on the climate-induced migration related to droughts; second, on how climatic phenomena such as El Niño and La Niña (ENSO) events can affect infant

¹It is predicted that the intensity and the frequency of phenomena such as El Niño will increase in the coming years (see Chapter 3 of the report "Global Warming of 1.5 Celsius degrees" of the Intergovernmental Panel of Climate Change and Wang et al. (2017))

health at birth in the context of air pollution and health; and third, on the adaptation mechanisms in agriculture when farmers face droughts and climatic variability.

The dissertation uses data from two particular countries, Malawi and Colombia, which have experienced frequent extreme events such as droughts. Additionally, they have very rich data availability that can help to assess the diverse ways in which climate change manifests. In fact, this allows to make a more profound analysis of climatic variability in rural and urban contexts. This is an advantage that permits to explore and analyze several economic aspects, but also serves as evidence of the magnitude of how climate change threatens human life. In the rural context of Malawi, for example, droughts impact agricultural activities and to the extent that rural households cannot adapt on the intensive margin (buying insurance or drought-resistant seeds), migration becomes an alternative (extensive margin). Here, drought acts as a push factor driving different gendered migration responses (for work and for marriage), and the age differences should be taken into account. For a middle income country like Colombia, the rural context and agricultural characteristics of the markets are different. The third chapter shows in fact that rural households in Colombia respond to droughts by increasing the total area planted in crops, but that this did not increase their agricultural productivity. Most likely, the type of rural markets in Colombia are different to the ones found in Malawi, the level of income is higher, and rural households diversify more by implementing crop-mixing, among other different factors. Even if there is evidence of drought induced migration in Colombia, the marriage institutions are not as important as in the case of Malawi (see the discussion in the first chapter) and child migration would happen mainly for work-related motives in Colombia. Additionally, particular phenomena as El Niño bring dry periods in the case of Colombia, creating extra distortions in the agricultural markets. The second chapter of the dissertation shows that ENSO events such as El Niño affect directly and indirectly the health of children during the time in utero, once the effect of air pollution on health is taken into account. We argue that part of the distortions of ENSO events happen through the general equilibrium effect it creates on agricultural markets and food prices. This discussion of the results of the dissertation shows not only the different interrelations in rural and urban contexts of climatic variability, but also, that the relationships can be very complex across different domains, not very straightforward and context-dependent.

Another important characteristic in common across the different chapters is the merging of remote sensing data (satellite images) with information of surveys and administrative data. The use of objective weather data in a standardized manner across locations has been discussed by Dell et al. (2012), in the agricultural studies of climate change adaptation (Deschênes and Greenstone, 2011; Schlenker and Roberts, 2009) and in the literature on migration by Millock (2015). The use of external sources allows to construct exogenous variables of weather and reduce the potential noise and measurement errors from self-reported extreme events in survey data. In this regard, this dissertation takes advantage of the explosion of the data available through satellite images, which have started to be more frequently used in several fields such as computer sciences, geography and economics (see Donaldson and Storeygard (2016) for a broader discussion on the use of remote sensing data in economics). On the one hand, chapters one and three use raw satellite images to construct standardized precipitation indices and define more adequate measures of droughts for the studies of migration in Malawi and agriculture in Colombia. On the other hand, the second chapter relies on data-sets constructed using monthly satellite images built on the spatial information of pollution and weather data of monitoring stations in Bogotá. The three chapters thus display a sample of the different uses that geographical tools can provide to economists, and more applications in this aspect are expected in the near future.

The dissertation contributes to the economic literature in different aspects. The first chapter merges the literature of climate-induced migration (Gray and Mueller, 2012; Hirvonen, 2016; Kubik and Maurel, 2016) with the literature of marriage institutions (Rosenzweig and Stark, 1989; Mobarak et al., 2013; Corno et al., 2020). In terms of the research question, we explore if there are differences in migratory responses following drought according to gender, age and migration motives, particularly at young age. The chapter adds evidence on the heterogeneous effects of climatic extreme events on migration (Gray and Mueller, 2012; Dillon et al., 2011; Baez et al., 2017a), gender differences and how marriage-related institutions affect such migration, which have not been explored yet. The second chapter links the literature of air pollution on health (Graff Zivin and Neidell, 2013; Lavaine and Neidell, 2017; Arceo et al., 2015) with the literature of extreme temperatures on health (Deschênes and Greenstone, 2011; Barreca et al., 2015), adding the effect of ENSO events on health, which has not been explored in the literature. Unlike other studies, the chapter

jointly investigates if ENSO events, local weather and air pollution affect urban health, particularly of newborns. ENSO events might impact health through different channels. They can impact local weather and then, health; ENSO events can affect pollution and then, health; they could also have a more direct effect on health and finally, they could also affect agriculture, food markets and food prices and generate loss of household income. In this regard, the impacts of air pollution on health can be seen as separate from the shocks that are mediated through ENSO events. The third chapter adds to the literature on responses to climate change in agriculture (Blanc and Schlenker, 2017; Burke and Emerick, 2016; Deschênes and Greenstone, 2007; Schlenker and Roberts, 2009), contributing to the scarce evidence in low and middle income countries, and proposing an alternative measure to evaluate climate impacts on agriculture. I answer the question of how droughts affect agricultural decisions and food production of rural households in Colombia in the short-term and in the medium-term and to what extent rural households adapt to these shocks. I also explore if there is any heterogeneity between households living in high versus low temperature areas, an aspect less explored in the literature. Not less important, the chapter gives evidence of the climate impacts outside of the common staple crop domain studied in the literature (maize, wheat, rice and soybeans) by exploring a broader set of crops and by analysing other farm inputs such as labor, the use of investments and access to water. As pointed out by Hertel and de Lima (2020), the literature should move beyond the yield impact (particularly from the common staple crops) where we have better data and models, and move towards other food products, farm inputs and nutritional impacts.

The first chapter is entitled "Gendered migration responses to droughts in Malawi" and written with Katrin Millock. We study two aspects of the literature: the marriage patterns according to bride price customs and the patterns of migration following drought, exploring its impacts according to sex, age and stated reasons to move of the migrants. Malawi is a country with a sizeable number of migrants for marriage and for work, and it suffers recurrent droughts that affect agricultural production and the population's food security. The chapter relates different aspects of the literature in the following way. On the one hand, there are already studies that show changes in migration patterns following drought (Gray and Mueller, 2012; Hirvonen, 2016; Kubik and Maurel, 2016), but there are still gaps in the knowledge of how gender affects climate-induced migration (Chindarkar, 2012), and how gender responses vary according to cultural contexts. While female migration is

constrained by temperature variability or droughts in Ethiopia and Nigeria (Dillon et al., 2011; Gray and Mueller, 2012), women migrate more following higher temperatures or drought in some Asian and Caribbean countries (Mueller et al., 2014; Thiede and Gray, 2017; Baez et al., 2017a,b). On the other hand, a recent literature in development economics addresses the role of marriage institutions on socioeconomic outcomes (Corno and Voena, 2017; Ashraf et al., 2020). As such, the bride price is money or goods paid by a man to a girl's family to secure a bride and it is more common in sub-Saharan African countries like Malawi. In other countries like India, it is the bride's family who gives money to the bridegroom's family (dowry) (Lowes and Nunn, 2018). In this regard, Corno et al. (2020) show that in countries with a culture of bride prices, adverse shocks during teenage years can increase the likelihood of early marriage (see Corno and Voena (2017) for the case of Tanzania); however, in countries with a culture of dowry payments, such adverse shocks can decrease the likelihood of early marriage. In this sense, the chapter provides evidence of the impact of droughts on gender in a context of bride price customs, particularly on marriage-related migration and work-related migration at different ages.

The chapter analyses the impact of drought on women's and men's migration, by constructing a retrospective panel on a yearly basis (a panel for individual migrant i from 2000-2016 that moves at time t). It combines these data with measures of droughts based on precipitation and temperature from satellite images. Relying on a quasi-experimental design, the chapter compares the probability of an individual in a particular age group migrating from an origin district to a destination district in a particular year, with or without drought. First, we explore the changes in the probability of migrating for marriage for girls aged 10 to 17 years old. This can provide evidence of gender differences in climate-induced migration and whether droughts can induce child marriages in a country with bride price customs. Second, we compare the results with men's work-related migration and women's migration for marriage to explore the differences further. In terms of the mechanism, since droughts affect crop yields and the income of families is dependent on agriculture, sending family members to live with other relatives or anticipating the marriage of daughters during those shocks could be a way to smooth consumption of the family.

Our results show rather small increases in the probability of women's within-district migration

because of marriage following moderate drought (a few percentage points), and we do not find sufficient evidence to reject the null hypothesis of drought not affecting the probability of child marriages. However, following severe droughts there is an increase in the probability of between-district migration for work-related reasons for boys aged 10 to 17 years old (9 percentage points), and to a lesser extent for girls (5 percentage points). Only when we combine both marriage-and work-related migration reasons for girls, we find an increase of 2.3 percentage points in migration following severe drought. The chapter makes a contribution in four ways. First, by analysing marriage-related migration and the timing of drought in the origin location of the individual, as evidence on the impact of drought on migration for marriage. Second, with respect to the literature on climate-induced migration, it analyses drought as a push factor driving different types of migration in Malawi, not only for work-related motives, but also for marriage as a way to smooth consumption of the family during droughts. Third, the analysis uses detailed information at the origin and destination districts over more than one and a half decade, exploiting the timing of the drought, unlike previous studies on Malawi. And fourth, it explores whether the type of migration changes depending on how far migrants move, comparing migration within the district of origin with migration between districts.

The second chapter "Climatic shocks, air quality, and health at birth in Bogotá" is written with Jorge Bonilla. The study, unlike any other previously, jointly investigates the effects of pollution, ENSO events and local weather on health. We examine this for the city of Bogotá for all the newborns over the period 1998-2015. The city is one of the largest urban centers in Latin America and it is exposed frequently to ENSO events and high levels of pollution. In terms of pollution, it reaches levels of cities like Delhi. According to the Real-time Air Quality Index from the World Air Quality Index Project, Bogotá reached a value of 141 in the index on March 8th of 2019 (very harmful for the individuals). For the same day, the index had a value of 151 (unhealthy) for Delhi, while it had a value of 23 (good) in Paris. The health costs of air pollution are not deniable. According to the report in 2017 of the National Department of Planning [DNP \(2017\)](#), the environmental degradation caused 10.5% (3,219) of the total deaths in Bogotá, with an estimated cost of 4.2 billion of Colombian pesos (2.5% of the city's GDP). However, this report does not consider the effects of ENSO on health. In fact, we estimate that ENSO effects on birth weight are several times larger than the impacts coming from pollution. Our analysis thus provides evidence that the health costs

from ENSO events can also be considerable.

Although the effects of air pollution on health have already been studied in the literature using similar instrumental variables like the ones we use (wind direction), the impact of ENSO events on health has not been explored yet. Our approach relies on papers such as [Hsiang et al. \(2011\)](#) and [Dingel et al. \(2020\)](#), where ENSO events are considered as global effects, disturbing not only weather factors, but also creating disturbances on food prices, and other general equilibrium effects that can alter household consumption and, in that manner, pregnant mothers and infant health. Our argument is that ENSO not only manifests itself as an extreme climatic shock that influences weather, but it can also have an impact on agriculture and disasters and hence, it might induce changes in food markets and affect household income. Therefore, the effect of pollution on health can be seen as separate from the shocks mediated through ENSO events. In addition, ENSO events can also affect air pollution levels, influencing health outcomes indirectly. In this aspect, the chapter links the literature of air pollution on health ([Graff Zivin and Neidell, 2013](#); [Lavaine and Neidell, 2017](#); [Arceo et al., 2015](#)) with the literature on extreme temperatures on health ([Barreca et al., 2015](#); [Deschênes et al., 2009](#); [Deschênes and Greenstone, 2011](#)), by adding the impact coming from ENSO events (El Niño and La Niña).

For the analysis, we merge pollution and weather data from monitoring stations with data of all the children born in Bogotá from 1998 to 2015. We focus on children during gestation in order to capture exposure to ENSO more instantaneously and to avoid any bias in the estimations created by not controlling for unobserved factors more common in adults (smoking, eating healthy food, exercising, etc.). As assigning air pollution and weather to health outcomes is not straightforward, we start by using the method of inverse distance interpolation to create a monthly set of maps of pollution and weather variables in a GIS software. As the mothers' addresses are not available, our best approach is to create a running-quarter panel aggregated by health centers, where the quarter corresponds to the trimester of gestation for the infants born in a specific month in the health center. As a final step, we construct a 2km buffer around the health center's address where the children were delivered, and assign the pollution and weather variables inside the buffer. The identification strategy relies on a two stage approach, where the first stage instruments the equation

of pollution by wind direction; the second stage establishes the relation of pollution and ENSO on health, controlling for weather factors, among other controls. Importantly, the system has the advantage of capturing the lagged effect on health of the exposure to pollution and ENSO during the three quarters of gestation. While the identification of the effects of air pollution on health relies on cross-section and time variation, the effect of ENSO events on health outcomes at birth relies entirely on the variation over time.

As a main result, across all specifications ENSO exposure decreases weight at birth and increases the likelihood of being low weight at birth, after including the channels of pollution and classical local weather. When estimations involve SO_2 , PM_{10} and $PM_{2.5}$, El Niño and La Niña events reduce birth weight when exposure happens in the last quarter of gestation, the period in which the mother is close to deliver. In the estimations with CO and NO_X , the separate effects of ENSO show that El Niño (in the third and first quarter) and La Niña (in the second quarter) increase the likelihood of having low weight at birth, and reduce the weight in grams at birth. Independently of the pollutant used, the cumulative effect of ENSO across gestational quarters are more often statistically different from zero at conventional significance levels for El Niño, but less for La Niña episodes. As a secondary result, we confirm the negative impact of air pollution on health. We find that exposure in uterus to average SO_2 concentrations in the first and third quarter of gestation influences health outcomes at birth, while $PM_{2.5}$ exposure during the third quarter of gestation increases the probability of being premature. In addition, PM_{10} appears to decrease birth weight when exposure occurs in the second quarter of gestation, while CO exposure (in the second and particularly, in the third quarter) is found to increase the probability of having low birth weight. When analysing the cumulative effect during the total gestational length, SO_2 and CO are the only pollutants that have a negative impact on birth weight at conventional significance levels. Moreover, the size of the impact of ENSO events is much larger than the consequences of pollution exposure. Being exposed to ENSO may decrease birth weight up to 1.3%, while an increase of 1 *ppb* of SO_2 or 1 $\mu g/m^3$ of PM_{25} might reduce birth weight up to 0.3% or 0.14%, respectively. Our paper thus sheds light on the magnitude of the estimate of the overall effect of extreme climatic shocks on health at birth via ENSO.

"Droughts and agricultural adaptation to climate change" corresponds to the third chapter of the dissertation. I analyse how droughts affect agricultural decisions and food production of rural households in Colombia, in the short-term and in the medium-term. The objective is to measure to what extent rural households adapt to these shocks and if there are differences in rural households' behaviour in the short-term versus in the medium-term. The analysis also focuses on an important aspect of heterogeneity: the differences between rural households located in municipalities with high average temperature and those located in municipalities with low average temperature. Colombia is a country with recurrent droughts and albeit only 6% of the GDP comes from agriculture, the sector accounts for 17.3% of total employment in June 2020. The country is also frequently affected by El Niño phenomena, which bring droughts in the case of Colombia. To put in context the impact of this phenomenon, according to the country's Environment ministry-IDEAM, the El Niño phenomenon might have reduced Colombia's rainfall by 80% in the first months of 2019. In this respect, rural households might try to cope with droughts on the intensive margin, by getting insurance or credits, smoothing consumption, mixing crops or postponing investment among others ways. Rural households might also try to adapt on the extensive margin by migrating.

Although there are several studies on climate change and agriculture, a vast literature has focused on the US (Deschênes and Greenstone, 2007; Schlenker and Roberts, 2009; Burke and Emerick, 2016), and less is known of the impact of climate change on agriculture for less developed countries. In terms of rural households' adaptation Burke and Emerick (2016) for US and Costinot et al. (2016) from a global perspective show that crop switching could be a possible response to climate change. Rural households might also react to these shocks by making investments in the unit of agricultural production, trying to get technical assistance or modifying the use of fertilizers. Similarly, Rosenzweig and Wolpin (1993) also highlight that households could respond to droughts on the intensive margin by increasing off-farm work (for the case of India see Jayachandran (2006)), selling cattle (see Fafchamps et al. (1998) for the case of Burkina Faso), or on the extensive margin, through migration (see Cattaneo et al. (2019) for a global perspective). This chapter makes a contribution in four ways. First, it adds to the scarce literature in low and middle income countries, and propose an alternative measure to evaluate climate impact and adaptation in agriculture in the short-term and in the medium-term in Colombia. Second, it explores the heterogeneity between

households living in high versus low temperature areas, an aspect less explored in the literature. Third, it shows that in the short-term and medium-term, rural households in Colombia adapt by using more available land, which could lead to higher gross agricultural productivity. Another novelty is to analyse consecutive droughts. Fourth, it explores in more detail how rural households adapt to droughts by analysing a broader set of crops: vegetables, fruits, cereals and coffee. As pointed out by [Hertel and de Lima \(2020\)](#), the FAO identifies 175 distinct crops but the majority of climate impact studies have focused on changes in yields for four staple crops: maize, wheat, rice and soybeans. These four staple crops account for only one quarter of the total value of agricultural output. The article gives evidence of climate impacts outside of the staple crop domain and explores other farm inputs such as labor, the use of investments and access to water. According to [Hertel and de Lima \(2020\)](#), the literature should move beyond the yield impact where we have better data and models, and move towards other food products, farm inputs and nutritional impacts.

The third chapter uses the Colombian panel survey (ELCA), which has very rich information on agricultural production and investments made in the land of rural households. The chapter merges the three waves of the ELCA (2010-2013-2016) with weather data from satellite images in order to avoid the measurement errors that could arise from using subjective self-reported shocks. As the ELCA panel follows the same household over three waves, it permits to study adaptation such as using more land on some crops than others or changes towards other types of agricultural activities over time. The analysis of the chapter can be divided into three parts: first, it uses the identification framework used in the literature that separate the effect of climate from yearly weather deviations ([Kelly et al., 2005](#); [Deschênes and Kolstad, 2011](#); [Burke and Emerick, 2016](#); [Bento et al., 2020](#)) for the Colombian rural households; the second part follows [Deschênes and Greenstone \(2007\)](#), [Schlenker and Roberts \(2009\)](#) and [Aragón et al. \(2021\)](#) to explore different adaptive margins in response to extreme temperature, using the standard measures of Growing Degree Days (GDD) and Harmful Degree Days (HDD); and third, it proposes to use the Standard Precipitation Index (SPI) as a measure of drought to evaluate the effects of extreme weather in agriculture. This alternative measure is used in a Differences-in-Differences (DID) framework to see how small rural households in Colombia adapt. The DID analyses the short-term and then, the medium-term adaptation.

The main findings of the chapter are five. First, with respect to the analysis of climate and yearly weather deviations, gross agricultural productivity in value terms is positively affected by the long-run mean temperature, while there is no effect from temperature shocks (deviations from the long-run mean); with respect to the marginal effects, higher temperature affects positively gross agricultural productivity in value terms in low-temperature municipalities but negatively in high-temperature municipalities. With the exception of [Aragón et al. \(2021\)](#), the heterogeneous effects of temperature have been explored for the case of U.S. ([Deschênes and Greenstone, 2007](#); [Schlenker and Roberts, 2009](#)) but less for developing countries. Second, across several of the analyses of the short-term and medium-term (climate-yearly weather deviations, GDD-HDD and the DID), the gross agricultural productivity in value terms of cereals and coffee benefits from droughts or higher temperatures. The gross agricultural productivity increases in low-temperature municipalities for cereals and coffee, while those in high-temperature areas are negatively affected. Vegetables seem to be more affected following droughts (DID analysis). Across the different analyses, the gross agricultural productivity in value terms of vegetables and fruits displays reductions, particularly in low-temperature municipalities (with increases in high-temperature municipalities). Third, with regard to the short-term and the medium-term adaptation, the drought of 2010 made the rural households re-allocate land by reducing the type of land left fallow and assigning it to production of crops and for livestock. This captures some of the trade-offs proposed in the theoretical framework. It also goes in line with the increase of farm inputs (land and labor) found in [Aragón et al. \(2021\)](#) and with increases in the total gross agricultural productivity in value terms. The drought of 2010 also affected labor market outcomes, making household heads and his/her partner to work more on the farm. In addition, rural households reduce the investments after droughts, which could be interpreted as postponing investment decisions and a way to smooth consumption. They also increase the use of external sources of water as their own water sources dry up during droughts. Fourth, after consecutive droughts in 2010 and in 2013, rural households partly implement the same strategies as in the short-term and medium-term, but the adaptation becomes more difficult. On the one hand, some strategies observed before such as working more in agriculture, using external sources of water and smoothing consumption by postponing investments, are less likely to be implemented. On the other hand, the rural households keep using more land available for crops and there is an increase in the land area with investments (excluding housing). In this sense, droughts

might make rural households aware of the potential future benefits of making investments, which is compatible with the findings of [Burke and Emerick \(2016\)](#). Fifth, the analysis of the climate-yearly weather deviations and the DID gave qualitatively similar findings, with a positive effect of temperature shocks and droughts (short-term and medium-term in the DID) on gross agricultural productivity in value terms, with a negative marginal effect in high-temperature municipalities, and a positive marginal effect in low-temperature municipalities. This article thus proposes an alternative manner to examine the short-term and medium-term adaptation decisions of the rural households.

The goal of this PhD dissertation is to analyze different ways in which climate change and climatic phenomena can impact economic outcomes in low and middle income countries. From my perspective, the main contributions of general interest of the dissertation are: as concerns the first chapter, it assesses evidence of the potentially adverse effects of migration as a coping mechanism following drought when other means of insurance do not exist. The results on marriage-related migration show that women migrate mainly within the district of origin and that the general effect of drought is to constrain migration. Marriage-related migration following moderate droughts happens at the within-district level but for women in older age groups, with little evidence of an increase in marriage-related migration for young girls (10 to 17 years old). This is an optimistic result indicating no large changes in the probability of child marriages following drought in Malawi. Regarding migration related to work, there is an increase in men's migration but at the between-district level, particularly for young boys following severe drought, and to a lesser extent for young girls. This is a pessimistic result providing evidence of children's migration for work following drought. Regarding the second chapter, it provides evidence of the negative effect of ENSO events (El Niño and La Niña) on health, which has not been explored neither in the literature on the effects of air pollution on health nor in the literature on the effects of extreme temperature on health. The dissertation addresses this and shows the damaging effects that ENSO events and air pollution can have on the health of children at birth. Moreover, the size of the impact of ENSO events is much larger than the consequences of pollution exposure. The third chapter shows that extreme heat affects agricultural productivity of rural households in a low-middle income country such as Colombia, particularly, higher temperature increases agricultural productivity in low-temperature municipali-

ties but reduces agricultural productivity in high-temperature municipalities for crop producers. It also proposes an alternative way to capture climate and weather shocks on agricultural productivity using a DID framework. In addition, it shows that rural households adapt to droughts by using more available land, increasing the probability of the household head or partner of working in their plots. This also allows the rural households to increase agricultural productivity following droughts, but some of these strategies are harder to implement when rural households experienced consecutive droughts.

The dissertation has policy implications along different angles. Regarding migration, safety nets and policies to provide a more stable income during droughts in Malawi could help to reduce child migration for the rural households affected by these shocks. With respect to the effects of ENSO events on health, policymakers could design ex-ante measures to mitigate the impact of these shocks, assuming that accurate predictions are available. In addition, policy interventions should also be targeted during pregnancy, as the period while children are in utero is one of the most important to their later development and a crucial component of human capital development (see [Almond and Currie \(2011a\)](#)). As concerns the responses to droughts in agriculture, policymakers should take into account the winners and losers of extreme heat. Living in areas with high temperature and having to face additional extreme temperature could make it more difficult to adapt in the new harsh conditions. In this sense, implementing adaptive measures against droughts becomes harder for the rural households. Additionally, as low-temperature areas could benefit from higher temperature, some crops might also benefit and have the possibility of being produced there. Food security policies might have to consider these aspects too, and probably, encourage switching the production from some crops to others, and move the production of specific crops from some areas to others, or provide safety nets in the most affected high-temperature areas.

Résumé

Cette thèse met en évidence les différents impacts économiques de la variabilité du climat et des phénomènes météorologiques dans les pays à faible revenu et à revenu intermédiaire. Le changement climatique affecte de nombreux aspects de la vie humaine, non seulement parce qu'il altère les régimes climatiques, mais aussi parce qu'il augmente la fréquence et l'intensité d'événements météorologiques extrêmes tels que l'oscillation australe El Niño (ENSO ou événements El Niño-La Niña).² Ces événements, parmi plusieurs autres conséquences, pourraient affecter les résultats agricoles, inciter les individus à migrer ou altérer leur santé. En ce qui concerne la migration, une abondante littérature documente la façon dont les populations peuvent se tourner vers la migration pour faire face à la sécheresse ou à d'autres chocs climatiques (Berlemann and Steinhardt, 2017; Beine and Jeusette, 2018; Cattaneo et al., 2019; Hoffmann et al., 2020). Une autre partie de la littérature analyse les effets négatifs du changement climatique et des températures extrêmes sur les nouveau-nés et la mortalité (Deschênes et al., 2009; Barreca and Schaller, 2020; Barreca et al., 2015). En termes d'agriculture, D'Agostino and Schlenker (2016) considèrent que le changement climatique accroît la volatilité de la production alimentaire, ce qui peut annuler les gains de rendements moyens obtenus avec la Révolution Verte. Tout récemment, Ortiz-Bobea et al. (2021) ont constaté que le changement climatique anthropogénique a réduit la productivité totale des facteurs (TFP) de l'agriculture mondiale d'environ 21% depuis 1961, ce qui équivaut à la perte des sept dernières années de croissance de la productivité. L'effet néfaste du changement climatique sur la TFP n'a pas été le même à travers les régions, affectant plus sévèrement les régions plus chaudes comme l'Afrique, l'Amérique latine et les Caraïbes (une réduction d'environ 26% à 34%). Dans

²Il est prévu que l'intensité et la fréquence de phénomènes tels qu'El Niño augmentent dans les années à venir (voir le chapitre 3 du rapport "Réchauffement climatique de 1,5 degré Celsius" du Panel intergouvernemental du changement climatique et Wang et al. (2017))

ce sens, cette thèse fournit des analyses des trois impacts économiques particuliers du changement climatique: en premier lieu, sur les migrations induites par le climat et liées aux sécheresses; deuxièmement, sur la manière dont les phénomènes climatiques tels que les événements El Niño et La Niña (ENSO) peuvent affecter les résultats à la naissance de l'enfant dans le contexte de la pollution de l'air et de la santé; et troisièmement, sur les mécanismes d'adaptation de l'agriculture lorsque les agriculteurs sont confrontés aux sécheresses et à la variabilité climatique.

Cette thèse mobilise des données de deux pays: le Malawi et la Colombie. Ces pays ont connu des événements météorologiques extrêmes et fréquents tels que la sécheresse. En outre, ils disposent de données très riches, nécessaires pour l'évaluation des diverses manières dont le changement climatique se manifeste. Il est ainsi possible de mener une analyse approfondie de la variabilité climatique dans des contextes ruraux et urbains. L'étude de ces deux cas permet non seulement d'explorer et d'analyser plusieurs aspects économiques, mais aussi de démontrer l'ampleur de la menace qui plane sur la vie humaine. Dans le contexte rural du Malawi, par exemple, les sécheresses ont un impact sur les activités agricoles et, dans la mesure où les ménages ne peuvent pas s'adapter en achetant des assurances ou des semences adaptées (marge intensive), la migration devient une alternative (marge extensive). La sécheresse agit donc comme un facteur qui entraîne des réponses migratoires différentes selon le genre (pour le travail et pour le mariage) et l'âge. Pour un pays à revenu intermédiaire comme la Colombie, le contexte rural et les caractéristiques agricoles des marchés sont différents. Le troisième chapitre montre qu'en effet les ménages ruraux colombiens réagissent aux sécheresses en augmentant la superficie totale des cultures, mais que cela ne leur permet pas d'accroître la productivité agricole. Le type de marchés ruraux en Colombie est probablement différent de celui trouvé au Malawi, le niveau de revenu est plus élevé en Colombie et les ménages ruraux diversifient davantage leurs cultures, parmi d'autres facteurs différents. Même si la migration induite par la sécheresse existe aussi en Colombie, les institutions matrimoniales ne sont pas aussi importantes que dans le cas du Malawi (voir la discussion du premier chapitre) et la migration des enfants se fait principalement pour des motifs liés au travail en Colombie. En outre, des phénomènes particuliers comme El Niño entraînent des périodes de sécheresse en Colombie, et créent des distorsions supplémentaires sur les marchés agricoles. Le deuxième chapitre de la thèse montre que les événements ENSO tels que El Niño affectent directement et indirectement la

santé des enfants pendant la période in utero, une fois que l'effet de la pollution atmosphérique sur la santé est pris en compte. Nous soutenons qu'une partie des distorsions des événements ENSO se produisent par l'effet d'équilibre général qu'ils créent sur les marchés agricoles et les prix des aliments. La discussion des résultats de la thèse montre non seulement les différentes interrelations dans les contextes ruraux et urbains de la variabilité climatique, mais aussi que les relations peuvent être très complexes dans différents domaines, difficiles à étudier et dépendantes du contexte.

Une autre caractéristique importante commune aux différents chapitres est la fusion de données de télédétection (images satellitaires), d'enquête et de sources administratives. L'utilisation de données météorologiques objectives de manière standardisée entre zones géographiques a été examinée par [Dell et al. \(2012\)](#), dans les études agricoles de l'adaptation au changement climatique ([Deschênes and Greenstone, 2011](#); [Schlenker and Roberts, 2009](#)) et dans la littérature sur la migration par [Millock \(2015\)](#). L'utilisation de sources externes permet de construire des variables météorologiques exogènes et de réduire le bruit potentiel et les erreurs de mesure des événements extrêmes auto-déclarés dans les enquêtes. À cet égard, cette thèse tire parti de la disponibilité croissante des données issues des images satellitaires, qui ont commencé à être plus fréquemment utilisées dans plusieurs domaines tels que l'informatique, la géographie et l'économie (voir [Donaldson and Storeygard \(2016\)](#) pour une discussion plus large sur l'utilisation des données de télédétection en économie). D'une part, les chapitres 1 et 3 utilisent des images satellitaires brutes pour construire des indices de précipitations standardisés et définir des mesures de sécheresse précises pour l'étude de la migration au Malawi et de l'agriculture en Colombie. D'autre part, le deuxième chapitre s'appuie sur un ensemble de données construit à l'aide d'images satellitaires mensuelles générées à partir d'informations spatiales sur la pollution et de données météorologiques des stations de surveillance de Bogotá. Ces trois chapitres sont donc une illustration de l'utilité des instruments géographiques pour les économistes. D'autres applications dans ce domaine sont attendues dans un avenir proche.

Cette thèse contribue à la littérature économique par différents aspects. Le premier chapitre regroupe la littérature sur les migrations induites par le climat ([Gray and Mueller, 2012](#); [Hirvonen, 2016](#); [Kubik and Maurel, 2016](#)) avec celle des institutions matrimoniales ([Rosenzweig and Stark, 1989](#); [Mobarak et al., 2013](#); [Corno et al., 2020](#)). L'objectif de cette étude, est d'explorer l'existence

de disparités dans les réponses migratoires suite à une sécheresse selon le sexe, l'âge et les motifs de migration, en particulier pour les jeunes. Ce chapitre met ainsi en évidence les effets hétérogènes des événements climatiques extrêmes sur la migration (Gray and Mueller, 2012; Dillon et al., 2011; Baez et al., 2017a). Par ailleurs, il met en exergue les différences entre les sexes et la manière dont les institutions liées au mariage influencent cette migration. Ces dimensions n'avaient jusqu'ici jamais été explorées par la littérature scientifique.

Le deuxième chapitre établit un lien entre deux branches de littérature: celles relatives aux effets sur la santé de la pollution de l'air (Graff Zivin and Neidell, 2013; Lavaine and Neidell, 2017; Arceo et al., 2015) et des températures extrêmes (Deschênes and Greenstone, 2011; Barreca et al., 2015), en ajoutant l'effet des événements ENSO, qui n'a pas été exploré dans la littérature. À la différence des autres études, le chapitre examine conjointement si les événements ENSO, les conditions météorologiques locales et la pollution de l'air affectent la santé urbaine, en particulier celle des nouveau-nés. En effet, les événements ENSO peuvent avoir un impact sur la santé par différents mécanismes. Ils peuvent avoir un impact sur la météo locale et ensuite sur la santé; les événements ENSO peuvent affecter la pollution et par la suite la santé; ils peuvent également avoir un effet plus direct sur la santé et enfin, ils peuvent aussi affecter l'agriculture, les marchés alimentaires, les prix des aliments et générer une perte de revenus pour les ménages, ce qui peut avoir un impact sur la santé. À cet égard, les impacts de la pollution de l'air sur la santé peuvent être considérés comme distincts des chocs induits par les événements ENSO.

Le troisième chapitre s'ajoute à la littérature sur les réponses au changement climatique dans l'agriculture (Blanc and Schlenker, 2017; Burke and Emerick, 2016; Deschênes and Greenstone, 2007; Schlenker and Roberts, 2009), en contribuant aux rares analyses scientifiques de l'adaptation agricole dans les pays à revenu faible et moyen, et en proposant une mesure alternative pour évaluer l'impact du climat sur l'agriculture. Ce chapitre aborde la question de l'impact des sécheresses sur les décisions agricoles, sur la production alimentaire des ménages ruraux en Colombie à court et à moyen terme, puis dans quelle mesure les ménages ruraux s'adaptent à ces chocs. Je cherche également à déterminer s'il existe une hétérogénéité entre les ménages vivant dans des zones à température élevée et ceux vivant dans des zones à faible température, un aspect moins étudié dans la littérature. Le chapitre met en évidence un autre point tout aussi important, celui des impacts climatiques en dehors du domaine des cultures vivrières étudiées dans la littérature (maïs,

blé, riz et soja) en explorant un ensemble plus large de cultures et en analysant d'autres facteurs de production agricole tels que la main-d'œuvre, l'utilisation des investissements et l'accès à l'eau. Comme l'a souligné [Hertel and de Lima \(2020\)](#), la littérature devrait aller au-delà de l'impact sur le rendement (en particulier des cultures vivrières), pour lequel nous disposons de meilleures données et de meilleurs modèles, et s'intéresser à d'autres produits alimentaires, aux différents facteurs de production agricole et aux impacts nutritionnels.

Le premier chapitre s'intitule "Réponses migratoires de genre aux sécheresses au Malawi" et a été coécrit avec Katrin Millock. Nous y étudions deux aspects de la littérature: les structures de mariage selon les coutumes de la dot et les mouvements migratoires suite à la sécheresse, en explorant ses impacts selon le sexe, l'âge et les raisons déclarées de se déplacer par les migrants. Le Malawi est un pays qui compte un nombre important de migrants pour le mariage et pour le travail, et il souffre de sécheresses récurrentes qui affectent la production agricole et la sécurité alimentaire de la population. Ce chapitre met en relation différents aspects de la littérature de la manière suivante. D'une part, il existe déjà des études qui montrent des changements dans les modèles de migration suite à la sécheresse ([Gray and Mueller, 2012](#); [Hirvonen, 2016](#); [Kubik and Maurel, 2016](#)), mais il existe encore des lacunes dans la connaissance de l'influence du genre sur les migrations induites par le climat ([Chindarkar, 2012](#)), et comment les réponses varient en fonction du genre selon les contextes culturels. Alors que la migration des femmes est limitée par la variabilité des températures ou les sécheresses en Éthiopie et au Nigéria ([Dillon et al., 2011](#); [Gray and Mueller, 2012](#)), les femmes migrent davantage à la suite de températures plus élevées ou de sécheresses dans certains pays d'Asie et des Caraïbes ([Mueller et al., 2014](#); [Thiede and Gray, 2017](#); [Baez et al., 2017a,b](#)). D'autre part, une littérature récente en économie du développement traite du rôle des institutions matrimoniales sur les résultats socio-économiques ([Corno and Voena, 2017](#); [Ashraf et al., 2020](#)). En tant que tel, le prix de la mariée est une somme d'argent ou des biens payés par un homme à la famille d'une fille pour assurer son mariage, et ce phénomène est plus courant dans les pays d'Afrique subsaharienne comme le Malawi. Dans d'autres pays comme l'Inde, c'est la famille de la mariée qui donne de l'argent à la famille du marié (dot) ([Lowes and Nunn, 2018](#)). À cet égard, [Corno et al. \(2020\)](#) montrent que dans les pays ayant une culture de prix de la mariée, des chocs défavorables pendant l'adolescence peuvent augmenter la probabilité d'un mariage précoce

(voir [Corno and Voena \(2017\)](#) pour le cas de la Tanzanie). Toutefois, dans les pays avec une culture de paiement de dot, de tels chocs négatifs peuvent réduire la probabilité d'un mariage précoce. Dans ce sens, le chapitre montre l'impact des sécheresses sur la migration liée au mariage ou au travail à différents âges, en fonction du genre, lorsque la coutume de prix de la mariée est pratiquée.

Le chapitre analyse l'impact de la sécheresse sur la migration des femmes et des hommes, en construisant un panel rétrospectif sur une base annuelle (un panel pour le migrant individuel i de 2000-2016 qui se déplace au temps t). Il combine ces données avec des mesures de sécheresse basées sur les précipitations et la température à partir d'images satellitaires. S'appuyant sur un design quasi-expérimental, le chapitre compare la probabilité qu'un individu d'un groupe d'âge particulier migre d'un district d'origine vers un district de destination au cours d'une année donnée, avec ou sans sécheresse. Premièrement, nous explorons les changements dans la probabilité de migrer pour se marier pour les filles âgées de 10 à 17 ans. Cela peut permettre de vérifier les différences de genre dans la migration induite par le climat et de tester si les sécheresses peuvent induire des mariages d'enfants dans un pays où le prix de la mariée est utilisé. Ensuite, nous comparons les résultats avec la migration des hommes liée au travail et la migration des femmes pour le mariage afin d'explorer davantage les différences. En termes de mécanisme, étant donné que les sécheresses affectent le rendement des cultures et le revenu des familles dépendant de l'agriculture, envoyer des membres de la famille vivre chez d'autres proches ou anticiper le mariage des filles pendant ces chocs pourrait être un moyen de lisser la consommation de la famille.

Nos résultats montrent des augmentations plutôt faibles de la probabilité de migration des femmes à l'intérieur du district pour cause de mariage suite à une sécheresse modérée (quelques points de pourcentage), et nos tests ne permettent pas de rejeter l'hypothèse nulle selon laquelle la sécheresse n'affecte pas la probabilité des mariages d'enfants. Cependant, à la suite de sécheresses sévères, on observe une augmentation de la probabilité de migration entre districts pour des raisons de travail pour les garçons âgés de 10 à 17 ans (9 points de pourcentage), et dans une moindre mesure pour les filles (5 points de pourcentage). Ce n'est que lorsque nous combinons les raisons de migration des filles liées au mariage et au travail que nous constatons une augmentation de 2,3 points de pourcentage de la migration suite à une sécheresse sévère. Ce chapitre apporte quatre

contributions. Premièrement, en analysant la migration liée au mariage et le moment où la sécheresse se produit dans le lieu d'origine de l'individu, afin d'analyser l'impact de la sécheresse sur la migration pour le mariage. Deuxièmement, en ce qui concerne la littérature sur la migration induite par le climat, nous analysons la sécheresse comme un facteur incitant à différents types de migration au Malawi, non seulement pour des motifs liés au travail, mais aussi pour le mariage comme moyen de lier la consommation de la famille pendant les sécheresses. Troisièmement, l'analyse utilise des informations détaillées dans les districts d'origine et de destination sur plus d'une décennie et demie, en exploitant la chronologie de la sécheresse, à la différence des études précédentes sur le Malawi. Et finalement, nous examinons si le type de migration change en fonction de la distance à laquelle les migrants se déplacent, en comparant la migration à l'intérieur du district d'origine et la migration entre districts.

Le deuxième chapitre "Chocs climatiques, qualité de l'air et santé à la naissance à Bogotá" est coécrit avec Jorge Bonilla. Cette étude contrairement à toute étude précédente, examine en même temps les effets de la pollution de l'air, des événements ENSO, et du climat local sur la santé. Nous l'examinons pour la ville de Bogotá pour tous les nouveau-nés sur la période 1998-2015. La ville est l'un des plus grands centres urbains d'Amérique latine et est fréquemment exposée aux événements ENSO et à des niveaux élevés de pollution. En termes de pollution, elle atteint des niveaux équivalents à des villes comme New Delhi. Selon l'Indice de Qualité de l'Air en temps réel tiré du Projet d'Indice Mondial de Qualité de l'Air, Bogotá a atteint une valeur de 141 dans l'indice le 8 mars 2019 (très nocif pour les individus). Pour le même jour, l'indice avait une valeur de 151 (malsain) pour New Delhi, alors qu'il avait une valeur de 23 (bon) à Paris. Les coûts sanitaires de la pollution de l'air sont indéniables. Selon le rapport de 2017 du Département National de la Planification [DNP \(2017\)](#), la dégradation de l'environnement a causé 10,5% (3,219) du total des décès à Bogotá, avec un coût estimé à 4,2 milliards de pesos colombiens (2,5% du PIB de la ville). Cependant, ce rapport n'examine pas les effets d'ENSO sur la santé. En effet, nous estimons que les effets d'ENSO sur le poids à la naissance sont bien plus importantes que ceux dus à pollution. Notre analyse montre ainsi que les coûts de la santé des événements ENSO peuvent également être considérables.

Bien que les effets de la pollution de l'air sur la santé aient déjà été étudiés dans la littérature

en utilisant des variables instrumentales similaires à celles que nous utilisons (direction du vent), l'impact des événements ENSO sur la santé n'a pas encore été exploré. Notre approche s'appuie sur des articles tels que [Hsiang et al. \(2011\)](#) et [Dingel et al. \(2020\)](#), où les événements ENSO sont considérés comme des effets globaux, perturbant non seulement les facteurs météorologiques, mais créant également des perturbations sur les prix des aliments, et d'autres effets d'équilibre général qui peuvent altérer la consommation des ménages et, de cette manière, la santé des mères enceintes et des enfants. Notre argument est qu'ENSO ne se manifeste pas seulement comme un choc climatique extrême qui influence la météo, mais qu'il peut aussi avoir un impact sur l'agriculture et les catastrophes et donc, qu'il pourrait induire des changements sur les marchés alimentaires et affecter le revenu des ménages. Par conséquent, l'effet de la pollution sur la santé peut être considéré comme distinct des chocs qui sont induits par les événements ENSO. En outre, les événements ENSO peuvent également affecter les niveaux de pollution de l'air, ce qui a un impact indirect sur la santé. À cet égard, le chapitre fait le lien entre la littérature sur la pollution de l'air et ses effets sur la santé ([Graff Zivin and Neidell, 2013](#); [Lavaine and Neidell, 2017](#); [Arceo et al., 2015](#)) et celle des températures extrêmes sur la santé ([Barreca et al., 2015](#); [Deschênes et al., 2009](#); [Deschênes and Greenstone, 2011](#)), en ajoutant l'impact provenant des événements ENSO (El Niño et La Niña).

Pour réaliser l'analyse, nous utilisons à la fois les données sur la pollution et la météo provenant des stations de surveillance et les données sur tous les enfants nés à Bogotá de 1998 à 2015. Nous nous concentrons sur les enfants pendant la gestation afin de capturer l'exposition à ENSO d'une manière plus instantanée et d'éviter tout biais dans les estimations créé par le fait de ne pas contrôler les facteurs non observés les plus courants chez les adultes (fumer, manger sainement, faire de l'exercice, etc.). Comme attribuer la pollution de l'air et les conditions météorologiques aux résultats de santé n'est pas chose simple, nous commençons par utiliser la méthode de l'interpolation de pondération par l'inverse de la distance pour créer un ensemble mensuel de cartes des variables de pollution et de variables météorologiques dans un logiciel SIG. Comme les adresses des mères ne sont pas disponibles, nous créons un panel de trimestres courants agrégés par centres de santé, où le trimestre correspond au trimestre de gestation des enfants nés au cours d'un mois spécifique dans le centre de santé. Comme dernière étape, nous construisons une zone tampon de 2 km autour de l'adresse du centre de santé où les enfants sont nés, et nous attribuons les variables de pollution et

de météo à l'intérieur de cette zone tampon. La stratégie d'identification repose sur une approche en deux étapes. La première étape instrumente l'équation de la pollution par la direction du vent. La deuxième étape établit la relation entre la pollution et ENSO sur la santé, en contrôlant les facteurs météorologiques, parmi d'autres contrôles. Il est important de noter que le système a l'avantage de capturer l'effet décalé sur la santé de l'exposition à la pollution et à l'ENSO pendant les trois trimestres de la gestation. Tandis que l'identification des effets de la pollution de l'air sur la santé repose sur la variation transversale et temporelle, l'effet des événements ENSO sur les résultats de santé à la naissance repose entièrement sur la variation temporelle.

Comme résultat principal, dans toutes les spécifications, après avoir inclus les canaux de la pollution et du climat local classique, l'exposition à l'ENSO diminue le poids à la naissance et augmente la probabilité d'avoir un poids faible à la naissance. Quand les estimations utilisent SO_2 , PM_{10} et $PM_{2,5}$, les événements El Niño et La Niña réduisent le poids à la naissance lorsque l'exposition se produit au cours du dernier trimestre de la gestation, période pendant laquelle la mère est sur le point d'accoucher. Dans les estimations avec CO et NO_X , les effets séparés de l'ENSO montrent qu'El Niño (au troisième et premier trimestre) et La Niña (au deuxième trimestre) augmentent la probabilité d'avoir un faible poids à la naissance, et réduisent le poids en grammes à la naissance. Indépendamment du polluant utilisé, l'effet cumulé d'ENSO sur les trimestres de gestation est plus souvent statistiquement différent de zéro aux niveaux de significativité statistique conventionnels pour El Niño, mais l'est moins pour les épisodes La Niña. Comme résultat secondaire, nous confirmons l'impact négatif de la pollution de l'air sur la santé. Nous trouvons que l'exposition *in utero* aux concentrations moyennes de SO_2 au cours du premier et du troisième trimestre de la gestation influence les mesures de santé à la naissance, tandis que l'exposition au $PM_{2,5}$ au cours du troisième trimestre de la gestation augmente la probabilité d'être prématuré. En outre, le PM_{10} semble diminuer le poids à la naissance lorsque l'exposition a lieu au cours du deuxième trimestre de la gestation, tandis que l'exposition au CO (au cours du deuxième et surtout du troisième trimestre) augmente la probabilité d'avoir un faible poids à la naissance. Lorsque l'on analyse l'effet cumulatif pendant la durée totale de la gestation, SO_2 et CO sont les seuls polluants qui ont un impact négatif sur le poids de naissance à des niveaux de significativité statistique conventionnels. De plus, l'impact des événements ENSO est beaucoup plus important que les conséquences de l'exposition

à la pollution. Le fait d'être exposé à l'ENSO peut réduire le poids à la naissance jusqu'à 1.3%, tandis qu'une augmentation de 1 *ppb* de SO_2 ou de 1 $\mu g/m^3$ de PM_{25} pourrait réduire le poids à la naissance jusqu'à 0.3% ou 0.14%, respectivement. Ce chapitre met donc en évidence l'effet global des chocs climatiques extrêmes sur la santé à la naissance via ENSO.

"Sécheresses et adaptation de l'agriculture au changement climatique" correspond au troisième chapitre de la thèse. J'y analyse comment les sécheresses affectent les décisions agricoles et la production alimentaire des ménages ruraux en Colombie, à court terme et à moyen terme. L'objectif est de déterminer dans quelle mesure les ménages ruraux s'adaptent à ces chocs et s'il existe des différences dans le comportement des ménages ruraux à court terme et à moyen terme. L'analyse se concentre également sur un aspect important de l'hétérogénéité: les différences entre les ménages ruraux situés dans des municipalités où la température moyenne est élevée et ceux situés dans des municipalités où la température moyenne est faible. La Colombie est un pays où les sécheresses sont récurrentes et bien que seulement 6% du PIB provienne de l'agriculture, ce secteur représente 17,3% de l'emploi total en juin 2020. Le pays est également fréquemment touché par le phénomène El Niño, qui entraîne des sécheresses dans le cas de la Colombie. Pour replacer dans son contexte l'impact de ce phénomène, selon le ministère de l'Environnement du pays-IDEAM, le phénomène El Niño pourrait avoir réduit de 80% les précipitations en Colombie au cours des premiers mois de 2019. À cet égard, les ménages ruraux peuvent, entre autres, essayer de faire face aux sécheresses sur la marge intensive, en souscrivant des assurances ou des crédits, en lissant leur consommation, en mélangeant les cultures ou en reportant les investissements, entre autres. Les ménages ruraux peuvent également essayer de s'adapter sur la marge extensive en migrant.

Bien qu'il existe plusieurs études sur le changement climatique et l'agriculture, une vaste littérature s'est concentrée sur les États-Unis (Deschênes and Greenstone, 2007; Schlenker and Roberts, 2009; Burke and Emerick, 2016), et l'impact du changement climatique sur l'agriculture dans des pays moins développés a été bien moins étudié. En termes d'adaptation des ménages ruraux, Burke and Emerick (2016) pour les États-Unis et Costinot et al. (2016) d'un point de vue global, montrent que l'alternance des cultures pourrait être une réponse possible au changement climatique. Les ménages ruraux peuvent également réagir à ces chocs en réalisant des investissements dans l'unité

de production agricole, en essayant d'obtenir une assistance technique ou en modifiant leur utilisation d'engrais. De la même manière, [Rosenzweig and Wolpin \(1993\)](#) soulignent également que les ménages peuvent réagir aux sécheresses sur la marge intensive en augmentant le travail non agricole (pour le cas de l'Inde, voir [Jayachandran \(2006\)](#)), vente de bétail (voir [Fafchamps et al. \(1998\)](#) pour le cas du Burkina Faso), ou sur la marge extensive, par la migration (voir [Cattaneo et al. \(2019\)](#) pour une perspective globale). Ce chapitre apporte quatre contributions. Premièrement, il ajoute à la littérature rare sur les pays à faible et moyen revenu, et propose une mesure alternative pour évaluer l'impact climatique et l'adaptation de l'agriculture à court et à moyen terme en Colombie. Deuxièmement, il explore l'hétérogénéité entre les ménages vivant dans des zones à haute et basse température, un aspect moins exploré dans la littérature. Troisièmement, il montre qu'à court terme, les ménages ruraux colombiens s'adaptent en utilisant davantage de terres disponibles, ce qui pourrait conduire à une augmentation de la productivité agricole brute. Une autre nouveauté consiste à analyser les sécheresses consécutives. Quatrièmement, il explore plus en détail la manière dont les ménages ruraux s'adaptent aux sécheresses en analysant un ensemble plus large de cultures: légumes, fruits, céréales et café. Comme l'a souligné [Hertel and de Lima \(2020\)](#), la FAO recense 175 cultures distinctes, mais la majorité des études sur l'impact du climat se sont concentrées sur les changements de rendement de quatre cultures de base, à savoir: le maïs, le blé, le riz et le soja. Ces quatre cultures de base ne représentent qu'un quart de la valeur totale de la production agricole. Ce chapitre analyse les impacts climatiques en dehors du domaine des cultures vivrières en explorant d'autres facteurs de production agricole tels que le travail, l'utilisation des investissements et l'accès à l'eau. Selon [Hertel and de Lima \(2020\)](#), la littérature devrait aller au-delà de l'impact sur le rendement, pour lequel nous disposons de meilleures données et de meilleurs modèles, et devrait s'intéresser aux autres produits alimentaires, aux intrants agricoles et aux impacts nutritionnels.

Le troisième chapitre utilise l'enquête de panel colombienne (ELCA), qui dispose d'informations très riches sur la production agricole et les investissements réalisés sur les terres des ménages ruraux. Le chapitre apparie les trois vagues de l'ELCA (2010-2013-2016) avec des données météorologiques provenant d'images satellitaires afin d'éviter les erreurs de mesure qui pourraient résulter de l'utilisation de chocs subjectifs auto-déclarés. Comme le panel ELCA suit le même ménage sur trois vagues, il permet d'étudier les adaptations telles que l'utilisation de plus de terres pour certaines cultures par

rapport à d'autres ou les changements vers d'autres types d'activités agricoles au fil du temps. Trois méthodes différentes sont utilisées dans le chapitre: premièrement, je utilise le cadre d'identification utilisé dans la littérature qui sépare l'effet du climat des déviations météorologiques annuelles (Kelly et al., 2005; Deschênes and Kolstad, 2011; Burke and Emerick, 2016; Bento et al., 2020); la deuxième partie suit Deschênes and Greenstone (2007), Schlenker and Roberts (2009) et Aragón et al. (2021) pour explorer les différentes marges d'adaptation en réponse aux températures extrêmes, en utilisant les mesures standard des degrés-jours de croissance (GDD) et des degrés-jours de dommage (HDD); et troisièmement, je propose d'utiliser l'indice de précipitation standard (SPI) comme mesure de la sécheresse pour évaluer les effets des conditions météorologiques extrêmes en agriculture. Cette mesure alternative est utilisée dans un cadre de Différence de Différences (DID) pour voir comment les ménages ruraux de Colombie s'adaptent. La méthode DID analyse l'adaptation à court terme, puis à moyen terme.

Les principales conclusions de ce chapitre sont au nombre de cinq. Premièrement, en ce qui concerne l'analyse du climat et des écarts météorologiques annuels, la productivité agricole brute en valeur augmente avec les écarts par rapport à la température moyenne à long terme, alors qu'il n'y a pas d'effet par rapport à l'écart de température (écart par rapport à la moyenne à long terme); par rapport aux effets marginaux, une température plus élevée fait croître la productivité agricole brute en valeur dans les municipalités à faible température mais fait décroître celle-ci dans les municipalités à température élevée. Comme mentionné par Aragón et al. (2021), les effets hétérogènes de la température ont été étudiés pour le cas des États-Unis (Deschênes and Greenstone, 2007; Schlenker and Roberts, 2009) mais moins pour les pays en développement. Deuxièmement, selon les différentes analyses du court et du moyen terme (climat-écarts climatiques annuels et DID), la productivité agricole brute en valeur des céréales et du café bénéficie des sécheresses ou des températures élevées. La productivité agricole brute augmente dans les municipalités à faible température pour les céréales et le café, et diminue dans celles des zones à température élevée. Les légumes semblent être plus affectés suite aux sécheresses (analyse DID). Dans les différentes analyses, la productivité agricole brute en valeur des légumes et des fruits présente des réductions, en particulier dans les municipalités à basse température (avec augmentations dans les municipalités à haute température). Troisièmement, en ce qui concerne le court et le moyen terme adaptation, la sécheresse de

2010 a poussé les ménages ruraux à réallouer les terres en réduisant la superficie des terres laissées en jachère et en les affectant aux cultures et à l'élevage. Cela reflète certains des mécanismes du cadre théorique du chapitre. Elle va également dans le sens de l'augmentation des intrants agricoles (terre et travail) trouvé dans [Aragón et al. \(2021\)](#) et de l'augmentation de la productivité agricole brute totale en valeur. La sécheresse de 2010 a également affecté le marché du travail, obligeant les ménages à travailler davantage sur l'exploitation. En outre, les ménages ruraux réduisent leurs investissements après les sécheresses, ce qui peut être interprété comme un report des décisions d'investissement et un moyen de lisser la consommation. Ils augmentent également l'utilisation de sources d'eau externes car leurs propres sources d'eau se tarissent pendant les sécheresses.

Quatrièmement, suite aux sécheresses consécutives en 2010 et en 2013, les ménages ruraux mettent partiellement en œuvre les mêmes stratégies, mais l'adaptation devient plus compliquée. D'une part, certaines stratégies observées auparavant, telles que travailler davantage dans l'agriculture, utiliser des sources d'eau externes et lisser la consommation en reportant les investissements, sont moins susceptibles d'être mises en œuvre. D'autre part, les ménages ruraux continuent à utiliser davantage de terres disponibles pour les cultures et on observe une augmentation de la superficie des terres dans lesquelles le ménage fait des investissements (à l'exclusion des logements). En ce sens, les sécheresses pourraient faire prendre conscience aux ménages ruraux des bénéfices potentiels futurs des investissements, ce qui est compatible avec les résultats de [Burke and Emerick \(2016\)](#). Cinquièmement, l'analyse du climat-écarts climatiques annuels et DID ont donné des résultats qualitativement similaires, avec un effet positif des chocs thermiques et des sécheresses (à court et moyen terme dans la DID) sur la productivité agricole brute en valeur, avec des effets marginaux négatifs dans les municipalités à haute température, alors qu'ils sont positifs dans celles à basse température. Cet article propose donc une manière alternative d'examiner les décisions d'adaptation à court et moyen terme des ménages ruraux.

L'objectif de cette thèse de doctorat est d'analyser les différentes façons dont le changement climatique et les phénomènes climatiques peuvent avoir un impact sur les résultats économiques dans les pays à revenu faible et intermédiaire. De mon point de vue, les principales contributions d'intérêt général de la thèse sont les suivantes: En ce qui concerne le premier chapitre, il montre les effets potentiellement négatifs de la migration comme mécanisme d'adaptation à la suite d'une sécheresse

lorsque d'autres moyens d'assurance n'existent pas. Les résultats sur la migration liée au mariage montrent que les femmes migrent principalement dans le district d'origine et que l'effet général de la sécheresse est de limiter la migration. La migration liée au mariage à la suite de sécheresses modérées se produit au niveau du district mais principalement pour les femmes des groupes d'âge plus élevés, et il y a très peu d'indications d'une augmentation de la migration liée au mariage pour les jeunes filles (10 à 17 ans). Il s'agit d'un résultat optimiste indiquant qu'il n'y a pas de grands changements dans la probabilité des mariages d'enfants suite à la sécheresse au Malawi. En ce qui concerne la migration liée au travail, on observe une augmentation de la migration des hommes mais au niveau inter-district, en particulier pour les jeunes garçons suite à une sécheresse sévère, et dans une moindre mesure pour les jeunes filles. Il s'agit d'un résultat pessimiste qu'illustre la migration des enfants vers le travail suite à la sécheresse. En ce qui concerne le deuxième chapitre, il permet d'attester l'effet négatif des événements ENSO (El Niño et La Niña) sur la santé, qui n'a été exploré ni dans la littérature sur les effets de la pollution de l'air sur la santé ni dans celle sur les effets des températures extrêmes sur la santé. La thèse aborde cette question et montre les effets dommageables que les événements ENSO et la pollution de l'air peuvent avoir sur la santé des enfants à la naissance. En outre, l'ampleur de l'impact des événements ENSO est beaucoup plus importante que les conséquences de l'exposition à la pollution. Le troisième chapitre montre que la chaleur extrême affecte la productivité agricole des ménages ruraux dans un pays à revenu intermédiaire ou faible comme la Colombie. En particulier, une température élevée augmente la productivité agricole dans les municipalités à faible température mais la réduit dans les municipalités à température élevée pour les producteurs de cultures. Il propose également une autre façon de capturer les chocs climatiques et météorologiques sur la productivité agricole en utilisant un cadre DID. En outre, il montre que les ménages ruraux s'adaptent aux sécheresses en utilisant davantage de terres disponibles, ce qui augmente la probabilité que les ménages travaillent sur leurs parcelles. Cela permet également aux ménages ruraux d'augmenter la productivité agricole après les sécheresses, mais certaines de ces stratégies sont plus difficiles à mettre en œuvre lorsque les ménages ruraux ont expérimenté des sécheresses consécutives.

Les résultats de cette thèse ont plusieurs implications en termes de politique économique. En ce qui concerne la migration, les filets de sécurité et les politiques visant à fournir un revenu plus

stable pendant les sécheresses au Malawi pourraient contribuer à réduire la migration des enfants des ménages ruraux touchés par ces chocs. En ce qui concerne les effets des événements ENSO sur la santé, les décideurs politiques pourraient concevoir des mesures ex ante pour atténuer l'impact de ces chocs, en supposant que des prévisions précises soient disponibles. De surcroît, les politiques d'intervention devraient également être ciblées pendant la grossesse, car la période de gestation est l'une des plus importantes pour le développement ultérieur des enfants et est un élément crucial du développement du capital humain (see [Almond and Currie \(2011a\)](#)). En ce qui concerne les réponses aux sécheresses dans l'agriculture, les décideurs politiques devraient tenir compte des gagnants et des perdants de la chaleur extrême. Le fait de vivre dans des régions où les températures sont élevées et de devoir faire face à des températures extrêmes supplémentaires pourrait rendre plus compliquée l'adaptation à ces nouvelles conditions difficiles. En ce sens, la mise en œuvre de mesures adaptatives contre la sécheresse devient plus difficile pour les ménages ruraux. De plus, comme les zones à faible température pourraient bénéficier d'une température plus élevée, certaines cultures pourraient également en bénéficier et avoir la possibilité d'y être produites. Les politiques de sécurité alimentaire devraient pouvoir tenir compte de ces aspects également, et probablement encourager le passage de certaines cultures à d'autres, le déplacement de la production de cultures spécifiques de certaines zones à d'autres, ou la mise en place de filets de sécurité dans les zones à température élevée les plus touchées.

Chapter 1

Gendered Migration Responses to Drought in Malawi¹

¹This chapter is the product of joint work with Katrin Millock.

Abstract

Migration is a common means of adaptation to weather shocks. Previous research has identified heterogeneous effects according to age, sex, and wealth, but little is still known about how marriage-related institutions affect such migration. Relying on a quasi-experimental identification strategy, we analyze marriage- and work-related migration in Malawi following large droughts, separating the effects for female and male migrants according to different age groups. The analysis based on stated motives of migration reveals marginal decreases in marriage-related migration among girls, but increases in marriage-related migration within districts for women in older age groups. We also find large increases in work-related between-district migration for boys, and to a smaller extent also for girls following severe drought. The results add to the evidence of the potentially adverse effects of migration as a coping mechanism following drought when other means of insurance do not exist.

Keywords: child marriage, climate change, droughts, internal migration, sub-Saharan Africa.

JEL: J12, O15, Q54, R23.

1.1 Introduction

A large literature documents how populations affected by drought or other adverse climatic shocks may turn to migration as a means to cope (Berlemann and Steinhardt, 2017; Beine and Jeusette, 2018; Cattaneo et al., 2019; Hoffmann et al., 2020). Migration is a potential adaptation option especially in countries where the population lacks formal insurance, or has little access to public safety nets. In this article, we aim at linking recent research on marriage patterns according to bride price customs with the patterns of migration following drought, exploring its impacts according to the sex and age of the migrants.

In fact, there are gaps in the knowledge of how gender affects climate-induced migration (Chindarkar, 2012). Existing analyses find different responses for men and women according to cultural contexts. Studies on Ethiopia and Nigeria show that female migration is constrained by variability in temperature or by drought (Dillon et al., 2011; Gray and Mueller, 2012), whereas some studies from Asia and the Caribbean find that women migrate more following higher temperatures or drought (Mueller et al., 2014; Thiede and Gray, 2017; Baez et al., 2017a,b). Whereas clear country-specific differences have been found, the specific mechanisms explaining the heterogeneous effects are not clear. In parallel, a recent literature in development economics documents the role of institutions such as bride price traditions and their impact on socioeconomic outcomes (Corno and Voena, 2017; Ashraf et al., 2020; Corno et al., 2020). In particular, Corno et al. (2020) study how drought affects the age of marriage and show differential impacts of the outcome depending on whether the country has a culture of dowry payments or bride prices. Corno and Voena (2017) find that adverse shocks - measured by rainfall shocks - during teenage years increases the likelihood of early marriage in Tanzania. Evidence such as this is important since it shows the benefits of developing formal insurance or social safety net policies to cope with drought.

The bride price is money or goods paid by a man to a girl's family to secure a bride. Contrary to countries like India where the bride's family gives money to the bridegroom's family (dowry), bride prizes are frequent in sub-Saharan Africa (Lowe and Nunn, 2018). In this article, we use data from Malawi over the period 2000-2016 to investigate how droughts affect men's and women's migration

differently, in particular marriage-related migration. In Malawi, a sizeable number of individuals migrate for marriage, as will be seen in the data. The country suffers recurrent episodes of droughts that affect agricultural production and the population's food security. In particular, the drought of 2005 was one of the most severe experienced recently in Malawi, leading the country to declare a state of disaster and a food emergency. According to the FAO, more than 4.2 million people, or around 34 percent of the population were not able to meet their food needs. Since droughts affect crop yields and the income of families dependent on agriculture, sending family members to live with other relatives or anticipating the marriage of daughters during those shocks could be a way to smooth consumption of the family.

To analyse the impact of drought on women's and men's migration, we construct a retrospective panel on a yearly basis (a panel for individual migrant i from 2000-2016 that moves at time t) based on data from the Living Standards Measurement Study - Integrated Surveys of Agriculture (LSMS-ISA) of Malawi. We combine these data with measures of droughts based on precipitation and temperature from satellite images. Relying on a quasi-experimental design, we compare the probability of an individual in a particular age group migrating from an origin district to a destination district in a particular year, with or without drought. In particular, we explore whether there was an increased probability of migrating for marriage among women aged 10-17. In doing so, we aim to provide more evidence on heterogeneous effects related to gender for climate change induced migration and the risk of drought inducing more child marriages in a country with bride price customs. We also compare the results with women's work-related migration and men's migration for marriage to explore the differences further.

Our results show rather small increases in the probability of women's within-district migration because of marriage following drought (a few percentage points). On the one hand, the results are somewhat optimistic in that we do not find sufficient evidence to reject the null hypothesis of drought not affecting the probability of child marriages. On the other hand, following severe drought, the results show increases in the probability of between-district migration for work-related reasons in the age category of boys (10 to 17 years old). Severe drought increases the probability of migrating for work by almost 9 percentage points for boys, and by 5 percentage points for girls.

Combining both marriage-and work-related migration reasons for girls shows an increase of 2.3 percentage points in migration following severe drought.

The article contributes to two different strands of the literature. Several household studies have found changes in migration patterns following drought (Gray and Mueller, 2012; Hirvonen, 2016; Kubik and Maurel, 2016) A previous article investigating rainfall shocks and migration of household heads in Malawi, (Lewin et al., 2012) finds that migration is reduced following rainfall shocks. This finding goes in line with other studies from Tanzania (Hirvonen, 2016) and a more recent study from Malawi (Jovanovic et al., 2019), that all show evidence of lower rainfall hindering migration through liquidity constraints. Lewin et al. (2012) use cross-section data from the Second Malawi Integrated Household Survey (IHS2) 2004-2005 together with self-reported assessment of whether the household suffered a weather shock during the last five years and data on rainfall variability and past rainfall, whereas Jovanovic et al. (2019) use panel data from 2009/10 (IHS3) and 2012/13 (IHSP) combined with rainfall data.

Another strand of the literature investigates how marriage patterns may be altered following income shocks. Already in the seminal article by Rosenzweig and Stark (1989), marriage of daughters into districts with uncorrelated weather shocks was seen as a form of insurance mechanism to deal with agricultural income risk.² In line with these results, Hotte and Marazyan (2019) find in Senegal that parents do within-kin-group (as opposed to outside-kin-group) marriages as a way to insure against adverse income shocks. Hoogeveen et al. (2011) find some evidence that the marriage of daughters is used as an alternative form of insurance in rural Zimbabwe following drought. The findings are based on data over the period 1994 to 2000 on 400 households. On the one hand, they find a positive but not significant effect of rainfall levels on the marriage rate. This can be explained as a market supply effect since there are fewer men who can afford to pay the bride price following drought. On the other hand, they find evidence that the marriage rate of daughters in poor households increases following a shock to livestock wealth. They do not focus on the age of marriage and the likelihood of an increased probability of child marriages. The age of the wife is the explicit

²Other related articles documenting evidence of income insurance affecting marriage patterns are Mobarak et al. (2013) who find that brides from households protected by flood barriers commanded higher dowry payments and were less likely to marry relatives.

concern of [Corno and Voena \(2017\)](#) who find that adverse income shocks in teenage years increases the likelihood of an early marriage among young women in rural Tanzania, a country where bride price customs are common, as in Malawi. [Corno et al. \(2020\)](#) extend this work to several countries in sub-Saharan Africa, where bride prices are common, and to India, where dowry payments are used instead, to show the opposite effects of bride price payments versus dowry payments on the timing of marriages.

Our article contributes to the empirical literature on marriage and development in four ways. The first contribution of the article is to analyse marriage-related migration and the timing of drought in the origin location of the individual, which provides for additional evidence on the impact of drought on migration for marriage. [Corno et al. \(2020\)](#) cannot control for the actual place of residence at the time of the drought, since they use Demographic and Health Survey (DHS) data and detailed information on the timing of marriage but have to assume that the individuals did not move after marriage but before the time of the survey.³ The second contribution related to the literature on climate-induced migration is to analyse drought as a push factor driving different types of migration in Malawi, not only for work-related motives, but also for marriage as a way to smooth consumption of the family during droughts. Third, the analysis uses detailed information at the origin and destination districts over more than one and a half decade, exploiting the timing of the drought, unlike previous studies on Malawi. And fourth, it explores whether the type of migration changes depending on how far migrants move, comparing migration within the district of origin with migration between districts.

The remainder of the article is organized as follows. Section [1.2](#) describes bride price customs and migration patterns in Malawi. Section [1.3](#) presents the empirical strategy, section [1.4](#) describes the data used for the analysis and section [1.5](#) presents the main results. Robustness tests are discussed in section [1.6](#). Section [1.7](#) gives descriptive evidence on some of the outcomes of marriage-related migration, and Section [1.8](#) concludes.

³In their analysis, the rainfall shock is measured in the current residence recorded in the survey. Based on previous literature, they argue that when migration occurs, it takes place across relatively short distances.

1.2 Background: migration and marriage patterns in Malawi

The bride price is money or goods paid by a man to a girl's family to secure a bride, whether she is willing or not. In Malawi, *lobola*, the bride price, is customary in patrilineal tribes, frequent in the Northern part of the country, but also in some Southern districts.⁴ In the Karonga district for instance, the Nyakyusa tradition says that if a family is poor and low in status, it can go to a rich man and get a loan, a cow or money in exchange for their daughter. The practice is called *Kupimbira* and the girl child is therefore used as a surety. As these practices started to reach the news, creating concern among the population, the Marriage and Family Relations Act was passed on 12th February 2015, and promulgated into law by the President on 17th April 2015. The Act raised the legal age of marriage from a minimum age of 15 to 18 for statutory marriage, but under-age marriages keep happening. Also, Malawi's constitution allows for children between 15 and 18 to get married with their parents' and guardians' consent. Hence, a constitutional amendment would be necessary to ban child marriages completely by legal means. In a recent study supported by the UNICEF of almost 7500 households across the country (Makwemba et al., 2019), 9% of the interviewed women were married before the age of 15, while 42% married before the age of 18.⁵ Marriage below 15 is particularly widespread in the Southern region while marriage below 18 is highest in the Northern region.

Marriage is often associated with relocation to other villages or towns (Beegle and Poulin, 2013), but Englund (2002) shows that most marriages in Malawi are formed within the woman's district of origin and that most rural-rural migration is for marriages. If the bride price can be paid in several years, it could be that even those affected by the same drought inside the district of origin can pay the bride price later on, hence avoiding the demand side effect for which Hoogeveen et al. (2011) found evidence in rural Zimbabwe. In the data on Malawi, we note that indeed, women migrate predominantly within the district of origin for marriage, as seen in Table 1.1. Table 1.1 describes the different reasons to migrate reported by the household head (or the spouse of the household

⁴Patrilineal tribes generally practise patrilocal marriage, under which the wife moves to the husband's household. Matrilineal tribes are mostly found in Central and Southern Malawi. The Chewa, one of the main groups practising matrilineal customs, have seen large changes during recent decades towards a patrilineal system (Beegle and Poulin, 2013).

⁵The marriage rates of their male counterparts were 1% and 6% respectively.

head) for each individual, separated by between-district and within-district migrants.⁶

Table 1.1: Reasons to migrate for the different type of migrants

Reason to migrate (percentage of sample)	Between Dist-MIG			Within Dist-MIG		
	Men	Women	Total	Men	Women	Total
WORK RELATED/LOOK FOR WORK	18.1	5.3	23.4	8.9	2.7	11.6
SCHOOL/STUDIES	0.8	0.6	1.3	0.5	0.4	0.9
MARRIAGE	3.0	16.2	19.2	8.9	26.6	35.5
FAMILY REASONS/FOLLOW JOIN FAMILY	26.2	25.9	52.1	22.6	22.0	44.5
LAND/PLOT/OTHER LAND RELATED	1.3	0.7	2.0	1.7	1.1	2.8
OTHERS	0.6	1.4	2.1	2.2	2.5	4.7
Total	49.9	50.1	100.0	44.7	55.3	100.0

Note: created based on the LSMS-ISA information. A few migrants have missing information on the reason to migrate. Initial categories were re-classified in these 6 categories. Migrants established only one reason to migrate.

The categories were re-classified in only these six groups, to assemble categories that were more similar, such as those related with work or migrants moving for family reasons, for instance. The numbers here coincide with the migrants reported in Table 1.21 in Appendix 1E, with a few missing observations. The main reasons to migrate are: to join family, followed by marriage and work-related. For between district migrants, 23.4% move for work and 52.1% to join relatives, and 19.2% for marriage reasons. Among the within district migrants, migrating for family reasons-follow family is the main category with 44.5%, followed by marriage with 35.5% of migrants and for work with 11.6% (see Table 1.1). Among those moving for marriage in Malawi, two thirds correspond to within-district migrants while one third are between-district migrants (see Table 1.1).

Marriage-related migration, and migration in general, is only one coping mechanism for drought. Several ex ante risk-management and ex post risk-coping strategies are used by households living in risky environments. Among the main ex ante risk-management strategies are informal insurance through networks, self-insurance in the form of savings, and income diversification in the same locality. On the Malawi sample, solidarity mechanisms through transfers account for only five percent of total household income, on average, as shown in Table 1.23 in Appendix 1G. In the case of covariate shocks such as drought, ex ante informal insurance and local income diversification

⁶Some reasons to migrate could be related, for instance, a person could have migrated for marriage but would also be working with her husband on his land.

have been shown to be ineffective (Dercon, 2002, 2005). Ex post, drawing down on assets such as livestock may not be sufficient either (Kazianga and Udry, 2006), and in the absence of public safety nets, migration is often the main risk-coping mechanism in poor countries. Tables 1.2 and 1.3 provide indications that migration is used as a coping mechanism following drought. Table 1.2 shows that around 90 % of heads of households that receive migrants in the destination district work outside of agriculture. Table 1.3 provides further evidence that migrants tend to depend less on agriculture in the destination district: for between- and within-district migrants, 49 % and 64 % of total income is agricultural income, respectively, compared to the non migrant average of 78%. The descriptive statistics of households with no migrants depending more on agricultural income than households with migrants, together with the fact that migrants go to households in destination that have a higher percentage of non-agricultural income (see Table 1.24 in Appendix 1G), indicate that migrants pursue activities less affected by drought after moving and may use migration to mitigate drought impacts.

Table 1.2: Household head works in agricultural activities in the destination
(by type of migrant)

Type of Migrant	HH head NOT working in Agri.		HH head working in Agri.		Total	
	Observ.	Percentage	Observ.	Percentage	Observ.	Percentage
WORK RELATED/LOOK FOR WORK	2892	91.6	265	8.4	3157	100.0
SCHOOL/STUDIES	139	96.5	5	3.5	144	100.0
MARRIAGE	2182	90.5	229	9.5	2411	100.0
FAMILY REASONS/FOLLOW JOIN FAMILY	5230	92.4	433	7.6	5663	100.0
LAND/PLOT/OTHER LAND RELATED	93	82.3	20	17.7	113	100.0
OTHERS	205	89.5	24	10.5	229	100.0
Total	10741	91.7	976	8.3	11717	100.0

Note: The table includes only migrants for whom the information was reported.

Table 1.3: Percentage of agricultural income in the household by person in destination
(by type of migrant)

	Obs.	Mean	SD	Min	Max
NoMIG.	65912	0.78	0.30	0	1
Between Dist-MIG.	7391	0.49	0.42	0	1
Within Dist-MIG.	7968	0.64	0.38	0	1
International MIG.	1448	0.73	0.33	0	1
Total	82719	0.74	0.33	0	1

Note: The table includes only those for whom the information was reported. Only for survey 2004-2005 and 2010-2011.

1.3 Empirical strategy

Identification is based on a quasi-experimental design that compares the probability of migration for individuals in specific age groups having been affected or not by drought. Similar to the approach in Baez et al. (2017a), the identifying assumption is that individuals in different districts but in the same age category would have similar migration probabilities in the absence of drought, given controls for education, sex, and district fixed effects and time trends. Such a quasi-experimental identification strategy relies on differences in migration propensities across age groups (Hatton and Williamson, 2003). In particular, for a sample of African countries, including Malawi, FAO (2016) records that 60 to 70 % of migrants are between 15 and 34 years old. Previous analyses of climate and migration have also found important heterogeneity in the migratory responses according to age (Baez et al., 2017a,b). We estimate the following specification:

$$\begin{aligned} \mathbf{M}_{i,a,od,t,s} = & \alpha + \beta \times drought_{o,t} + \gamma_a \times drought_{o,t} \times age_a + \delta_a \times age_a \\ & + \eta \times X_{i,d,t} + \theta \times C_{od} + \phi_{s,i} + \psi_o + \psi_d + \lambda_t + \varepsilon_{i,a,od,t,s} \end{aligned} \quad (1.1)$$

The dependent variable $\mathbf{M}_{i,a,od,t,s}$ is equal to 1 when individual i of age group a , migrated from district o to district d at time t , as recorded in survey s , and zero for all the years when the individual did not move. The dummy variable "drought" equals 1 if a drought occurred in the district based on the Standardized Precipitation Index (SPI) constructed from satellite images (Funk et al., 2015). Age_a is a vector of five age groups at the time of migration (10 – 17, 18 – 27, 28 – 37, 38 – 47, 48 – 57, where we have excluded children below 10 and individuals above 57 years old); the coefficients δ_a thus capture age fixed effects that influence migration. $X_{i,d,t}$ is a vector of individual control variables such as sex and whether the individual has primary schooling or not.⁷ Migration costs are proxied by the distance between district o and district d , C_{od} . The specification also includes fixed effects by survey-time $\phi_{s,i}$, since a migrant is recorded in the survey in year s , but the migration took place in time t , time-invariant origin-district and destination-district characteristics (ψ_o and ψ_d) to capture additional attracting factors of the destination district d and push factors in the origin district o . λ_t are the time fixed effects. The variable $\varepsilon_{i,a,od,t,s}$ denotes the residuals,

⁷The index is d since data on characteristics are available only at destination, but they are measured in the year of migration.

and standard errors are clustered by district of origin.⁸

Previous analyses have shown that the impact of drought is ambiguous (Millock, 2015). On the one hand, $\beta > 0$ since droughts push migrants out of the district in search of better conditions for agriculture elsewhere, or make them migrate within-district to the non-agricultural sector. On the other hand, migration is costly, and between-district migration more so than within-district migration, and the effect of drought could be to decrease the probability of migration, i.e., $\beta < 0$ (Henry et al., 2004a,b). As regards the interaction with age, γ_a should indicate a positive probability of migration for the age groups 18 – 27 and 28 – 37, compared to the reference group of individuals aged between 48 and 57 years.⁹ The net effect of drought on the probability of migration of children in the age group 10-17 is given by the estimate of $\beta + \gamma_{10-17}$. The coefficient γ_{10-17} on its own measures the expected change following drought in the probability of migration of children in the age category 10-17 compared to the reference group aged 48-57:

$$E[Mig|drought = 1, age_{10-17}] - E[Mig|drought = 1, age_{48-57}] \\ - (E[Mig|drought = 0, age_{10-17}] - E[Mig|drought = 0, age_{48-57}])$$

The estimations are done both with a linear probability model and with a Probit model. The specification in Equation (1) assumes an idiosyncratic error term. Using a factor error model instead helps to control for any missing variables that could make the migratory response to drought vary over time and district. Examples of such common shocks could be low-level conflicts, the development of industry which would give alternative employment possibilities, or changes in infrastructure and policy that help cushion the impact of drought on agricultural households. In the case of Malawi, the Malawi Social Action Fund (MASAF) safety net program payments could be one such time-varying factor that is specific to some districts and that makes the drought response

⁸The attractiveness of alternative destinations will make the error terms for each migration pair correlated. By clustering the standard errors, we should control at least for the correlation in choices of destination districts by migrants from the same origin district (Fafchamps and Shilpi, 2013). In robustness tests we add estimations controlling for the main factor we are investigating - drought - at destination.

⁹Kernel estimations of the probability of migration show that this age category has the highest propensity to migrate, followed by the 10-17 age category. As shown in the Figure 1.5 in Appendix 1D, this high migration propensity is found in particular for reasons to join family at old age (Reason 4). For marriage-related migration (Reason 3), the density is the highest in the youngest age interval, then decreases slightly to stabilize before finally declining sharply after the age of 57 years.

interdependent across districts.¹⁰ In the OLS estimations, we will thus use the interactive fixed effects approach (Bai, 2009) to account for common correlated effects.

1.4 Data

1.4.1 Migration

We use the Living Standards Measurement Study-Integrated Surveys on Agriculture (LSMS-ISA), implemented by the World Bank, jointly with national governments of Sub-Saharan African countries. The data give detailed information of migration movements with the districts of destination and origin and are nationally representative of Malawi.¹¹ Since the country did not experience wars or severe conflict during the period analyzed, it enables us to identify more clearly the impact of drought on migration. The LSMS-ISA were conducted taking into account two agricultural seasons. As the rainy season in Malawi goes from November to May in general,¹² the Integrated Household Panel Survey (IHPS) 2013 was conducted between April 2013 (post-planting period) to October 2013 (post-harvest). As the time-line of the surveys shows (see Figure 1.4 in Appendix 1C), they were conducted in similar periods in order to capture the different agricultural seasons and make them comparable. This allows us to construct comparable measures at the level of district by year during the entire period.

The LSMS surveys provide two types of questions related to migration: a) in which region/district/country the person was born and b) in which region/district/country the person lived before moving there. Both questions are the same for the different waves, stating also the number of years the individual has been living there and how many years ago the migrant moved.¹³ We use option b) to define

¹⁰Since 1995, the MASAF aims at providing food security through a mix of cash transfers and subsidies to the use of fertilizers and other agricultural inputs during the lean season. As from 2012, the program was scaled up to cover about 500,000 households per year and to provide infrastructure (see Beegle et al. (2017) for a detailed evaluation of the MASAF program of 2012/2013).

¹¹We use the district classification of the LSMS-ISA surveys for Malawi, which did not have any change in the administrative organization during the different waves. There are 32 districts defined in the surveys (28 official districts + 4 cities), and we exclude the district island of Likoma, which is located in the lakes. Table 1.21 in Appendix 1E shows the district classification.

¹²According to the report of the Third Integrated Household Survey (IHS3), the start and end dates of rains vary spatially, but happen around this period for the majority of the country; however, earlier or later harvests are possible, depending on the type of crop, rainfall and other location and climatic conditions. By definition, the agricultural season includes harvest.

¹³We exclude temporal movements of migrants moving less than one year before the survey and those moving

migration. Table 1.4 shows the total number of migrants and non migrants by survey. Internal migration is divided into within-district migrants and between-district migrants.¹⁴ In Malawi, international migrants account for a small fraction of the total of migrants (around 3-5% of migrants in Table 1.4). The numbers of internal migrants are large, with a smaller share of total migration in 2016-2017 (around 21% of the total sample of individuals in that year, see Table 1.4). The size of within-district and between-district migration flows changes across the different waves, showing variability over time (see Table 1.4). The majority of movements have rural origin, rather than urban: 76.1% moved from villages in other districts (between-district) and 82.1% moved from villages in the same district (within-district).¹⁵

Table 1.4: Total Number of Migrants by survey

	2004-2005			2010-2011			2013			2016-2017		
	Total	%	% migrants	Total	%	% migrants	Total	%	% migrants	Total	%	% migrants
No Migrants	34283	65.0%		35752	72.4%		11122	62.9%		41729	78.8%	
Between District Migrants	7540	14.3%	40.9%	7243	14.7%	53.2%	3375	19.1%	51.4%	4710	8.9%	41.9%
Within District Migrants	10072	19.1%	54.6%	5651	11.4%	41.5%	3016	17.1%	45.9%	6140	11.6%	54.6%
International Migrants	821	1.6%	4.5%	715	1.4%	5.3%	175	1.0%	2.7%	388	0.7%	3.5%
Total	52716	100.0%		49361	100.0%		17688	100.0%		52967	100.0%	

Note: Own calculations based on LSMS surveys. LSMS Malawi 2004-2005 does not have information on district of origin for between district migrants. The table includes duplicate migrants appearing in more than one wave. They are dropped in the main analysis to arrive to a sample of migrants as in Table 1.21 in Appendix 1E

Although a migrant is recorded in the year of the survey s , her migration decision was taken based in year $s - t$ (the year of movement).¹⁶ As the surveys ask for the year of movement and record the year and month of the interview, we can track down more precisely the year of migration. In order to avoid double counting of migrants, we discard repeated cases when a person is recorded as migrant in wave s and in a wave before, but reported the same origin of district. Those are movements of the same person captured in more than one wave, so we will count only the last registration in the surveys. Table 1.21 in Appendix 1E has a more detailed description of the number of migrants used

before 2000 as the first survey was conducted in 2004-2005 and the questions of self-reported droughts in enumeration areas (EA) are for 5 years before.

¹⁴We use the question: [where did you move from?] 1. This village, 2. Other village in this district, 3. Village in other district, 4. This town or urban center, 5. Other town or urban centre in this district, 6. Town or urban centre in other district, 7. Outside Malawi. Between-district migrants correspond to options 3 and 6, within-district migrants to options 2 and 5, and international migrants to option 7.

¹⁵Data not shown here but available on request.

¹⁶There could be measurement error in the sense that a person said the movement was two years back, but the movement was two years and a half or more or less.

in the analysis by origin and by destination, which excludes those duplicate migrants (Table 1.4 includes them). Unfortunately, the LSMS-ISA 2004-2005 did not collect information on the district of origin for the between-district migrants and they cannot be included in the between-district analysis.

1.4.2 Drought measures

We measure drought using standardized indices calculated from gridded satellite data.¹⁷ Two measures are extensively used in the literature to capture droughts: the Standardized Precipitation Index (SPI) proposed by McKee et al. (1993) and the Standardized Precipitation-Evapotranspiration Index (SPEI) of Vicente-Serrano et al. (2010).

The SPI measures the number of standard deviations that observed cumulative precipitation deviates from the average and it can be calculated for any monthly time scale; it considers the long-term time series of precipitation accumulated over the desired time scale to estimate an appropriate probability density function. The method entails a transformation of the frequency distribution (e.g., Gamma, Pearson III, etc.) to another frequency distribution (normal, or Gaussian). A long period of data is necessary for the calculation (longer than 30 years is desirable). Compared to the SPI, the SPEI incorporates temperature data also. Thus, the SPEI is based on a climatic water balance determined by the difference between precipitation and potential evapotranspiration (PET) for each month in each location. It thus combines the changes in evapotranspiration demand of air temperature fluctuations and trends with the multiple temporality of the SPI. In both cases, the SPEI and SPI range from around -3 to 3, with lower values of the index being associated with more intense droughts and higher values with more rainfall.

A potential source of weather data to calculate both the SPI and the SPEI is the data from the Climatic Research Unit (CRU) of the University of East Anglia with monthly information since 1900 to 2016 and for an extensive sample of variables such as precipitation, vapor pressure, potential

¹⁷Table 1.22 in Appendix 1F shows statistics on the reports of the four most important events that happened in the enumeration areas (EA) based on the community surveys. It shows that droughts are the major events affecting the communities during this period. Based on the literature (Millock, 2015), we prefer to use the objective weather data to measure relative drought in a standardized manner across locations, rather than the community survey data. Table 1.5 shows the importance of using a standardized measure rather than the community survey measure.

evapotranspiration, etc. (Harris and Jones, 2017). Unfortunately, the resolution of this data is of only 0.5×0.5 degree - around $50\text{km} \times 50\text{km}$ at the Equator - and the construction of drought indices for the size of the districts in Malawi will not capture correctly spatial variability of the districts which have an average size of $54.3\text{km} \times 54.3\text{km}$.¹⁸ For these reasons, we use the satellite images of the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) which provide information of rainfall from 1981 to 2016 at a higher resolution of 0.05×0.05 degree - around $5\text{km} \times 5\text{km}$ at the Equator (Funk et al., 2015). Figure 1.1 in Appendix 1A also shows that the average temperature in the country stayed relatively stable during the period of analysis so there should be less need to adjust the drought measure by temperature changes and the SPI will be our preferred measure.¹⁹

We used the software ArcGIS and Google Earth Engine to manipulate the raster images, calculate the averages of precipitation at the level of district and by month and generate some additional maps with the help of the Python platform inside the software. The drought indices are calculated using the R-software and the package provided by Vicente-Serrano et al. (2010) and Beguería et al. (2014), which allows us to generate the SPI by months and for different locations (in this case, districts). Several monthly scales were calculated representing the number of months that the water deficits accumulate. Figure 1.2 in Appendix 1A shows the SPI index using the CHIRPS data for the monthly scale 12 and 24, where the annual observation corresponds to the average during the agricultural growing season, from November of the previous year until March of that year, as in Asfaw and Maggio (2017).²⁰

¹⁸Based on data from <http://www.statoids.com/umw.html> , the average area of the districts in Malawi is $2\,948.8\text{km}^2$ or $54.3\text{km} \times 54.3\text{km}$.

¹⁹An alternative index could be constructed by using CHIRPS data for precipitation and temperature from the MODIS-TERRA dataset (satellite MOD11A1.006 Terra Land Surface for Temperature and Emissivity Daily Global) at 0.01×0.01 degree - around $1\text{km} \times 1\text{km}$. The construction of the potential evapotranspiration could add noise to the estimation, however, if it does not include also information of wind speed and solar radiation. This gives additional arguments for why we prefer to use the SPI as the main measure to capture droughts.

²⁰A monthly scale of 12 goes in line better with the period of time when the surveys were collected, see Figure 1.4 in Appendix 1C. Both the SPI with scale of 12 and 24 were used in the empirical analysis, but the tables below present only the results using the SPI-12 measure, since shorter time scales capture better agricultural drought.

1.4.3 Socio-economic characteristics and other covariates

The analysis includes controls for the sex of the migrant (one for female, zero otherwise) and whether the individual has primary education or not. It also includes a variable of the distance between the district of origin and the district of destination as a proxy for the migration costs that the migrant would have faced. We use the Vincenty measure of [Nichols \(2003\)](#) for the distance between districts. This measure takes into account geodesic distances between a pair of points on the surface of the Earth (specified in signed decimal degrees latitude and longitude), using an ellipsoidal model of the Earth. Alternative measures using the Euclidean distance gave very close results.

1.4.4 Summary statistics

The final sample consists of 127,000 individual-year observations of between-district migration and 182,000 individual-year observations of within-district migration.²¹ [Table 1.25](#) in [Appendix 1G](#) presents the balance table for the data measuring drought using the SPI for the 309,000 individual-year observations of the sample. The migrants in a drought year and the migrants in a non-drought year display no significant differences apart from rainfall, which we control for in the estimations using the drought indicator. This holds for both between-district and within-district migration.

1.5 Results

Before investigating the gender effects of migration, we start by showing the aggregate response of migration in Malawi to weather shocks, which goes in line with previous work on liquidity constraints for migration. [Table 1.5](#) shows all migration, for both men and women, with the drought measure based on the SPI presented in columns (1) and (2) and the measure based on the community surveys in columns (3) and (4). We show only the results from the estimations using interactive fixed effects, since the purpose of [Table 1.5](#) is to underline the difference in conclusions between the two drought measures.²² The estimations based on the community survey measure show a significant increase in the probability of migration when drought was reported by the community chief in at

²¹Although the total sample size of migrants is around 30,000 for 16 years (480,000), we reach 309,000 individual-year observations due to missing observations in the explanatory variables.

²²[Appendix 1L](#) describes the type of measurement error that could arise when self-reported measures of droughts are used.

least 20% (or 10%) of the enumeration areas in the origin district. This push effect of drought is not different according to age category (columns (3) and (4)). Using the SPI measure of drought instead shows evidence that drought decreases the probability of migration, in line with evidence on liquidity traps (Bazzi, 2017; Hirvonen, 2016), but only for moderate droughts (SPI below -1) which are more frequent in Malawi according to the data (column (1)). These results also show differences according to age category, with a small positive net effect in the age interval 28-37 years, and a negative or marginally significant net effect for other age categories. Since the drought measure based on community surveys in the enumeration area is less precise, due to the need to aggregate the measure at the district level, and the threshold for drought may seem ad hoc (10 or 20 %), we prefer to interpret the results based on the SPI measure only, from now on. The results based on this measure are also in line with previous evidence of liquidity constraints for migration in Malawi (Lewin et al., 2012) and in Tanzania (Hirvonen, 2016).

Our specific objective is to understand how women’s migration respond to drought and, in particular, to investigate the risk of child marriages following drought. Table 1.6 shows the probability of female migration overall, without distinguishing between- and within-district migration, and for all reasons of migration. The sample is now 166,014 migrants, around half of the full migrant sample. The explanatory power of the OLS interactive fixed effects model is 14.6 %.²³ The proportion of correctly predicted outcomes in the Probit model is 94 %. Table 1.6 shows that the general effect of drought on the reference group of women aged 48 to 57 years is to constrain migration rather than increasing it. When adding the marginal effect²⁴ of each age category compared to the reference group, we see small increases in the probability of migration, in particular in the older age intervals for moderate drought, but with significant effects also in the age intervals 10-17 and 18-27 for severe droughts (with the SPI below -2). The effects are not large, though. Based on the marginal effects of the probit estimation (columns (3) and (4)), the increase in the probability of migration following a severe drought is 0.35 percentage points in the age intervals 18-27 and 0.18 in the age interval 28-37. For girls in the ages 10-17, the marginal effect of a severe drought is a 0.33

²³Estimations with standard OLS have very low explanatory power, indicating potential missing variables. Using OLS with interactive fixed effects enables us to control better for district-specific time trends and heterogeneity.

²⁴As probit estimations correspond to non-linear models, marginal effects are not constant over the sample. In this case, the marginal effects in all the probit estimations are calculated for each observation and then averaged (average marginal effect).

Table 1.5: Between and Within District Migration

	SPI-Cumulative drought for 12 months		Drought using EA	
	OLS Interactive FE		OLS Interactive FE	
	(1)	(2)	(3)	(4)
spi12-drought-agro (< -1)	-0.0111** (0.00531)			
spi12-drought-agro (< -1) x age (10-17)	0.00943 (0.00691)			
spi12-drought-agro (< -1) x age (18-27)	0.0120* (0.00609)			
spi12-drought-agro (< -1) x age (28-37)	0.0136** (0.00615)			
spi12-drought-agro (< -1) x age (38-47)	0.0167* (0.00825)			
spi12-drought-agro (< -2)		-0.0139 (0.0101)		
spi12-drought-agro (< -2) x age (10-17)		0.0159* (0.00924)		
spi12-drought-agro (< -2) x age (18-27)		0.0180** (0.00823)		
spi12-drought-agro (< -2) x age (28-37)		0.0173 (0.0111)		
spi12-drought-agro (< -2) x age (38-47)		0.0190 (0.0126)		
drought1 (10%)			0.0914*** (0.0121)	
drought1 (10%) x age (10-17)			0.0169 (0.0126)	
drought1 (10%) x age (18-27)			0.00974 (0.0127)	
drought1 (10%) x age (28-37)			0.0203 (0.0149)	
drought1 (10%) x age (38-47)			0.0180 (0.0176)	
drought2 (20%)				0.0946*** (0.0149)
drought2 (20%) x age (10-17)				0.0216 (0.0156)
drought2 (20%) x age (18-27)				0.0123 (0.0166)
drought2 (20%) x age (28-37)				0.0188 (0.0174)
drought2 (20%) x age (38-47)				0.0256 (0.0186)
Observations	309510	309510	300636	300636
R^2	0.147	0.147	0.163	0.159
R^2_a	0.147	0.147	0.163	0.159
Ncluster	31	31	31	31

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Includes dummies for year, district of origin, district of destination and for survey, as well as controls for education and sex (in the case it is relevant).

Enumeration area is the lower level of district in Malawi.

drought1 (10%) is a dummy equal to one if a drought is reported in 10% or more of the enumeration areas.

drought2 (20%) is a dummy equal to one if a drought is reported in 20% or more of the enumeration areas.

spi12-drought-agro (< -1) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -1 during at least one month of the agricultural growing season.

spi12-drought-agro (< -2) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -2 during at least one month of the agricultural growing season.

Baseline category of age: 48-57. Age <10 and Age >57 are excluded. Distance or Vicenty-distance calculated using geodesic distances between a pair of points on the surface of the Earth. Standard errors clustered at origin-district. Standard deviations in parenthesis.

percentage point increase in the probability of migration, and only a 0.09 percentage point increase following moderate drought.

Figure 1.3 in Appendix 1B shows the marginal effects of women's and men's migration for marriage following moderate drought. Whereas there are no differences between migrants from rural versus urban districts, one clearly sees the higher probability of migration for individuals with pri-

mary education compared to those without such education. But above all, Figure 1.3 in Appendix 1B shows the small increase in women’s marriage-related migration following drought and the even more modest impact of drought on men’s marriage-related migration, with no evidence of a different impact on marriage-related migration in the lowest age group of women compared to the older age intervals. If we constrain the sample to women’s marriage-related migration only, there is a significant increase in the probability of marriage-related migration in the age intervals above 18 years following moderate drought, ranging from 0.3 to 0.9 percentage points (Table 1.7). But any hypothesis of an increase in marriage-related migration of children following drought can be rejected according to these estimates, since we actually see a small but significant net reduction of 0.35 percentage points in the probability of marriage-related migration in the age interval 10-17.

To investigate this result further, we separate marriage-related migration outside and within the woman’s district of origin in Table 1.8. Here we see strong evidence of the effect coming only from within-district migration, with little effect of drought on between-district migration for marriage. This goes in line with the literature showing that women in Malawi migrate for marriage mainly within the district of origin (Beegle and Poulin, 2013). The marginal effects from the Probit model show a decrease in marriage-related within-district migration of children in the age interval 10-17 of 0.45 percentage points following moderate drought (column (7)). The negative effect of drought on marriage-related migration of children could potentially be interpreted as evidence of a marriage-market equilibrium effect. If marriage-related migration is mainly local, and all households are affected by a covariate weather shock, then fewer grooms will be able to pay the bride price.²⁵ This interpretation is supported by the fact that between-district migration displays no significant change following drought in the same age-category. There is also a reduction by 0.62 percentage points in the probability of marriage-related migration within the district for women in the age interval 28-37 following moderate drought. The probability of within-district migration for marriage increases in the other age categories, with the marginal effect ranging from 0.10 to 1.35 percentage points.

²⁵Corno et al. (2020) show the conditions for a net decrease in child marriages in marriage markets with bride price: the income elasticity of demand should exceed the income elasticity of supply, or the price elasticity of supply should exceed that of demand in case that the income elasticities are equal. Unfortunately, we only have data on marriage payments for a few hundred migrants in the sample, and cannot test any effect on bride prices.

Table 1.6: Between and Within District Migration - Women

	SPI-Cumulative drought for 12 months			
	OLS Interactive FE		PROBIT-MG EFFECTS	
	(1)	(2)	(3)	(4)
spi12-drought-agro (< -1)	-0.0178*		-0.0241**	
	(0.00900)		(0.0114)	
spi12-drought-agro (< -1) x age (10-17)	0.0174		0.0250**	
	(0.0109)		(0.0115)	
spi12-drought-agro (< -1) x age (18-27)	0.0217**		0.0273**	
	(0.00920)		(0.0113)	
spi12-drought-agro (< -1) x age (28-37)	0.0229**		0.0286***	
	(0.00967)		(0.0107)	
spi12-drought-agro (< -1) x age (38-47)	0.0280**		0.0376***	
	(0.0126)		(0.0131)	
spi12-drought-agro (< -2)		-0.0121		-0.0427**
		(0.0171)		(0.0208)
spi12-drought-agro (< -2) x age (10-17)		0.0128		0.0460**
		(0.0189)		(0.0209)
spi12-drought-agro (< -2) x age (18-27)		0.0160		0.0462**
		(0.0167)		(0.0227)
spi12-drought-agro (< -2) x age (28-37)		0.0176		0.0445**
		(0.0161)		(0.0192)
spi12-drought-agro (< -2) x age (38-47)		-0.000528		0.0488**
		(0.0210)		(0.0226)
Observations	166014	166014	166014	166014
R2	0.146	0.146		
$R2_a$	0.146	0.146		
$R2_p$			0.0113	0.0112
Ncluster	31	31	31	31
Pcorr	.	.	94.44	94.44

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Includes dummies for year, district of origin, district of destination and for survey, as well as controls for education and sex (in the case it is relevant).

spi12-drought-agro (< -1) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -1 during at least one month of the agricultural growing season.

spi12-drought-agro (< -2) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -2 during at least one month of the agricultural growing season.

Baseline category of age: 48-57. Age <10 and Age >57 are excluded. Distance or Vicenty-distance calculated using geodesic distances between a pair of points on the surface of the Earth. Standard errors clustered at origin-district. Standard deviations in parenthesis.

These results should be compared with the effect of drought on women's work-related migration. The sample is now much smaller, since women migrate mainly for marriage. Table 1.9 shows that only severe drought increases women's work-related migration (column (4)). The marginal net effect for girls in the age interval 10-17 is a 5.1 percentage point increase in the probability of work-related migration between districts. For all other age intervals the marginal effect of drought is to reduce

Table 1.7: Between and Within District Migration - Women migrating for marriage (reason 3)

	SPI-Cumulative drought for 12 months			
	OLS Interactive FE		PROBIT-MG EFFECTS	
	(1)	(2)	(3)	(4)
spi12-drought-agro (< -1)	-0.0319*** (0.0102)		-0.0450*** (0.0160)	
spi12-drought-agro (< -1) x age (10-17)	0.0283*** (0.00921)		0.0415*** (0.0142)	
spi12-drought-agro (< -1) x age (18-27)	0.0357*** (0.0101)		0.0483*** (0.0159)	
spi12-drought-agro (< -1) x age (28-37)	0.0401*** (0.0110)		0.0539*** (0.0162)	
spi12-drought-agro (< -1) x age (38-47)	0.0403** (0.0168)		0.0545*** (0.0202)	
spi12-drought-agro (< -2)		-0.0261 (0.0182)		-0.0415 (0.0278)
spi12-drought-agro (< -2) x age (10-17)		0.0210 (0.0196)		0.0378 (0.0295)
spi12-drought-agro (< -2) x age (18-27)		0.0336 (0.0209)		0.0477 (0.0311)
spi12-drought-agro (< -2) x age (28-37)		0.0418** (0.0197)		0.0562** (0.0276)
spi12-drought-agro (< -2) x age (38-47)		0.0424 (0.0250)		0.0580* (0.0314)
Observations	94104	94104	94104	94104
R2	0.148	0.148		
$R2_a$	0.147	0.147		
$R2_p$			0.0122	0.0120
Ncluster	31	31	31	31
Pcorr	.	.	94.44	94.44

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Includes dummies for year, district of origin, district of destination and for survey, as well as controls for education and sex (in the case it is relevant).

spi12-drought-agro (< -1) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -1 during at least one month of the agricultural growing season.

spi12-drought-agro (< -2) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -2 during at least one month of the agricultural growing season.

Agricultural growing season from November to March. Baseline category of age: 48-57. Age <10 and Age >57 are excluded. Distance or Vicenty-distance calculated using geodesic distances between a pair of points on the surface of the Earth. Standard errors clustered at origin-district. Standard deviations in parenthesis.

the likelihood of migration. This result indicates potential migration of children for work reasons or, possibly, a mix of marriage and work-motivated reasons for women's migration, something we probe further in Table 1.11.

Table 1.9 also shows a similar pattern for men's between-district migration for work reasons,

Table 1.8: Between and Within District Migration (split sample)
Women migrating for marriage (reason 3)

	SPI-Cumulative drought for 12 months (Between)				SPI-Cumulative drought for 12 months (Within)			
	OLS Interactive FE		PROBIT-MG EFFECTS		OLS Interactive FE		PROBIT-MG EFFECTS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
spi12-drought-agro (< -1)	-0.0416** (0.0202)		-0.0439 (0.0325)		-0.0358*** (0.00977)		-0.0461*** (0.0162)	
spi12-drought-agro (< -1) x age (10-17)	0.0435* (0.0220)		0.0437 (0.0326)		0.0305*** (0.00831)		0.0416*** (0.0143)	
spi12-drought-agro (< -1) x age (18-27)	0.0431** (0.0184)		0.0487 (0.0296)		0.0365*** (0.00958)		0.0471*** (0.0162)	
spi12-drought-agro (< -1) x age (28-37)	0.0618*** (0.0205)		0.0670** (0.0300)		0.0300** (0.0115)		0.0399** (0.0173)	
spi12-drought-agro (< -1) x age (38-47)	0.0350 (0.0241)		0.0434 (0.0347)		0.0495** (0.0186)		0.0596*** (0.0227)	
spi12-drought-agro (< -2)		0.0547 (0.105)		-0.0168 (0.0599)		-0.0343** (0.0126)		-0.0526 (0.0354)
spi12-drought-agro (< -2) x age (10-17)		-0.0967 (0.0985)		0.00348 (0.0546)		0.0359* (0.0197)		0.0552 (0.0424)
spi12-drought-agro (< -2) x age (18-27)		-0.0536 (0.0982)		0.0176 (0.0549)		0.0432** (0.0169)		0.0608 (0.0398)
spi12-drought-agro (< -2) x age (28-37)		-0.0488 (0.101)		0.0246 (0.0552)		0.0498*** (0.0174)		0.0673* (0.0360)
spi12-drought-agro (< -2) x age (38-47)		-0.0639 (0.0837)		0.0179 (0.0537)		0.0585** (0.0215)		0.0753* (0.0392)
Observations	28728	28728	28728	28728	65376	65376	65376	65376
R2	0.175	0.175			0.166	0.166		
R2 _a	0.172	0.172			0.165	0.165		
R2 _p			0.0200	0.0192			0.0174	0.0173
Ncluster	31	31	31	31	31	31	31	31
Pcorr	.	.	94.44	94.44	.	.	94.44	94.44

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Includes dummies for year, district of origin, district of destination and for survey, as well as controls for education and sex (in the case it is relevant).

spi12-drought-agro (< -1) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -1 during at least one month of the agricultural growing season.

spi12-drought-agro (< -2) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -2 during at least one month of the agricultural growing season.

Agricultural growing season from November to March. Baseline category of age: 48-57. Age<10 and Age>57 are excluded. Distance or Vicenty-distance calculated using geodesic distances between a pair of points on the surface of the Earth. Standard errors clustered at origin-district. Standard deviations in parenthesis.

which also increases following drought in the youngest age interval (columns (7) and (8)).²⁶ Following a moderate drought, the work-related migration of boys increases by 3.6 percentage points, and by 8.7 percentage points following severe drought. The finding of an increase in the migration of children for work reasons following drought adds to the evidence of the negative social consequences of drought. It could potentially indicate that child labour is used as an ex post coping mechanism

²⁶The effects of drought on men's migration for work is entirely driven by between-district migration, as drought never has a significant effect on men's within-district migration (see additional Table 1.26 in Appendix 1H). The only other significant effect found on men's between-district migration for work is a smaller increase in the probability of work-related migration (+2.8 percentage points) following severe drought for men in the age interval 38-47 (column (8) in Table 1.9).

following drought.²⁷

Table 1.9: Between District Migration - Women and men migrating for work (reason 1)

	SPI-Cumulative drought for 12 months (Women)				SPI-Cumulative drought for 12 months (Men)			
	OLS Interactive FE		PROBIT-MG EFFECTS		OLS Interactive FE		PROBIT-MG EFFECTS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
spi12-drought-agro (< -1)	-0.0130 (0.0270)		-0.0211 (0.0318)		-0.000853 (0.0181)		-0.00616 (0.0162)	
spi12-drought-agro (< -1) x age (10-17)	-0.0121 (0.0302)		0.0173 (0.0379)		0.0193 (0.0244)		0.0422** (0.0198)	
spi12-drought-agro (< -1) x age (18-27)	0.00847 (0.0281)		0.0140 (0.0311)		-0.00201 (0.0202)		0.00380 (0.0171)	
spi12-drought-agro (< -1) x age (28-37)	0.0122 (0.0281)		0.0213 (0.0341)		0.00713 (0.0177)		0.0120 (0.0158)	
spi12-drought-agro (< -1) x age (38-47)	0.0186 (0.0380)		0.0258 (0.0393)		0.00356 (0.0201)		0.0201 (0.0166)	
spi12-drought-agro (< -2)		-0.0418 (0.0510)		-0.418*** (0.0122)		-0.0421 (0.0320)		-0.0339 (0.0267)
spi12-drought-agro (< -2) x age (10-17)		0.0201 (0.0702)		0.469*** (0.0397)		0.143* (0.0789)		0.121*** (0.0317)
spi12-drought-agro (< -2) x age (18-27)		0.0314 (0.0552)		0.407*** (0.0164)		0.0454 (0.0318)		0.0404* (0.0242)
spi12-drought-agro (< -2) x age (28-37)		0.00723 (0.0522)		0.386*** (0.0143)		0.0446 (0.0287)		0.0359 (0.0247)
spi12-drought-agro (< -2) x age (38-47)		0.0250 (0.0626)		0.387*** (0.0350)		0.0530 (0.0387)		0.0621** (0.0295)
Observations	9522	9522	9522	9522	31914	31914	31914	31914
R2	0.183	0.183			0.183	0.183		
R2 _a	0.175	0.175			0.181	0.181		
R2 _p			0.0379	0.0401			0.0266	0.0275
Ncluster	31	31	31	31	31	31	31	31
Pcorr	.	.	94.44	94.44	.	.	94.44	94.44

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Includes dummies for year, district of origin, district of destination and for survey, as well as controls for education and sex (in the case it is relevant).

spi12-drought-agro (< -1) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -1 during at least one month of the agricultural growing season.

spi12-drought-agro (< -2) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -2 during at least one month of the agricultural growing season.

Agricultural growing season from November to March. Baseline category of age: 48-57. Age<10 and Age>57 are excluded. Distance or Vicenty-distance calculated using geodesic distances between a pair of points on the surface of the Earth. Standard errors clustered at origin-district. Standard deviations in parenthesis.

To check that it is mainly women's marriage-related migration that is affected by drought, Table 1.10 shows that men's marriage-related migration across districts is affected only by severe drought. Note the small sample size, though, for male marriage-related migration between districts (5,382 individuals). These estimations cannot be used to test a marriage-market equilibrium effect either, since they concern male migration for marriage-related reasons, which would not be compatible

²⁷In order to fully establish such an effect, we would need to check the labour status of the same children in the place of origin. Unfortunately, the panel of individuals built on the three surveys used here has a much smaller sample size, so does not enable a relevant comparison.

with patrilocality and bride price customs (under which it is the bride who moves to the groom's family). The small number of men moving between-district for marriage compared to women - as well as within district - reflects the prevalence of patrilocality today in Malawi. [Robinson and Gottlieb \(2019\)](#) notes that the Chewa in Northern Malawi and in the Western parts of Central Malawi, despite being matrilineal traditionally, practice patrilocality today. We do note much smaller effects of drought on men's marriage-related migration, though: a net increase of 0.9 percentage points following severe drought in the age interval 28-37 only (column (8)), with drought reducing the probability of marriage-related migration for all other age intervals. The net effect on marriage-related migration of boys in the age interval 10-17 is also close to zero, with a reduction of 0.7 percentage points. Most marriage-related migration occurs within districts, though. [Table 1.27](#) in [Appendix 1H](#) compares the previous results on women's marriage-related migration within the district of origin to that of men's, and finds no significant effect on men's marriage-related migration within the district of origin.

Finally, since the motives for migration may be mixed, and not clearly reported by some household heads (who may be reluctant to give marriage as a reason for under-age children), we do estimations on all female migration related to either marriage or work reasons in [Table 1.11](#). The estimation results show that severe droughts may increase women's migration for both reasons, but only for between-district migration (column (4)), as expected if a drought-coping mechanism is at work. The effect is large in the age interval 10-17, with a net increase of 2.23 percentage points in the probability of work- and marriage-related between-district migration of women. The only other age interval that displays an increase in the probability of work- and marriage-related migration to another district is the age interval 38-47 years, whereas severe drought reduces the probability of migration for all other age groups. The results could indicate mixed reasons for marriage and work on the husband's farm, but we cannot disentangle the results further. It is thus only when combining the work and marriage-related motives that we find a sizeable increase in the probability of migration of girls aged 10-17 following drought.

Table 1.10: Between District Migration - Women and men migrating for marriage (reason 3)

	SPI-Cumulative drought for 12 months (Women)				SPI-Cumulative drought for 12 months (Men)			
	OLS Interactive FE		PROBIT-MG EFFECTS		OLS Inbteractive FE		PROBIT-MG EFFECTS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
spi12-drought-agro (< -1)	-0.0416** (0.0202)		-0.0439 (0.0325)		-0.00756 (0.0358)		-0.00525 (0.0394)	
spi12-drought-agro (< -1) x age (10-17)	0.0435* (0.0220)		0.0437 (0.0326)		-0.0220 (0.0343)		-0.0202 (0.0410)	
spi12-drought-agro (< -1) x age (18-27)	0.0431** (0.0184)		0.0487 (0.0296)		0.0187 (0.0371)		0.0222 (0.0415)	
spi12-drought-agro (< -1) x age (28-37)	0.0618*** (0.0205)		0.0670** (0.0300)		0.0305 (0.0333)		0.0234 (0.0349)	
spi12-drought-agro (< -1) x age (38-47)	0.0350 (0.0241)		0.0434 (0.0347)		0.0116 (0.0456)		0.0120 (0.0480)	
spi12-drought-agro (< -2)		0.0547 (0.105)		-0.0168 (0.0599)		-0.123*** (0.0298)		-0.418*** (0.0183)
spi12-drought-agro (< -2) x age (10-17)		-0.0967 (0.0985)		0.00348 (0.0546)		0.0788 (0.0793)		0.411*** (0.0564)
spi12-drought-agro (< -2) x age (18-27)		-0.0536 (0.0982)		0.0176 (0.0549)		0.0646** (0.0262)		0.392*** (0.0215)
spi12-drought-agro (< -2) x age (28-37)		-0.0488 (0.101)		0.0246 (0.0552)		0.141*** (0.0473)		0.427*** (0.0273)
spi12-drought-agro (< -2) x age (38-47)		-0.0639 (0.0837)		0.0179 (0.0537)		0.0351 (0.0494)		0.375*** (0.0552)
Observations	28728	28728	28728	28728	5382	5382	5382	5382
R2	0.175	0.175			0.151	0.152		
R2 _a	0.172	0.172			0.136	0.137		
R2 _p			0.0200	0.0192			0.0190	0.0193
Ncluster	31	31	31	31	31	31	31	31
Pcorr	.	.	94.44	94.44	.	.	94.44	94.44

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Includes dummies for year, district of origin, district of destination and for survey, as well as controls for education and sex (in the case it is relevant).

spi12-drought-agro (< -1) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -1 during at least one month of the agricultural growing season.

spi12-drought-agro (< -2) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -2 during at least one month of the agricultural growing season.

Agricultural growing season from November to March. Baseline category of age: 48-57. Age<10 and Age>57 are excluded. Distance or Vicenty-distance calculated using geodesic distances between a pair of points on the surface of the Earth. Standard errors clustered at origin-district. Standard deviations in parenthesis.

1.6 Robustness tests

In this section, we start by giving descriptive evidence on the marital status of women who migrated, in order to support the use of the self-reported reasons for migration as reflecting actual migration for marriage. To test the robustness of the results, we then present estimations including also drought in the destination district, potential household interaction effects, and the results from a multinomial model of the different migration choices. We also discuss another potential mechanism behind the results, the role of ethnicity as reason for marriage.

Table 1.11: Between and Within District Migration (split sample)
Women migrating for marriage (reason 3) and work (reason 1)

	SPI-Cumulative drought for 12 months (Between)				SPI-Cumulative drought for 12 months (Within)			
	OLS Interactive FE		PROBIT-MG EFFECTS		OLS Interactive FE		PROBIT-MG EFFECTS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
spi12-drought-agro (< -1)	-0.00850 (0.0158)		-0.0129 (0.0148)		-0.0153* (0.00786)		-0.0152 (0.00930)	
spi12-drought-agro (< -1) x age (10-17)	0.0101 (0.0171)		0.0200 (0.0149)		0.0101 (0.00859)		0.0105 (0.00961)	
spi12-drought-agro (< -1) x age (18-27)	0.00876 (0.0164)		0.0145 (0.0144)		0.0143* (0.00708)		0.0144* (0.00852)	
spi12-drought-agro (< -1) x age (28-37)	0.0168 (0.0152)		0.0229* (0.0138)		0.0116* (0.00580)		0.0112* (0.00657)	
spi12-drought-agro (< -1) x age (38-47)	0.0113 (0.0179)		0.0219 (0.0155)		0.0243*** (0.00831)		0.0242*** (0.00912)	
spi12-drought-agro (< -2)		-0.0340 (0.0261)		-0.0470** (0.0223)		0.0000230 (0.0131)		0.00136 (0.0139)
spi12-drought-agro (< -2) x age (10-17)		0.00531 (0.0272)		0.0693*** (0.0221)		0.00119 (0.0152)		0.000966 (0.0160)
spi12-drought-agro (< -2) x age (18-27)		0.0328 (0.0254)		0.0469** (0.0200)		0.0109 (0.0141)		0.00940 (0.0151)
spi12-drought-agro (< -2) x age (28-37)		0.0319 (0.0204)		0.0458** (0.0183)		0.00826 (0.0105)		0.00656 (0.0111)
spi12-drought-agro (< -2) x age (38-47)		0.0343 (0.0264)		0.0583*** (0.0221)		0.0327* (0.0162)		0.0266* (0.0155)
Observations	75546	75546	75546	75546	115038	115038	115038	115038
R2	0.176	0.176			0.166	0.166		
R2 _a	0.175	0.175			0.165	0.165		
R2 _p			0.0226	0.0226			0.0179	0.0181
Ncluster	31	31	31	31	31	31	31	31
Pcorr	.	.	94.44	94.44	.	.	94.44	94.44

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Includes dummies for year, district of origin, district of destination and for survey, as well as controls for education and sex (in the case it is relevant).

spi12-drought-agro (< -1) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -1 during at least one month of the agricultural growing season.

spi12-drought-agro (< -2) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -2 during at least one month of the agricultural growing season.

Agricultural growing season from November to March. Baseline category of age: 48-57. Age<10 and Age>57 are excluded. Distance or Vicenty-distance calculated using geodesic distances between a pair of points on the surface of the Earth. Standard errors clustered at origin-district. Standard deviations in parenthesis.

1.6.1 Marital status after migration corroborates the stated reason for migration

As support of the self-reported data on the reason for migration, the descriptive statistics in Tables 1.12 and 1.13 show that the marital status in the destination is monogamous marriage for the majority of migrants (58%), with the second largest category being polygamous marriage (3.8%). For those who declared marriage as reason to migrate, 92% are married in monogamous relationships (Table 1.13). Compared to non migrants, migrants are more likely to be married (Table 1.12).

Table 1.12: Marital status in destination- By category

	Monogamous	Polygamous	Separated	Divorced	Widowed	Never married	Total
	Observ.	Observ.	Observ.	Observ.	Observ.	Observ.	Observ.
NoMIG.	25643	2540	2145	2112	3801	28848	65089
Between Dist-MIG.	6001	259	301	228	293	4044	11126
Within Dist-MIG.	7608	643	238	286	297	3258	12330
International MIG.	1072	175	60	69	249	262	1887
Total	40324	3617	2744	2695	4640	36412	90432

Note: The table includes only those for whom the information was reported. The question asked is: What is [NAME]'s present marital status?.

Table 1.13: Marital status in destination- By reason to migrate

	Monogamous	Polygamous	Separated	Divorced	Widowed	Never married	Total
	Observ.	Observ.	Observ.	Observ.	Observ.	Observ.	Observ.
WORK RELATED/LOOK FOR WORK	3712	153	210	143	196	716	5130
SCHOOL/STUDIES	47	2	4	2	1	242	298
MARRIAGE	7597	556	38	17	66	14	8288
FAMILY REASONS/FOLLOW JOIN FAMILY	1281	88	181	145	187	6184	8066
LAND/PLOT/OTHER LAND RELATED	547	61	21	22	37	18	706
OTHERS	415	41	85	184	102	122	949
Total	13599	901	539	513	589	7296	23437

Note: The table includes only those for whom the information was reported. The question asked is: What is [NAME]'s present marital status?.

1.6.2 The influence of drought in the destination district

The main tables relied on drought as a push factor, acting only in the place of origin. Drought in the destination district could deter migration, though, under the assumption that migrants are informed about drought in potential destinations. Table 1.28 in Appendix II tests the robustness of the main results of women migrating for marriage between-district when including jointly drought in the origin and in the destination districts. Recall that women migrate mainly within district for marriage in Malawi. Since we do not have the location of the villages themselves but use district-level drought data, we can only test this specification on between-district migration, and the sample is thus identical to the between-district migration sample in Table 1.8. The marginal effects of severe drought are insignificant, just as in the main estimation with drought in the origin district only (column (4) in Table 1.8). The effects of moderate drought at origin are similar in terms of statistical significance and also in magnitude for the age group with a significant response (28-37 years). We conclude that the main results are robust to controlling for drought at destination. In addition, we note that drought at destination has the expected net deterring effect on migration for the reference group, and most other age intervals (except for the age group 38-47 years).

As an additional robustness check, we have tested a more general specification including destination district by year fixed effects to control for varying conditions in the destination districts more generally (Table 1.29 in Appendix II). The results in Table 1.29 correspond to the main results in Table 1.7. They are robust for all other age categories, but indicate a very small net increase in the probability of migration of girls (0.04 percentage points).

1.6.3 Household interaction effects

In the main estimations, we estimate the effect of drought on individual moves, but they could be interrelated if two or more persons in the household move at the same time. In Table 1.30 in Appendix IJ, we show estimations with interactions for whether the household head and partner migrated together. For moderate drought, the results show very similar results for women's migration and no significant interaction effect, so compared to Table 1.7 the main results do not change. For severe drought, though, it is interesting to note that including this control increases the probability

of migrating among the older age groups by 1 to 1.5 percentage points and that the interaction effect with the dummy has a negative effect. There is still no effect on the youngest age category, though.

Other potentially relevant effects could come from the presence of brothers and sisters (Corno et al., 2020) or the relation of the individual to the household head (Beegle and Poulin, 2013), or characteristics of the household, such as the household wealth or the level of education of the household head. Unfortunately, these interactions could not be investigated on the data used here since such characteristics are recorded in the destination district, and not at origin. A more detailed investigation of such mechanisms and their role in explaining the results found here should be the subject of future research.

1.6.4 Results from a multinomial model

The different destinations, within and between districts, could be considered outcomes of the same utility maximization problem and dependent on migration costs. To test the robustness of our main results on women migrating for marriage (Table 1.8), Table 1.31 in Appendix 1K presents the results from a multinomial estimation on the same sample adding the non migrants and accounting for the two migration destination choices simultaneously. As in the main results, it is moderate drought that affects migration for marriage rather than severe drought, and the effects are found only for marriage-related migration within the district of origin. It is only in the age interval 28-37 years that drought decreases also women's between-district migration, but the marginal effects are small, less than one percentage point. The main result of a small decrease following drought in the migration propensity for marriage of girls who are between 10 and 17 years old is robust. In fact, the decrease is now even smaller, bordering on zero, indicating no evidence of increases in marriage-related migration following drought. The main tables are not presented with multinomial estimation, since the inclusion of fixed effects for district of origin or survey cause problems of convergence in the estimation and were omitted in the multinomial model. The results in Table 1.31 in Appendix 1K should thus be interpreted with caution.

1.6.5 The role of ethnicity

In the context of coping mechanisms for drought, one may wonder whether the patterns of marriage-related migration found here simply reflect a tendency for inter-ethnic marriage, which has been shown to function as an alternative insurance mechanism against shocks (Hotte and Marazyan, 2019). Ethnicity may be an important reason for marriage in sub-Saharan Africa, but compared to other countries in sub-Saharan Africa (Crespin-Boucaud, 2020), we can rule out ethnicity as a major reason for marriage. Table 1.14 shows that the rates of inter-ethnic marriages in Malawi are relatively low.²⁸

Table 1.14: DHS Malawi

DHS-survey	Inter-ethnic marriage		Age 1st cohabitation
	no	yes	
2000	66.3	33.7	17.4
2004-2005	69.6	30.4	17.3
2010	69.5	30.5	17.4
2015-2016	66.9	33.1	17.8

1.7 Descriptive evidence on outcomes following migration

Migration for marriage, in particular at an early age, makes the individual - the girl in this case - unable to finish her schooling, less likely to participate in the labour market and potentially, more likely to get unwanted pregnancies when she is younger. This and the following subsections will discuss some of the potential outcomes for the migrants, particularly for those migrating for marriage, based on descriptive statistics.

1.7.1 School attendance after migration

One of the most important consequences of migration for marriage or work is potential drop-out of school. Table 1.15 shows the drop-out rates in the destination household, by category of migrant status. We see a clear correlation in the data between migration for marriage and drop-out rates from school.²⁹ We also note that almost half of the migrant sample dropped out of school, 13

²⁸We thank Juliette Crespin-Boucaud who kindly calculated this table from the DHS in Malawi.

²⁹A formal comparison of the percentage of drop-out for marriage-pregnancy between non-migrants and within-migrants yields a p-value of 0.00, whereas the t-test for equality of the percentage of drop-out for marriage-pregnancy

532 (Table 1.16) compared to the entire sample of migrants in Table 1.21, Appendix 1E. Also, the largest share of those dropping-out of school for marriage or pregnancy are migrants who moved for marriage (Table 1.16).³⁰

Table 1.17 shows that a large proportion of such drop-out of school occurs for those migrating recently and up to 5 years ago. Although these are descriptive statistics only, they could be interpreted as an indication of some of the consequences of marriage-related migration.

Table 1.15: Drop-out school in destination- By category

	Drop-out school - other reasons		Drop-out school for marriage-pregnancy		Total	
	Observ.	Percentage	Observ.	Percentage	Observ.	Percentage
NoMIG.	27115	89.7	3129	10.3	30244	100.0
Between Dist-MIG.	5538	89.0	685	11.0	6223	100.0
Within Dist-MIG.	6305	86.2	1010	13.8	7315	100.0
International MIG.	970	90.5	102	9.5	1072	100.0
Total	39928	89.0	4926	11.0	44854	100.0

Note: The table includes only those for whom the information was reported. The question asked is: why did you not continue your education? There are 18 options, among which the option "married-became pregnant".

Table 1.16: Drop-out school in destination - By reason to migrate

Reason to migrate	Drop-out school - other reasons		Drop-out school for marriage-pregnancy		Total	
	Observ.	Percentage	Observ.	Percentage	Observ.	Percentage
WORK RELATED/LOOK FOR WORK	3704	96.0	155	4.0	3859	100.0
SCHOOL/STUDIES	57	85.1	10	14.9	67	100.0
MARRIAGE	4863	80.0	1217	20.0	6080	100.0
FAMILY REASONS/FOLLOW JOIN FAMILY	2285	91.1	223	8.9	2508	100.0
LAND/PLOT/OTHER LAND RELATED	459	93.9	30	6.1	489	100.0
OTHERS	469	88.7	60	11.3	529	100.0
Total	11837	87.5	1695	12.5	13532	100.0

Note: The table includes only those for whom the information was reported. The question asked is: why did you not continue your education? There are 18 options, among which the option "married-became pregnant".

between non-migrants and between-migrants yields a p-value of 0.06.

³⁰A formal test of equality between the proportions of migrants moving for marriage-related reasons compared to migrants moving for all other reasons yields a p-value of 0.00.

Table 1.17: Drop-out school in destination- Years ago to migrate

Years ago to migrate	Drop-out for other reasons	Drop-out for marriage-pregnancy	Total
	Observ.	Observ.	Observ.
0	677	83	760
1	2111	308	2419
2	1667	259	1926
3	1425	213	1638
4	1237	180	1417
5	1116	164	1280
6	517	68	585
7	448	68	516
8	409	58	467
9	473	57	530
10	527	66	593
11	314	49	363
12	172	21	193
13	126	20	146
14	167	23	190
15	163	20	183
16	188	21	209
17	106	16	122
Total	11843	1694	13537

Note: The table includes only those for whom the information was reported. The question asked is: why did you not continue your education? There are 18 options, among which married-became pregnant.

1.7.2 Do female migrants have higher probability of being pregnant after migrating?

Tables 1.18 to 1.20 show the likelihood of pregnancy among migrant categories. In the main analysis we find a higher probability of marriage-related migration within the district of origin. Table 1.18 shows that within-district migrants, in particular, followed by between-district migrants have higher likelihood to have been pregnant in the last 24 months compared to non-migrants (both p-values=0.00). Once they moved, 49.5% of those migrating for marriage were pregnant in the past 24 months (2094 migrants for marriage compared to the 4234 marriage migrants in Table 1.19). This proportion is significantly different from that of the migrants for other motives (p-value=0.00).

Table 1.20 shows that most migrants with a pregnancy within the last 24 months are recent marriage-related migrants, having migrated between 1 to 5 years ago. We conclude that the descriptive statistics on some migrant outcomes in the destination are indicative that the self-reported migration for marriage reasons do reflect actual marriage-related moves. The statistics also indicate some of the adverse consequences of marriage in terms of dropping out of school because of

pregnancy. All these consequences could usefully be investigated in future work that would need to combine migration data with DHS data on health outcomes, for example.

Table 1.18: Have a pregnancy in the last 24 months in destination - By category

	Pregnancy		No pregnancy		Total	
	Observ.	Percentage	Observ.	Percentage	Observ.	Percentage
NoMIG.	4029	23.6	13068	76.4	17097	100.0
Between Dist-MIG.	1056	27.9	2731	72.1	3787	100.0
Within Dist-MIG.	1673	38.2	2711	61.8	4384	100.0
International MIG.	125	27.0	338	73.0	463	100.0
Total	6883	26.7	18848	73.3	25731	100.0

Note: The table includes only those for whom the information was reported. The question asked is: In the past 24 months, did you give birth to a child, even if born dead?.

Table 1.19: Have a pregnancy in the last 24 months in destination - By reason to migrate

Reason to migrate	Pregnancy		No Pregnancy		Total	
	Observ.	Percentage	Observ.	Percentage	Observ.	Percentage
WORK RELATED/LOOK FOR WORK	189	24.3	588	75.7	777	100.0
SCHOOL/STUDIES	3	3.6	80	96.4	83	100.0
MARRIAGE	2094	49.5	2140	50.5	4234	100.0
FAMILY REASONS/FOLLOW JOIN FAMILY	276	10.8	2274	89.2	2550	100.0
LAND/PLOT/OTHER LAND RELATED	52	31.9	111	68.1	163	100.0
OTHERS	111	31.5	241	68.5	352	100.0
Total	2725	33.4	5434	66.6	8159	100.0

Note: The table includes only those for whom the information was reported. The question asked is: In the past 24 months, did you give birth to a child, even if born dead?.

Table 1.20: Have a pregnancy in the last 24 months in destination- Years ago to migrate

Years ago to migrate	Pregnancy	No Pregnancy	Total
	Observ.	Observ.	Observ.
0	83	209	292
1	604	1218	1822
2	491	835	1326
3	407	704	1111
4	344	531	875
5	277	524	801
6	97	275	372
7	83	215	298
8	73	195	268
9	70	187	257
10	105	274	379
11	49	125	174
12	20	61	81
13	15	51	66
14	11	38	49
Total	2729	5442	8171

Note: The table includes only those for whom the information was reported. The question asked is: In the past 24 months, did you give birth to a child, even if born dead?.

1.8 Conclusion

In this article, we aimed at linking the literature on migration as an adaptation mechanism following drought with the literature on marriage institutions, and provide more evidence on gender-related differences in the response of migration to drought. In a country with bride prices, such as Malawi, one consequence of drought may be an increase in the likelihood of child marriages. Combining data from the LSMS-ISA with satellite gridded weather data over the period 2000-2016, we estimate the probability of migration for reasons of work and marriage for men and women and find significant differences in the response to drought.

The results on marriage-related migration show that women migrate mainly within the district of origin and that the general effect of drought is to constrain migration. But the net migration probability increases after drought in some age intervals, in particular in the age intervals 18-27 and 28-37. We find little evidence of an increase in marriage-related migration in the lowest age category, 10-17. As concerns migration related to work, we see an increase in men's between-district migration, and a particularly large increase of almost 9 percentage points in the age interval 10-17 following severe drought. The between-district migration for work of girls in the same age interval

also increases following severe drought, by 5 percentage points. On the one hand, the article thus shows evidence of an optimistic result indicating no large changes in the probability of child marriages following drought in Malawi. On the other hand, we also conclude on a pessimistic note in providing evidence of increases in children's migration for work following drought, mainly for boys, and to a lesser extent for girls.

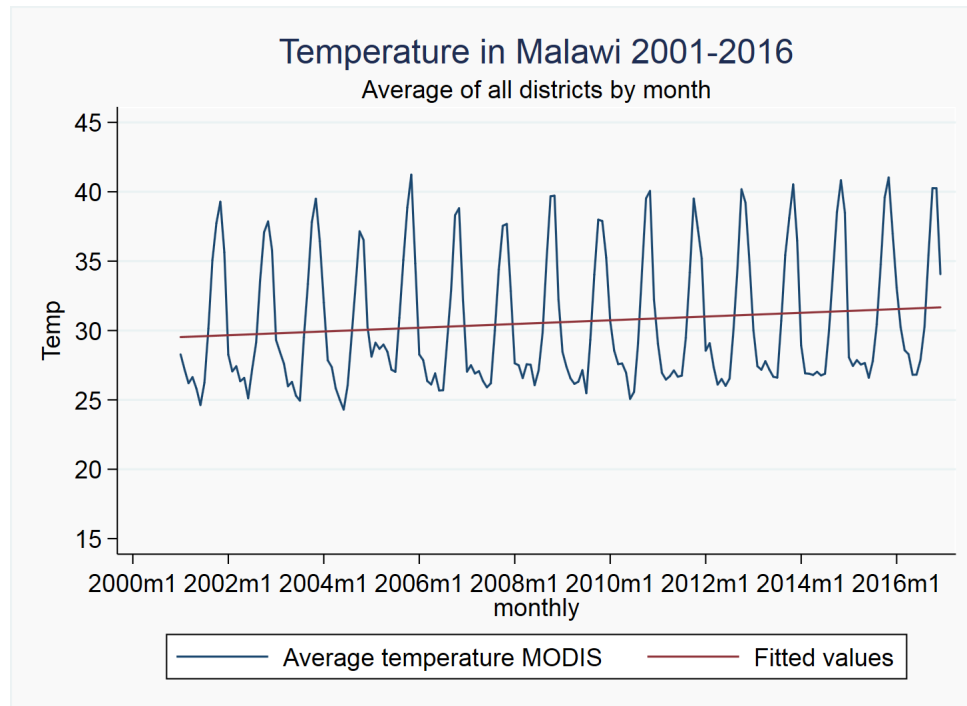
Though the rich data on the district of origin and destination as well as the time of migration is an important advantage of the article, unfortunately, the lack of household information at origin before moving limits the study. For instance, it would be relevant to investigate the effect of wealth and the level of education of the household head in marriage-related migration decisions. Observing the majority of the information at the destination does not allow us to check whether or not household composition, such as the number of siblings, or the presence of sisters in particular, might reduce such migration.

Future research directions include investigating in more detail potential interrelations in different household members' migration, and exploring further the migrants moving for family reasons or to join family. It is possible that such moves of older individuals constitute another coping strategy of rural households. Also, whereas we cannot reject the null hypothesis that drought does not increase the probability of marriages in the lowest age category – which is a positive result – the indication of potential increases in work-related migration in the lowest age groups should be investigated further, since it would go in line with existing evidence on the use of child labour following income shocks (Edmonds and Shrestha, 2013). Finally, the consequences of marriage-related migration need to be investigated further in order to find causal effects on the outcomes. The choice of Malawi for the case study was motivated by the large droughts that the country has suffered during the time period studied. In consequence, the evidence here only stems from one country and could be extended by further work on social outcomes of drought-induced migration for marriage in other countries.

Appendices

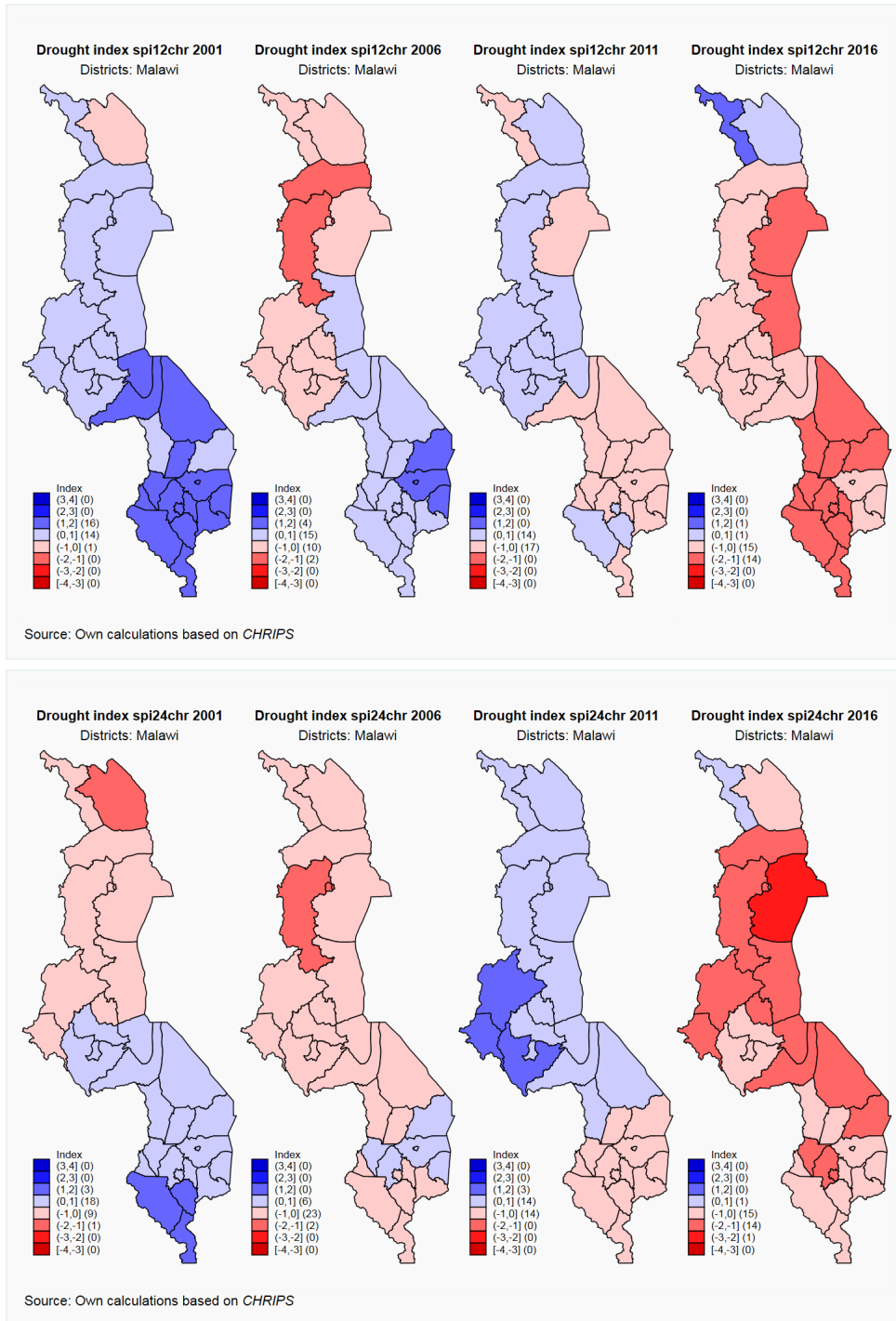
1A Temperature in Malawi and SPI maps

Figure 1.1: Average Temperature in Malawi



Note: created based on MODIS Land Surface Temperature Using images of the satellite MOD11A1.006 at 1km of resolution.

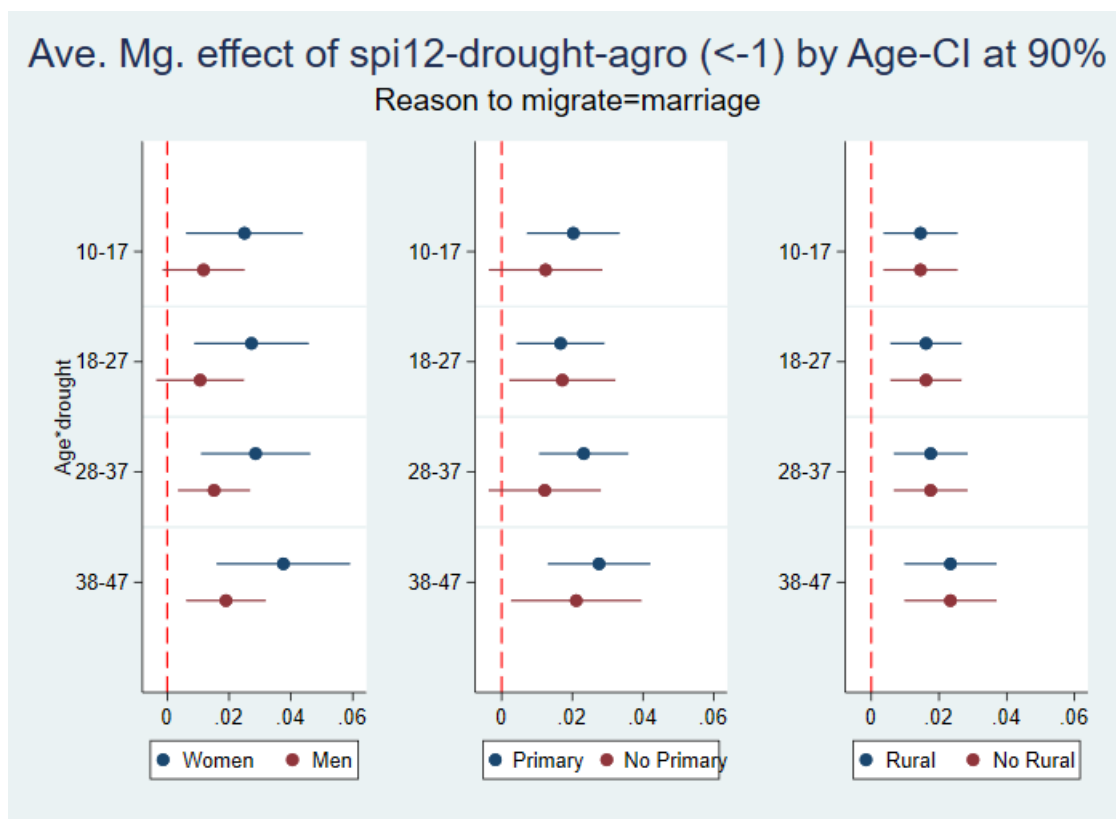
Figure 1.2: SPI-CHIRPS by district (scale 12-24)



Note: created based on the SPI-CHIRPS. The monthly scale gives the number of months over which water deficits accumulate.

1B Marginal Effects by socioeconomic characteristics

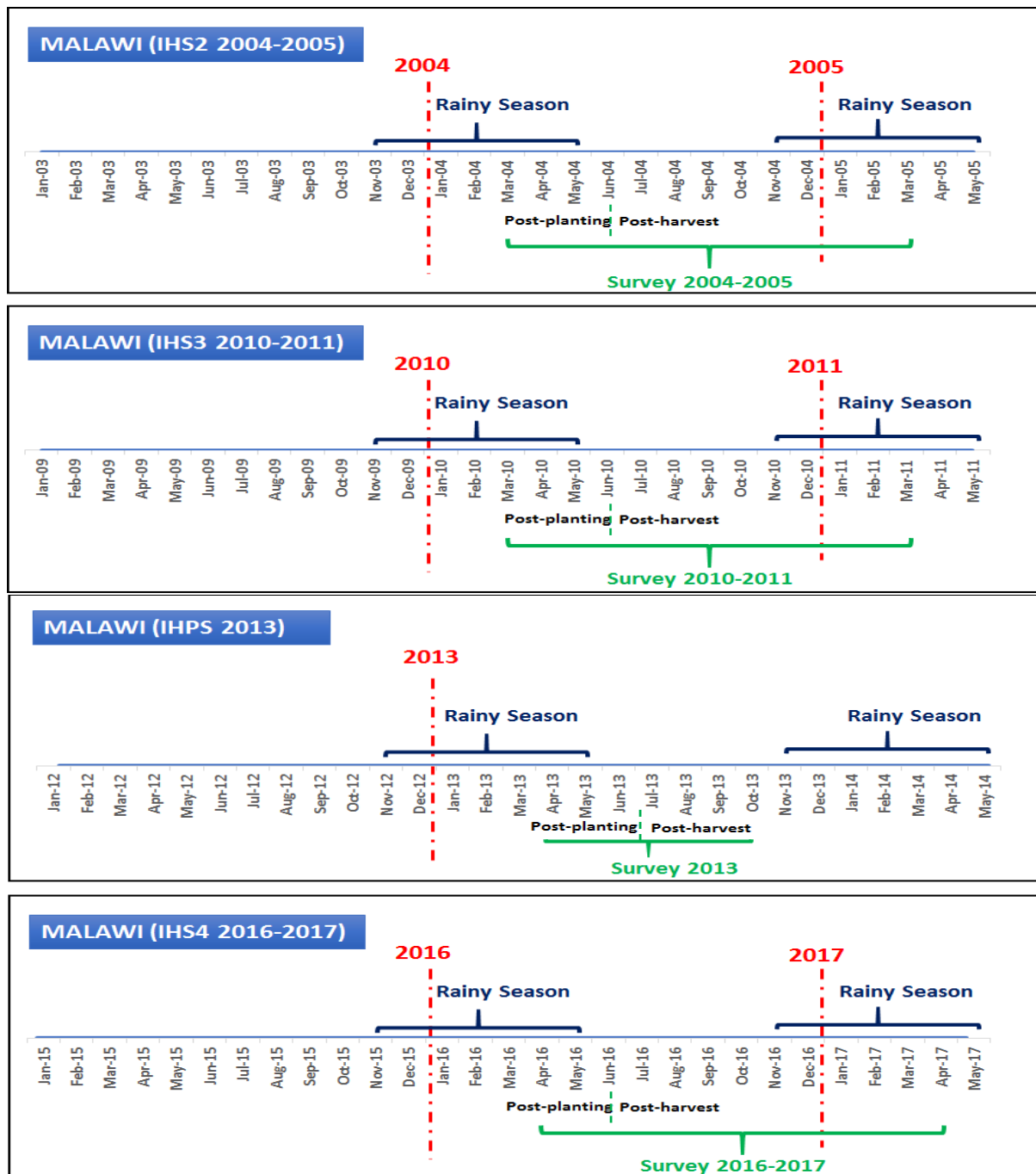
Figure 1.3: Average marginal effects by different categories
Migrating for Marriage



Note: created using the *spi12-drought-agro* (< -1), which is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -1 during at least one month of the agricultural growing season. Baseline category of age: 48-57. Age <10 and Age >57 are excluded. Results for a Probit model for between and within migrants and migrating for marriage. The first figure corresponds to the punctual estimations restricting for women (blue) and men (brown). The same holds for Primary-No Primary and Rural-No Rural.

1C Timeline surveys and rainy seasons in Malawi

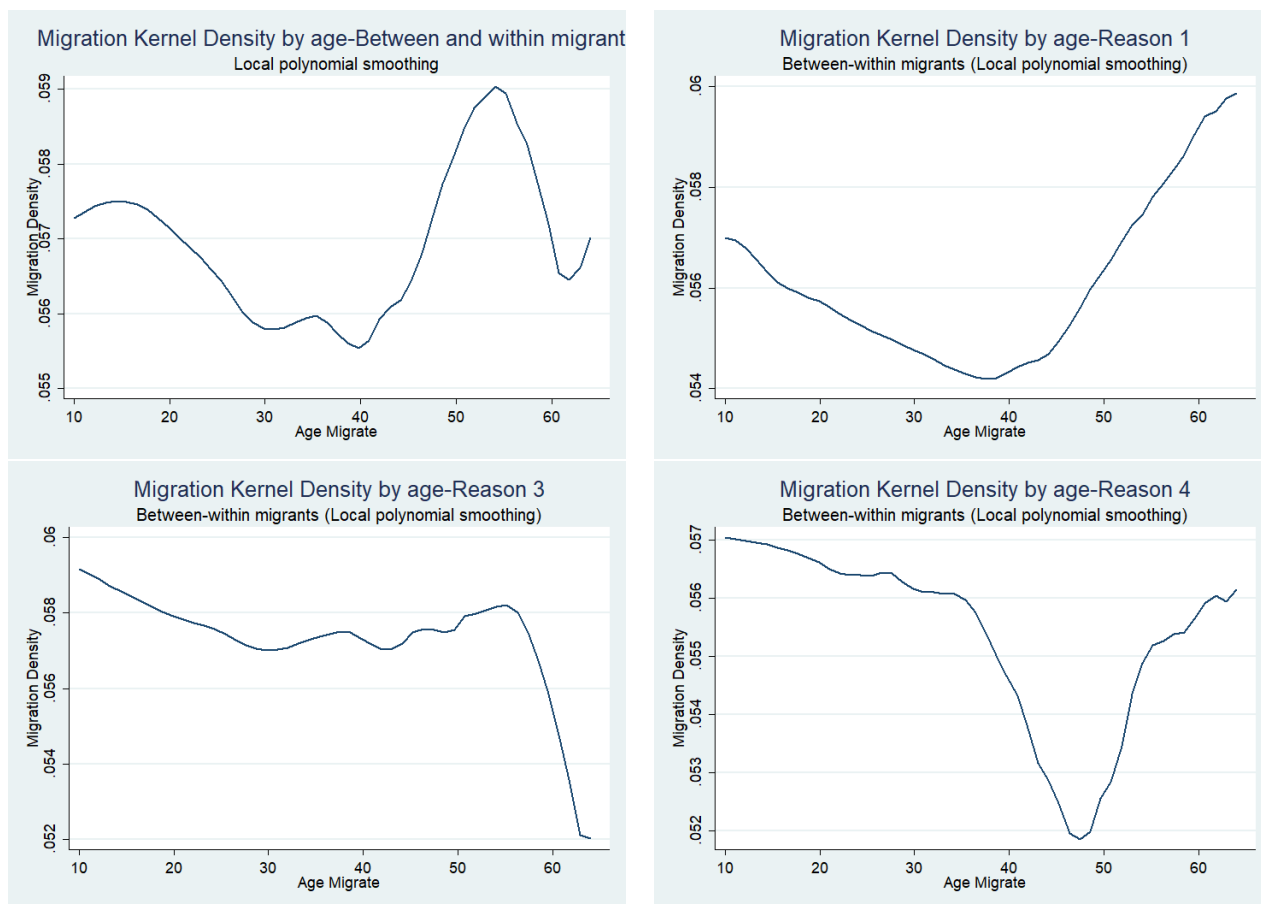
Figure 1.4: Time-line of Surveys and Rainy seasons. LSMS-ISA for MALAWI



Note: Own construction based on IHS and IHPS wave reports MALAWI LSMS-ISA 2016-2017, MALAWI LSMS-ISA 2013, MALAWI LSMS-ISA 2010-2011 and MALAWI LSMS-ISA 2004-2005.

1D Kernel density of migration by age

Figure 1.5: Additional graphs-Kernel density of migration by age
(Main reasons to migrate: 1 for work, 3 for marriage, 4 for family reasons)



Note: Own construction based on IHS and IHPS wave reports MALAWI LSMS-ISA 2016-2017, MALAWI LSMS-ISA 2013, MALAWI LSMS-ISA 2010-2011 and MALAWI LSMS-ISA 2004-2005.

1E District Classification

Table 1.21: District Classification for the LSMS-ISA

District	By district of origin			By district of destination		
	Within Dist-MIG Obs.	Between Dist-MIG Obs.	Total Obs.	Within Dist-MIG Obs.	Between Dist-MIG Obs.	Total Obs.
1101 (CHITIPA)	290	713	1003	136	713	849
1102 (KARONGA)	367	754	1121	256	754	1010
1103 (NKHATABAY)	268	618	886	472	618	1090
1104 (RUMPHI)	378	501	879	562	501	1063
1105 (MZIMBA)	663	976	1639	473	976	1449
1107 (MZUZU CITY)	303	831	1134	1428	831	2259
1201 (KASUNGU)	520	767	1287	690	767	1457
1202 (NKHOTA KOTA)	254	513	767	481	513	994
1203 (NTCHISI)	151	517	668	341	517	858
1204 (DOWA)	324	605	929	464	605	1069
1205 (SALIMA)	258	243	501	305	243	548
1206 (LILONGWE)	617	982	1599	432	982	1414
1207 (MCHINJI)	220	409	629	273	409	682
1208 (DEDZA)	396	359	755	260	359	619
1209 (NTCHEU)	403	192	595	414	192	606
1210 (LILONGWE CITY)	587	1312	1899	1692	1312	3004
1301 (MANGOCHI)	383	420	803	394	420	814
1302 (MACHINGA)	302	386	688	251	386	637
1303 (ZOMBA)	348	450	798	314	450	764
1304 (CHIRADZULU)	262	285	547	258	285	543
1305 (BLANTYRE)	251	337	588	272	337	609
1306 (MWANZA)	86	133	219	256	133	389
1307 (THYOLO)	554	348	902	302	348	650
1308 (MULANJE)	485	471	956	382	471	853
1309 (PHALOMBE)	126	397	523	236	397	633
1310 (CHIKWAWA)	259	409	668	237	409	646
1311 (NSANJE)	179	409	588	178	409	587
1312 (BALAKA)	265	170	435	327	170	497
1313 (NENO)	36	58	94	236	58	294
1314 (ZOMBA CITY)	230	566	796	776	566	1342
1315 (BLANTYRE CITY)	973	566	1539	1172	566	1738
Total	10738	15697	26435	14270	15697	29967

Note: created based on the LSMS-ISA information. Likoma Island is removed from the sample. District of origin is missing for LSMS-ISA Survey 2004-2005

1G Characteristics of migrants and Balance Tables

Table 1.23: Percentage of transfers in the household by person in destination
(by type of migrant)

	Obs.	Mean	SD	Min	Max
NoMIG.	67524	0.05	0.10	0	1
Between Dist-MIG.	7974	0.04	0.10	0	1
Within Dist-MIG.	8388	0.05	0.12	0	1
International MIG.	1484	0.08	0.15	0	1
Total	85370	0.05	0.10	0	1

Note: The table includes only those for whom the information was reported. Only for survey 2004-2005 and 2010-2011.

Table 1.24: Percentage of no agricultural income (no agricultural wage and self-employment) in
the household by person in destination
(by type of migrant)

	Obs.	Mean	SD	Min	Max
NoMIG.	66329	0.17	0.30	0	1
Between Dist-MIG.	7531	0.46	0.43	0	1
Within Dist-MIG.	8055	0.30	0.38	0	1
International MIG.	1457	0.19	0.32	0	1
Total	83372	0.21	0.33	0	1

Note: The table includes only those for whom the information was reported. Only for survey 2004-2005 and 2010-2011.

Table 1.25: Balance Tables for droughts at origin-district using SPI measure

Variable	Sample: Between-district Migrants				Sample: Within-district Migrants			
	(1) No Drought	(2) Drought	(3) Diff	(4) N	(1) No Drought	(2) Drought	(3) Diff	(4) N
Sex of the individual	0.493 (0.500)	0.497 (0.500)	0.000 (0.000)	127,044	0.559 (0.496)	0.583 (0.493)	0.000 (0.000)	182,466
Number of years of education	9.130 (3.881)	9.114 (3.822)	-0.000 (0.000)	127,044	7.718 (3.581)	7.628 (3.464)	-0.000 (0.000)	182,466
Dummy for at least primary education	0.660 (0.474)	0.663 (0.473)	0.000 (0.000)	127,044	0.513 (0.500)	0.509 (0.500)	0.000 (0.000)	182,466
Vincenty distance dist-origin to dist-destination	162.698 (135.481)	162.234 (134.325)	0.000 (0.000)	127,044	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	182,466
Percentage of kids less than 5 years old in the Household	0.169 (0.162)	0.172 (0.162)	-0.000 (0.000)	127,044	0.203 (0.170)	0.206 (0.171)	-0.000 (0.000)	182,466
Dummy for migrants coming from rural areas	0.753 (0.431)	0.777 (0.416)	-0.000 (0.000)	127,044	0.862 (0.345)	0.893 (0.309)	-0.000 (0.000)	182,466
Dummy for migrants coming from urban areas	0.246 (0.431)	0.223 (0.416)	-0.000 (0.000)	127,044	0.138 (0.345)	0.107 (0.309)	0.000 (0.000)	182,466
Size of the Household	5.024 (2.240)	5.065 (2.250)	-0.000 (0.000)	127,044	5.015 (2.247)	5.038 (2.276)	-0.000 (0.000)	182,466
Mean of monthly maximum rainfall during year(mm)-CHIRPS	16.966 (4.213)	15.702 (3.837)	-0.874 (0.227)***	127,044	16.745 (4.196)	15.873 (3.997)	-0.710 (0.239)***	182,466
Mean of monthly average rainfall during year(mm)-CHIRPS	2.915 (0.519)	2.627 (0.578)	-0.299 (0.037)***	127,044	2.936 (0.527)	2.693 (0.587)	-0.284 (0.035)***	182,466
Mean of monthly minimum rainfall during year(mm)-CHIRPS	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	127,044	0.000 (0.000)	0.000 (0.000)	-0.000 (0.000)	182,466
Mean of monthly cumulative rainfall during year(mm)-CHIRPS	88.744 (15.643)	80.051 (17.572)	-9.160 (1.168)***	127,044	89.379 (15.923)	82.029 (17.847)	-8.678 (1.085)***	182,466
Anomalies-Mean of monthly average rainfall during year(mm)-CHIRPS	-0.029 (0.188)	-0.133 (0.210)	-0.109 (0.014)***	127,044	-0.021 (0.191)	-0.109 (0.213)	-0.103 (0.013)***	182,466
SD of monthly average rainfall during year(mm)-CHIRPS	3.619 (0.607)	3.275 (0.767)	-0.322 (0.058)***	127,044	3.629 (0.634)	3.364 (0.798)	-0.269 (0.049)***	182,466
SD of monthly cumulative rainfall during year(mm)-CHIRPS	110.193 (18.104)	100.116 (23.415)	-9.886 (1.850)***	127,044	110.539 (19.022)	102.829 (24.348)	-8.114 (1.576)***	182,466
Mean of monthly maximum temperature during year(degrees)-MODIS	34.021 (2.969)	34.925 (3.005)	0.166 (0.191)	112,928	33.718 (3.196)	34.406 (3.248)	0.120 (0.223)	162,192
Mean of monthly average temperature during year(degrees)-MODIS	30.121 (2.277)	30.799 (2.396)	0.093 (0.140)	112,928	29.902 (2.455)	30.394 (2.543)	0.049 (0.161)	162,192
Mean of monthly minimum temperature during year(degrees)-MODIS	25.588 (1.641)	25.977 (1.809)	-0.013 (0.102)	112,928	25.475 (1.755)	25.705 (1.855)	-0.067 (0.112)	162,192
Anomalies-Mean of monthly average temperature during year(degrees)-MODIS	-0.082 (0.507)	0.069 (0.533)	0.021 (0.031)	112,928	-0.131 (0.547)	-0.021 (0.566)	0.011 (0.036)	162,192
SD of monthly average temperature during year(degrees)-MODIS	4.667 (1.464)	5.108 (1.174)	-0.035 (0.090)	112,928	4.575 (1.515)	4.978 (1.332)	-0.035 (0.104)	162,192
Observations	93,680	33,364	127,044		132,997	49,469	182,466	

Note: Includes dummies for year, district of origin, district of destination and for survey.

Drought dummy calculated using the SPI for 12 months scale; 1 for at least one month of drought ($SPI < -1$) during the growing season, 0 otherwise. Sample restricted to age 10-57 years old.

Standard errors clustered at origin-district (column 3). Standard deviations in parenthesis (column 1 and 2)

1H Additional Robustness checks

Table 1.26: Within District Migration - Women and Men migrating for work (reason 1)

	SPI-Cumulative drought for 12 months (Women)				SPI-Cumulative drought for 12 months (Men)			
	OLS Interactive FE		PROBIT-MG EFFECTS		OLS Interactive FE		PROBIT-MG EFFECTS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
spi12-drought-agro (< -1)	-0.00481 (0.0214)		0.000565 (0.0227)		-0.00478 (0.0188)		-0.00105 (0.0193)	
spi12-drought-agro (< -1) x age (10-17)	0.00850 (0.0342)		0.00698 (0.0334)		-0.00406 (0.0310)		-0.0232 (0.0334)	
spi12-drought-agro (< -1) x age (18-27)	0.0164 (0.0273)		0.00817 (0.0281)		0.000892 (0.0182)		-0.00204 (0.0186)	
spi12-drought-agro (< -1) x age (28-37)	0.00134 (0.0248)		-0.00320 (0.0252)		0.0125 (0.0142)		0.00897 (0.0141)	
spi12-drought-agro (< -1) x age (38-47)	0.0442 (0.0340)		0.0396 (0.0307)		0.0122 (0.0147)		0.00872 (0.0157)	
spi12-drought-agro (< -2)		0.00499 (0.0590)		0.00818 (0.0470)		0.0242 (0.0229)		0.0253 (0.0191)
spi12-drought-agro (< -2) x age (10-17)		-0.0408 (0.0784)		-0.0413 (0.0723)		-0.0422 (0.0421)		-0.0639 (0.0530)
spi12-drought-agro (< -2) x age (18-27)		-0.0107 (0.0571)		-0.0115 (0.0454)		-0.00114 (0.0278)		-0.000977 (0.0224)
spi12-drought-agro (< -2) x age (28-37)		-0.0215 (0.0609)		-0.0270 (0.0500)		0.00338 (0.0187)		0.00115 (0.0147)
spi12-drought-agro (< -2) x age (38-47)		-0.0227 (0.0641)		-0.0197 (0.0537)		0.0331 (0.0275)		0.0203 (0.0217)
Observations	6948	6948	6948	6948	22284	22284	22284	22284
R2	0.182	0.182			0.167	0.167		
R2 _a	0.171	0.171			0.163	0.164		
R2 _p			0.0244	0.0227			0.0211	0.0225
Ncluster	31	31	31	31	31	31	31	31
Pcorr	.	.	94.44	94.44	.	.	94.44	94.44

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Includes dummies for year, district of origin, district of destination and for survey, as well as controls for education and sex (in the case it is relevant).

spi12-drought-agro (< -1) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -1 during at least one month of the agricultural growing season.

spi12-drought-agro (< -2) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -2 during at least one month of the agricultural growing season.

Agricultural growing season from November to March. Baseline category of age: 48-57. Age <10 and Age >57 are excluded. Distance or Vicenty-distance calculated using geodesic distances between a pair of points on the surface of the Earth. Standard errors clustered at origin-district. Standard deviations in parenthesis.

Table 1.27: Within District Migration - Women and men migrating for marriage (reason 3)

	SPI-Cumulative drought for 12 months (Women)				SPI-Cumulative drought for 12 months (Men)			
	OLS Interactive FE		PROBIT-MG EFFECTS		OLS Interactive FE		PROBIT-MG EFFECTS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
spi12-drought-agro (< -1)	-0.0358*** (0.00977)		-0.0461*** (0.0162)		-0.0271* (0.0145)		-0.0316* (0.0187)	
spi12-drought-agro (< -1) x age (10-17)	0.0305*** (0.00831)		0.0416*** (0.0143)		0.0390 (0.0247)		0.0412 (0.0264)	
spi12-drought-agro (< -1) x age (18-27)	0.0365*** (0.00958)		0.0471*** (0.0162)		0.0182 (0.0158)		0.0208 (0.0198)	
spi12-drought-agro (< -1) x age (28-37)	0.0300** (0.0115)		0.0399** (0.0173)		0.00386 (0.0145)		0.00390 (0.0184)	
spi12-drought-agro (< -1) x age (38-47)	0.0495** (0.0186)		0.0596*** (0.0227)		0.00961 (0.0152)		0.0118 (0.0187)	
spi12-drought-agro (< -2)		-0.0343** (0.0126)		-0.0526 (0.0354)		-0.0229 (0.0372)		-0.0280 (0.0536)
spi12-drought-agro (< -2) x age (10-17)		0.0359* (0.0197)		0.0552 (0.0424)		0.0807 (0.0686)		0.0719 (0.0630)
spi12-drought-agro (< -2) x age (18-27)		0.0432** (0.0169)		0.0608 (0.0398)		0.0428 (0.0352)		0.0456 (0.0513)
spi12-drought-agro (< -2) x age (28-37)		0.0498*** (0.0174)		0.0673* (0.0360)		-0.0105 (0.0348)		-0.0178 (0.0532)
spi12-drought-agro (< -2) x age (38-47)		0.0585** (0.0215)		0.0753* (0.0392)		0.0398 (0.0449)		0.0435 (0.0565)
Observations	65376	65376	65376	65376	20430	20430	20430	20430
R2	0.166	0.166			0.167	0.168		
R2 _a	0.165	0.165			0.164	0.164		
R2 _p			0.0174	0.0173			0.0248	0.0248
Ncluster	31	31	31	31	31	31	31	31
Pcorr	.	.	94.44	94.44	.	.	94.44	94.44

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Includes dummies for year, district of origin, district of destination and for survey, as well as controls for education and sex (in the case it is relevant).

spi12-drought-agro (< -1) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -1 during at least one month of the agricultural growing season.

spi12-drought-agro (< -2) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -2 during at least one month of the agricultural growing season.

Agricultural growing season from November to March. Baseline category of age: 48-57. Age <10 and Age >57 are excluded. Distance or Vicenty-distance calculated using geodesic distances between a pair of points on the surface of the Earth. Standard errors clustered at origin-district. Standard deviations in parenthesis.

11 Droughts in the destination district

Table 1.28: $Drought_o$ and $Drought_d$ together.

Women migrating for marriage (reason 3) Between District

	SPI-Cumulative drought for 12 months (WOMEN)			
	OLS Interactive FE		PROBIT-MG EFFECTS	
	(1)	(2)	(3)	(4)
spi12-drought-agro (< -1) dist-o	-0.0253*** (0.00688)		-0.0139 (0.0152)	
spi12-drought-agro (< -1) x age (10-17) dist-o	0.0324** (0.0128)		0.0212 (0.0180)	
spi12-drought-agro (< -1) x age (18-27) dist-o	0.0271*** (0.00699)		0.0183 (0.0133)	
spi12-drought-agro (< -1) x age (28-37) dist-o	0.0466*** (0.0128)		0.0386** (0.0173)	
spi12-drought-agro (< -1) x age (38-47) dist-o	0.00839 (0.0207)		0.000899 (0.0260)	
spi12-drought-agro (< -1) dist-d	-0.0300 (0.0191)		-0.0445** (0.0218)	
spi12-drought-agro (< -1) x age (10-17) dist-d	0.0193 (0.0220)		0.0311 (0.0261)	
spi12-drought-agro (< -1) x age (18-27) dist-d	0.0273 (0.0216)		0.0427* (0.0245)	
spi12-drought-agro (< -1) x age (28-37) dist-d	0.0261 (0.0213)		0.0399* (0.0242)	
spi12-drought-agro (< -1) x age (38-47) dist-d	0.0457* (0.0238)		0.0604** (0.0277)	
spi12-drought-agro (< -2) dist-o		0.0682 (0.0829)		0.0134 (0.0507)
spi12-drought-agro (< -2) x age (10-17) dist-o		-0.115 (0.0773)		-0.0268 (0.0487)
spi12-drought-agro (< -2) x age (18-27) dist-o		-0.0651 (0.0776)		-0.0117 (0.0470)
spi12-drought-agro (< -2) x age (28-37) dist-o		-0.0672 (0.0748)		-0.0112 (0.0450)
spi12-drought-agro (< -2) x age (38-47) dist-o		-0.0780 (0.0656)		-0.00737 (0.0474)
spi12-drought-agro (< -2) dist-d		-0.0201 (0.0442)		-0.0484 (0.0309)
spi12-drought-agro (< -2) x age (10-17) dist-d		0.0313 (0.0468)		0.0494 (0.0334)
spi12-drought-agro (< -2) x age (18-27) dist-d		0.0162 (0.0404)		0.0469* (0.0277)
spi12-drought-agro (< -2) x age (28-37) dist-d		0.0315 (0.0483)		0.0603* (0.0323)
spi12-drought-agro (< -2) x age (38-47) dist-d		0.0214 (0.0440)		0.0378 (0.0339)
Observations	28728	28728	28728	28728
R2	0.175	0.175		
$R2_a$	0.172	0.172		
$R2_p$			0.0203	0.0194
Ncluster	31	31	31	31
Pcorr	.	.	94.44	94.44

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Includes dummies for year, district of origin, district of destination and for survey, as well as controls for education and sex (in the case it is relevant).

$spi12$ -drought-agro (< -1) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -1 during at least one month of the agricultural growing season.

$spi12$ -drought-agro (< -2) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -2 during at least one month of the agricultural growing season.

Agricultural growing season from November to March. Baseline category of age: 48-57. Age <10 and Age >57 are excluded. Distance or Vicenty-distance calculated using geodesic distances between a pair of points on the surface of the Earth. Standard errors clustered at origin-district. Standard deviations in parenthesis.

Table 1.29: Adding $dist_d \times year$ fixed effects. Between and Within District Migration
Women migrating for marriage (reason 3)

	SPI-Cumulative drought for 12 months			
	OLS Interactive FE		PROBIT-MG EFFECTS	
	(1)	(2)	(3)	(4)
spi12-drought-agro(< -1)	-0.0297** (0.0112)		-0.0430** (0.0168)	
spi12-drought-agro(< -1) x age(10-17)	0.0294*** (0.00937)		0.0434*** (0.0145)	
spi12-drought-agro(< -1) x age(18-27)	0.0353*** (0.0106)		0.0486*** (0.0164)	
spi12-drought-agro(< -1) x age(28-37)	0.0386*** (0.0113)		0.0527*** (0.0166)	
spi12-drought-agro(< -1) x age(38-47)	0.0384** (0.0165)		0.0525*** (0.0202)	
spi12-drought-agro(< -2)		-0.0288 (0.0222)		-0.0428 (0.0310)
spi12-drought-agro(< -2) x age(10-17)		0.0294 (0.0208)		0.0452 (0.0302)
spi12-drought-agro(< -2) x age(18-27)		0.0370 (0.0223)		0.0501 (0.0317)
spi12-drought-agro(< -2) x age(28-37)		0.0423* (0.0211)		0.0551* (0.0286)
spi12-drought-agro(< -2) x age(38-47)		0.0423 (0.0258)		0.0556* (0.0320)
Observations	94104	94104	93374	93374
R2	0.155	0.155		
$R2_a$	0.154	0.154		
$R2_p$			0.0299	0.0298
Ncluster	31	31	31	31
Pcorr	.	.	94.40	94.40

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Includes dummies for year, district of origin, district of destination, for survey and district of destination \times year, as well as controls for education and sex (in the case it is relevant).

spi12-drought-agro (< -1) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -1 during at least one month of the agricultural growing season.

spi12-drought-agro (< -2) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -2 during at least one month of the agricultural growing season.

Baseline category of age: 48-57. Age<10 and Age>57 are excluded. Distance or Vicenty-distance calculated using geodesic distances between a pair of points on the surface of the Earth. Standard errors clustered at origin-district. Standard deviations in parenthesis.

1J Household interaction effects

Table 1.30: Including control for whether household head and partner migrated together Between and Within District Migration - Women migrating for marriage (reason 3)

	SPI-Cumulative drought for 12 months			
	OLS Interactive FE		PROBIT-MG EFFECTS	
	(1)	(2)	(3)	(4)
spi12-drought-agro (< -1)	-0.0330*** (0.0102)		-0.0484*** (0.0163)	
spi12-drought-agro (< -1) x age (10-17)	0.0282*** (0.00922)		0.0436*** (0.0144)	
spi12-drought-agro (< -1) x age (18-27)	0.0362*** (0.0103)		0.0508*** (0.0164)	
spi12-drought-agro (< -1) x age (28-37)	0.0372*** (0.0113)		0.0529*** (0.0168)	
spi12-drought-agro (< -1) x age (38-47)	0.0410** (0.0176)		0.0573*** (0.0217)	
spi12-drought-agro (< -1)-couple HH mig	0.00908 (0.0403)		0.0306 (0.0444)	
spi12-drought-agro (< -1) x age (10-17)-couple HH mig	0.0406 (0.0458)		0.0239 (0.0461)	
spi12-drought-agro (< -1) x age (18-27)-couple HH mig	0.0000994 (0.0466)		-0.0154 (0.0498)	
spi12-drought-agro (< -1) x age (28-37)-couple HH mig	0.0152 (0.0445)		-0.00465 (0.0482)	
spi12-drought-agro (< -1) x age (38-47)-couple HH mig	-0.00448 (0.0473)		-0.0240 (0.0483)	
spi12-drought-agro (< -2)		-0.0386*** (0.0138)		-0.0711** (0.0343)
spi12-drought-agro (< -2) x age (10-17)		0.0333 (0.0201)		0.0668 (0.0414)
spi12-drought-agro (< -2) x age (18-27)		0.0432** (0.0168)		0.0744* (0.0382)
spi12-drought-agro (< -2) x age (28-37)		0.0523*** (0.0179)		0.0835** (0.0353)
spi12-drought-agro (< -2) x age (38-47)		0.0536** (0.0224)		0.0867** (0.0387)
spi12-drought-agro (< -2)-couple HH mig		0.170 (0.173)		0.155** (0.0785)
spi12-drought-agro (< -2) x age (10-17)-couple HH mig		-0.153 (0.193)		-0.127 (0.102)
spi12-drought-agro (< -2) x age (18-27)-couple HH mig		-0.120 (0.180)		-0.115 (0.0802)
spi12-drought-agro (< -2) x age (28-37)-couple HH mig		-0.155 (0.168)		-0.139* (0.0770)
spi12-drought-agro (< -2) x age (38-47)-couple HH mig		-0.153 (0.157)		-0.145** (0.0590)
Observations	94104	94104	94104	94104
R2	0.148	0.148		
R2 _a	0.147	0.147		
R2 _p			0.0125	0.0123
Ncluster	31	31	31	31
Pcorr	.	.	94.44	94.44

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Includes dummies for year, district of origin, district of destination and for survey.

spi12-drought-agro (< -1) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -1 during at least one month of the agricultural growing season.

spi12-drought-agro (< -2) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -2 during at least one month of the agricultural growing season.

Agricultural growing season from November to March. Baseline category of age: 48-57. Age<10 and Age>57 are excluded. Distance or Vicenty-distance calculated using geodesic distances between a pair of points on the surface of the Earth. Standard errors clustered at origin-district. Standard deviations in parenthesis. The "couple HH mig" variable is a dummy when the couple (household head and partner) of the household migrated together to the same district in the same year.

1K A multinomial estimation

Table 1.31: Multinomial Logit: No migrants and Between-Within Migrants - Women migrating for marriage (reason 3)

a) Multinomial Logit - Relative Risk Ratio to baseline (No migrate)				
	1		2	
	MLOGIT SPI		MLOGIT SPI Ext	
	Between	Within	Between	Within
	(1)	(2)	(3)	(4)
spi12-drought-agro (< -1)	0.416 (0.278)	0.406** (0.154)		
spi12-drought-agro (< -1) x age (10-17)	2.514 (1.685)	2.422** (0.913)		
spi12-drought-agro (< -1) x age (18-27)	2.503 (1.532)	2.504** (0.972)		
spi12-drought-agro (< -1) x age (28-37)	3.806** (2.255)	2.068* (0.800)		
spi12-drought-agro (< -1) x age (38-47)	2.193 (1.503)	2.809** (1.331)		
spi12-drought-agro (< -2)			0.887 (0.855)	0.466 (0.394)
spi12-drought-agro (< -2) x age (10-17)			1.127 (1.122)	2.215 (1.973)
spi12-drought-agro (< -2) x age (18-27)			1.205 (1.138)	2.310 (2.241)
spi12-drought-agro (< -2) x age (28-37)			1.397 (1.289)	2.827 (2.720)
spi12-drought-agro (< -2) x age (38-47)			1.677 (1.505)	3.959 (3.875)
Observations	495094		495094	
r^2_p	0.144		0.144	
Ncluster	31		31	

Exponentiated coefficients; Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Includes dummies for year, district of destination, as well as controls for education and sex (in the case it is relevant). Dummies for district of origin and survey were excluded as they generated a problem of convergence in the likelihood of the Multinomial Logit. Although the results could be compared with Table 1.8, caution has to be taken as the Multinomial Logit does not include the fixed effects mentioned before. Also, the table here compares the within and between migrants (excluding the migrants for other reasons different than reason 3) with all the no migrants.

spi12-drought-agro (< -1) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -1 during at least one month of the agricultural growing season.

spi12-drought-agro (< -2) is a dummy equal to one if the SPI index (accumulated over 12 months) takes values below -2 during at least one month of the agricultural growing season.

Agricultural growing season from November to March. Baseline category of age: 48-57. Age <10 and Age >57 are excluded. Distance or Vicenty-distance calculated using geodesic distances between a pair of points on the surface of the Earth. Standard errors clustered at origin-district. Standard deviations in parenthesis.

b) Marginal Effects-Multinomial Logit (from previous table)

	MLOGIT-MG EFFECTS	
	(1)	(2)
0-No Mig	0.00906** (0.00360)	0.00581 (0.00615)
1-Between	-0.00262 (0.00213)	-0.000340 (0.00291)
2-Within	-0.00644** (0.00272)	-0.00547 (0.00608)
spi12-drought-agro (< -1) x age (10-17)		
0-No Mig	-0.00907*** (0.00342)	
1-Between	0.00275 (0.00216)	
2-Within	0.00632** (0.00271)	
spi12-drought-agro (< -1) x age (18-27)		
0-No Mig	-0.00930*** (0.00353)	
1-Between	0.00274 (0.00197)	
2-Within	0.00656** (0.00278)	
spi12-drought-agro (< -1) x age (28-37)		
0-No Mig	-0.00918*** (0.00344)	
1-Between	0.00401** (0.00197)	
2-Within	0.00517* (0.00277)	
spi12-drought-agro (< -1) x age (38-47)		
0-No Mig	-0.00972** (0.00411)	
1-Between	0.00234 (0.00215)	
2-Within	0.00738** (0.00340)	
spi12-drought-agro (< -2) x age (10-17)		
0-No Mig		-0.00604 (0.00589)
1-Between		0.000339 (0.00302)
2-Within		0.00570 (0.00640)
spi12-drought-agro (< -2) x age (18-27)		
0-No Mig		-0.00654 (0.00692)
1-Between		0.000541 (0.00286)
2-Within		0.00600 (0.00696)
spi12-drought-agro (< -2) x age (28-37)		
0-No Mig		-0.00842 (0.00667)
1-Between		0.000980 (0.00280)
2-Within		0.00744 (0.00691)
spi12-drought-agro (< -2) x age (38-47)		
0-No Mig		-0.0114 (0.00713)
1-Between		0.00152 (0.00272)
2-Within		0.00985 (0.00702)
Observations	495094	495094
Ncluster	31	31

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1L Measurement error in the variable of droughts

As mentioned in the section of data³¹, the drought measures constructed with the community questionnaires of the LSMS-ISA could be subject to measurement error and recall bias. Those questionnaires are answered by the chief of the community who might have thought that the drought events were harder than they were, if the particular drought hit harder her community, compared to other years or compared to droughts in other communities. This reason justified the use of drought measures such as the Standard Precipitation Index (SPI) which is an exogenous and more precise measure, less prone to these errors. The section of main results shows that drought measures using the LSMS-ISA tend to over-estimate the effect on migration, compared with the more appropriate measures of droughts (SPI). In particular, we find higher coefficients using the LSMS-ISA drought measures than when using the SPI.

According to the *classical measurement error* in the explanatory variables, the attenuation bias should be between zero and one and the coefficient $\hat{\beta}$ will be biased towards zero (smaller coefficient). The results found here go in the opposite direction of the *classical measurement error*, and this appendix will try to explain the reason behind this.

We will start with the simplest regression model with one independent variable x and y as the dependent variable. Both will have zero mean for the easiness of the exposition, so the relationship is:

$$y = \beta x + \epsilon \tag{L.1}$$

We have data observed with additive measurement error:

$$\tilde{x} = x + \mu \tag{L.2}$$

$$\tilde{y} = y + \nu \tag{L.3}$$

³¹This section is constructed using the lecture notes of Professor Jörn-Steffen Pischke, for the course "Ec524: empirical methods in applied economics" at the London School of Economics.

Let us make these simplifying assumptions:

$$E(\mu) = 0 \tag{L.4}$$

$$plim \frac{1}{n} (y' \mu) = 0 \tag{L.5}$$

$$plim \frac{1}{n} (x' \mu) = 0 \tag{L.6}$$

$$plim \frac{1}{n} (\epsilon' \mu) = 0 \tag{L.7}$$

These assumptions plus the assumption $\sigma_v^2 = 0$ (only measurement error in the explanatory variable x) define the *classical measurement error*. By substituting (L.2) into (L.1):

$$y = \beta(\tilde{x} - \mu) + \epsilon = \beta\tilde{x} + (\epsilon - \beta\mu) \tag{L.8}$$

The measurement error in x is the source of endogeneity. As \tilde{x} and μ are positively correlated the OLS estimation will lead to negative bias in $\hat{\beta}$ if the true β is positive (or positive bias if β is negative). In this classical case, using the OLS estimator we have:

$$\hat{\beta} = \frac{cov(\tilde{x}, y)}{var(\tilde{x})} = \frac{cov(x + \mu, \beta x + \epsilon)}{var(x + \mu)}$$

Taking the limit in probability,

$$plim \hat{\beta} = \beta \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\mu^2} = \beta\lambda \tag{L.9}$$

As $0 < \lambda < 1$, $\hat{\beta}$ estimated will be biased towards zero (the attenuation bias) and λ is known as the attenuation factor.

In the particular case that concerns us, the assumption (L.6) does not hold. As mentioned in the first paragraph of this appendix, it might be that as a drought occurs and as the impact is more severe, the chief of the community perceives that as a more extreme drought (in x), compared with other droughts occurred in the past or in other communities, making the chief to declare the event (and making μ bigger) or think that the event lasted for more years than what it actually did.

Hence, $cov(x, \mu) \neq 0$ (positive in this case). By relaxing this assumption on the OLS estimator as before, and taking the limit in probability, we get:

$$plim\hat{\beta} = \frac{\beta\sigma_x^2 + \beta\sigma_{x\mu}}{\sigma_x^2 + \sigma_\mu^2 + 2\sigma_{x\mu}} \quad (\text{L.10})$$

Notice that (L.10) becomes (L.9) when $\sigma_{x\mu} = 0$. (L.10) can be re-expressed as:

$$plim\hat{\beta} = \beta \left(1 - \frac{\sigma_\mu^2 + \sigma_{x\mu}}{\sigma_x^2 + \sigma_\mu^2 + 2\sigma_{x\mu}} \right) = \beta(1 - b_{\mu x}) \quad (\text{L.11})$$

with $(1 - b_{\mu x})$ as the attenuation factor. In order to know the bias of $\tilde{\beta}$, let us take the derivative of (L.11) with respect to $\sigma_{x\mu}$:

$$\frac{\partial\beta(1 - b_{\mu x})}{\partial\sigma_{x\mu}} = \beta \left(\frac{\sigma_\mu^2 - \sigma_x^2}{(\sigma_\mu^2 + \sigma_x^2 + 2\sigma_{x\mu})^2} \right) \quad (\text{L.12})$$

Here, we can have 4 cases:

1. If $\sigma_\mu^2 > \sigma_x^2$, the $\uparrow \sigma_{x\mu}$ (when $cov(x, \mu) > 0$) will \uparrow the attenuation factor.
2. If $\sigma_\mu^2 > \sigma_x^2$, the $\downarrow \sigma_{x\mu}$ (when $cov(x, \mu) < 0$) will \downarrow the attenuation factor.
3. If $\sigma_\mu^2 < \sigma_x^2$, the $\uparrow \sigma_{x\mu}$ (when $cov(x, \mu) > 0$) will \downarrow the attenuation factor.
4. If $\sigma_\mu^2 < \sigma_x^2$, the $\downarrow \sigma_{x\mu}$ (when $cov(x, \mu) < 0$) will \uparrow the attenuation factor.

Our situation corresponds to case 1, so the chiefs of the community might over-perceive the droughts, much harder than they were, so $\sigma_\mu^2 > \sigma_x^2$. Also, as we argue before, $cov(x, \mu) > 0$. Hence, the attenuation factor will increase, making the $\tilde{\beta}$ much bigger as we found in the results.

The case 2 is quoted by Bound and Krueger (1991) as the *mean reverting measurement error*. According to them, in earnings data, if \tilde{x} has mainly measurement error ($\sigma_\mu^2 > \sigma_x^2$), a more negative $\sigma_{x\mu}$ (the covariance between x and μ is negative) implies a lower attenuation factor and could even reverse the sign of β . Notice that the reversal of the sign can happen also in case 3.

Chapter 2

Climatic shocks, air quality, and health at birth in Bogotá¹

¹This chapter is the product of joint work with Jorge Bonilla.

Abstract

We contribute to the literature on air pollution and health by assessing an additional channel, the effect of El Niño Southern Oscillation (ENSO) on health. Currently, there is a vast literature on the effects of urban pollution on health. Our research, unlike other studies, jointly investigates the effects of pollution, ENSO and local weather on health. On the one hand, ENSO manifests itself as an extreme climatic shock that follows certain seasonality and influences weather. It may also have an impact on floods, droughts and agriculture inducing changes in food markets or a loss of household income, which also affect health. On the other hand, health outcomes are affected by other factors which follow separate mechanisms to the previous ones. Therefore, pollutant impacts on health may be interpreted as separate effects from other shocks mediated through ENSO. Using a database from 1998 to 2015 on air quality and vital statistics for Bogotá, and ENSO information, we find that across several specifications, ENSO affects birth weight and the probability of low birth weight after separating pollution and classical local weather impacts. Interestingly, the effect on birth weight of ENSO are several times larger than the impacts of pollution. Being exposed to ENSO may decrease birth weight up to 1.3%, while an increase of 1 *ppb* of SO_2 or 1 $\mu g/m^3$ of PM_{25} might reduce birth weight up to 0.3% or 0.14%, respectively. From a policy point of view, these results are relevant because regardless of the measure of pollution that we employ, the amount of the impacts exhibited by climatic shocks via ENSO events dominate.

Keywords: climate change, health, ENSO Index, El Niño, La Niña, weather, pollution, Bogotá.
JEL: Q54, Q53, J13, I15, I18.

2.1 Introduction

Two of the major climatic phenomena on Earth are El Niño Southern Oscillation (ENSO²) and the North Atlantic Oscillation (NAO). They create important fluctuations in weather in the surroundings of the Pacific Ocean and the North Atlantic, respectively. With regard to ENSO, the occurrence and impacts of its events have been extensively studied from the meteorological and geographical side (see for instance [Stenseth et al. \(2003\)](#)). Both the intensity and the frequency of El Niño have been forecasted to increase in the coming years (see chapter 3 of the report "Global Warming of 1.5 Celsius degrees" of the Intergovernmental Panel of Climate Change and [Wang et al. \(2017\)](#)). Despite this, studies that explore the impact of ENSO on socio-economic outcomes are rare. One of the first studies to do so was [Hsiang et al. \(2011\)](#), who find that the ENSO events affects one-fifth of all civil conflicts since 1950 in the tropics. They also show that El Niño may double the probability of having new civil conflicts with respect to La Niña events. More recently, [Dingel et al. \(2020\)](#) use ENSO as a natural experiment to show that when cereal productivity is more spatially correlated, more productive countries experience larger gains from trade than the less productive ones.

In addition to conflicts, ENSO also affects health. [Caminade et al. \(2016\)](#) assess the impact of El Niño on the Zika outbreak in Latin America. Also, [Kovats et al. \(2003\)](#) find a relationship between ENSO events and the spread of malaria in South America and South Asia and cholera in Bangladesh. Their findings suggest that high temperatures during El Niño in 2015-2016 fostered the disease transmission by the *Aedes Aegypti* mosquito. In Colombia, [Brando and Santos \(2015\)](#) explore the impact of the strong La Niña event in 2010-2011 on birth weight, finding that the climatic shock reduces birth weight. They also argue that households react to the shock by decreasing investment in education for children. When it comes to economic variables, in another study for Colombia, [Abril-Salcedo et al. \(2020\)](#) find that El Niño had a significant impact on consumer food prices during the strong event in 2015, by using a smooth transition nonlinear model. In that study, weather shocks are transitory and asymmetric, affecting inflation growth from five to nine months after the shock. This could be an example of the potential channel through which El Niño affects

²The episodes of cold and warm weather are measured by the Oceanic Niño Index (ONI) that captures the El Niño Southern Oscillation (ENSO). For the case of Colombia, El Niño is associated with warm episodes and La Niña with cold episodes, but the same does not hold in other countries.

food prices and hence, consumption of households.

Another strand of the literature studies the impacts of climate change measured as temperature variability in general, but not ENSO in particular. [Deschênes et al. \(2009\)](#) find that extreme temperatures increase low birth weight probability and reduce birth weight. Similarly, [Barreca and Schaller \(2020\)](#) find that hot temperatures reduce gestational weeks. In the case of mortality, [Barreca et al. \(2015\)](#) show that high temperatures over the period 1900-2004 had an effect on deaths, but the effect was lower in places that often face high temperatures. Our article contributes to two strands of the literature: first, about the effects of climate variability on health by assessing the influence of ENSO on health variables; and second, regarding the effect of air pollution on health adding the impact of climatic shocks.³

The ENSO phenomena may influence health outcomes through several channels: (1) ENSO manifests as a climatic shock affecting weather, hence, extreme weather may have an impact on health (see [Barreca et al. \(2015\)](#)). (2) ENSO affects air pollution, which may influence health. There is a vast literature on the effect of pollutants on health ([Arceo et al., 2015](#); [Knittel et al., 2016](#); [Schlenker and Walker, 2016](#); [Deschênes et al., 2017](#)). It becomes particularly relevant in areas where pollution levels are high, as it is the case of urban environments. (3) There might exist a direct impact of ENSO on health. Although one may hypothesize about the presence of this channel, it is difficult to find practical examples to cite in this regard, particularly considering that effects are mainly transmitted via changes in weather. (4) ENSO also affects the economy. For example, floods and droughts induce changes in agriculture and food prices (see [Abril-Salcedo et al. \(2020\)](#) for the effects of ENSO on prices in Colombia) that lead to a loss of household income and consumption or effects on other socioeconomic variables. As [Hsiang et al. \(2011\)](#) argue, ENSO may be seen as a mechanism that assembles economic shocks. Such economic impacts may be considered as a set of general equilibrium effects that may also influence health. In this article, we identify channels (1) and (2). We also believe that our approach partly captures channels (3) and (4) through the overall estimate of the effect of ENSO on health.

³Only recently, [Elorreaga et al. \(2020\)](#) found that the ENSO phenomenon could have increased childhood stunting on the coast of Piura in Peru.

Our article investigates different mechanisms: the effect of local weather on health, the impact of ENSO on health, and the influence of pollution on health. Unlike previous studies, we jointly explore the effects of pollution, ENSO and local weather on birth weight, probability of low birth weight, and gestational length. In our approach, pollutant impacts on health may be interpreted as separate effects from other shocks mediated through ENSO. We conduct our analysis using data over the period 1998-2015 from Bogotá, one of the largest urban centers in Latin America. Pollution levels in the city reach annual means of $26 \mu\text{g}/\text{m}^3$ for PM_{10} in areas of low exposition and $87 \mu\text{g}/\text{m}^3$ in the most polluted places, which exceed the maximum value of $20 \mu\text{g}/\text{m}^3$ recommended by the World Health Organization (WHO).⁴ In the case of daily concentrations, pollution may increase to $160 \mu\text{g}/\text{m}^3$, also surpassing the WHO's air quality standard of $50 \mu\text{g}/\text{m}^3$ (see [Secretary-Environment \(2015\)](#)). Estimating the effect on health is complex because, in addition to the effects of local weather and ENSO, these factors may alter air pollutant concentrations. Apart from annual weather fluctuations (see [Deschênes and Greenstone \(2011\)](#)), no one seems to have investigated further the link between pollution and health, considering the interaction it has with climatic shocks as ENSO events. El Niño or La Niña might exacerbate pollution. The intuition behind this is that a higher intensity or steeper trend of ENSO events could induce temperature and rainfall changes that favor higher concentrations.⁵ On our data, we find that ENSO events may affect some pollutants. Similar results are shown by [Grundström et al. \(2011\)](#) in the North-West of Europe indicating that climate shifts of the NAO affect air pollution. Our analysis for ENSO indicates that intraday variation matters, which motivates us to suggest maximum and peak hour concentrations as pollutant exposure measures on health besides the classical daily average.

Another challenge in assessing the effect of air pollution on health is that pollution is endogenous due to omitted variables and unobserved confounders. Therefore, our empirical strategy is based on a two stage approach. In a first stage, we estimate pollution as a function of local weather and ENSO shocks exploiting exogenous variation of wind direction as an instrument. Wind direction is included as a set of indicator variables and interactions with wind speed. As an extra set of

⁴Mobile sources account for more than 50% of air pollution in the city.

⁵See [Watson et al. \(1988\)](#) for the difference in meaning between concentration, exposure and dose.

instruments, we use changes in driving restrictions over time. In the second stage, we regress health outcomes on instrumented pollution, ENSO shocks and local weather. We aim to capture size and length of pollutant exposure and strong climatic variability exposure *in utero* by computing measures for each of the three gestational quarters. Note that ENSO shocks represent effects of strong climatic variability (or nonlinear transmission effects) through channels different from pollution exposure. Thus, the estimates of the extreme climate variability effects in our specifications provide a measure of the aggregated impacts from channels (1) and (2), and partially from channels (3) and (4). The lack of information about food prices and household consumption during the entire period of our sample prevent us to separate economic shocks from other channels.

In the review of the state of the art for pollution and individual well-being by [Graff Zivin and Neidell \(2013\)](#), they acknowledge the use of quasi-experimental techniques to study causal estimates of pollution and the role of avoidance behaviour to reduce the effect of pollutants on health. However, they consider that such studies should expand to focus toward the effects of pollution on human capital. Lately, the literature is heading from studying labor supply outcomes (see [Graff Zivin and Neidell \(2013\)](#)) to cognitive formation and performance (see [Hanna and Oliva \(2015\)](#)). Recently, [de la Mata and Gaviria \(2019\)](#) studied the effects of air pollution on health for the case of Bogotá, using a smaller sample of children born in the city and attending kinder-gardens between 2010 to 2014. These studies highlight the importance of human capital for economic growth, in the sense that negative shocks from pollution on health should affect labor productivity (see [Graff Zivin and Neidell \(2012\)](#)).

Our study thus sheds light on the magnitude of extreme climate variability and pollution on birth outcomes, as an indicator of human capital. It is relevant since better health at birth has been linked with lower health-care costs after birth and later in life (see [Almond et al. \(2005\)](#)). As has been pointed out by [Almond and Currie \(2011a\)](#), different early life conditions, particularly before the age of five, can have persistent and profound impacts on later life, affecting educational outcomes and future earnings. In fact, health at birth is a crucial component of human capital development and the evidence discussed by [Almond and Currie \(2011a\)](#) suggest that the period *in utero* is one of the most important stages for children's later development. With respect to

educational attainment, [Conley et al. \(2001\)](#) find that a low birth weight child who spent 6 years at the poverty line is less likely to graduate from high school compared to a normal birth weight child who spent 6 years at the poverty line. In addition, [Conley et al. \(2006\)](#) find that increases in birth weight might lead to decreases in infant and neonatal mortality. In a wider perspective, [Almond and Currie \(2011b\)](#) review the literature of the "fetal origins" hypothesis which establishes that the effects of fetal conditions are persistent and that the intrauterine environment can lead to future disease (diabetes, overweight, cardiovascular diseases), and affect future earnings. In more economical terms, their later-life impacts can extend to "bread and butter" economic outcomes, educational attainment and wages. Taken together, this could imply that in addition to interventions implemented at young age, children might also need to be targeted during pregnancy, as early life conditions can have long-term impacts in the accumulation of human capital.

The article is organized as follows: Section [2.2](#) discusses the importance of analyzing health outcomes and air pollution in Bogotá, Section [2.3](#) describes the data-sets, some stylized facts that motivate the study of the relationship between ENSO, local weather and pollution in Bogotá, and the descriptive statistics. Section [2.4](#) discusses the empirical strategy while section [2.5](#) presents the main results of the effects of pollution and ENSO events on health. Finally, Section [2.6](#) concludes.

2.2 The importance of health, air pollution and ENSO in Bogotá

According to a report in 2017 of the National Department of Planning [DNP \(2017\)](#), the health costs of the environmental degradation in Colombia are COP \$20.7 billions, equivalent to 2.6% of the GDP in 2015. Of that, the costs associated to urban air pollution correspond to 75%. For the particular case of Bogotá, pollution is one of the most important environmental problems, getting more attention in the media. [DNP \(2017\)](#) indicates that 10.5% (3.219) of the total deaths in the city are attributed to air pollution, with an estimated cost of COP \$4.2 billions (2.5% of the city's GDP). Pollution in Bogotá reach under certain conditions levels of cities like Delhi. According to the Real-time Air Quality Index⁶ from the World Air Quality Index Project, Bogotá showed a value of 141 in the index on March 8th of 2019 (very harmful for the individuals). For the same day, the

⁶The index takes into account different pollutants emitted locally.

index had a value of 151 (unhealthy) in Delhi that contrasted with a value of 23 (good) for Paris; this gives a picture of the dimension of the problem in the city.

In 2011, the Decennial Pollution Abatement Plan in Bogotá conducted by [Behrentz et al. \(2010\)](#) (University of Los Andes and University of La Salle) for the Secretary of Environment of the city found that implementing the plan (by renewing the type of buses, introducing diesel particle filters, among other measures) could avoid 27,500 hospitalizations for respiratory diseases among the population of children, 75,000 case entries to emergency rooms and around 7,500 in intensive care. In the base scenario without measures, infant deaths could reach 3,700 for the period 2010 to 2020, but with the implementation of the plan this number could have declined in 1,500 potential deaths during the same years. For adults, the plan could have avoided 14,000 deaths, 40,000 cases in emergency rooms and 11,000 cases of hospitalizations for respiratory diseases. Although the cost of the plan could have reached COP \$1.7 billions, the benefits would have totaled COP \$16 billions. Despite the benefits of the plan would exceed its costs, the plan only implemented few of its reduction measures. It motivated to re-build the plan in the first quarter of 2021.

These reports make evident the large costs that air pollution can cause on health, for both children and adults. However, they do not consider the effects of ENSO on health as our article is one of the first addressing it. As we show in the results section, ENSO effects on birth weight are several times larger than the impacts coming from pollution. As we argue, ENSO might have an impact on health but can also affect health through the effects it has on pollution, through the effect on weather and then on health, or through the distortions it generates in the economy which can affect health. If the environmental degradation from air pollution already has a huge cost for the city, considering the ENSO impact on health could add an extra cost that has not been taken into account, which is important from a policy perspective.

In this article we will focus only on children during gestation because first, it allows us to capture more instantaneous effects during the nine months of gestation, and second, we do not have to control for unobserved factors more common in adults, such as whether or not the individual smokes, the habits of eating healthy food, exercising and having a healthy life in general. By choosing children

in gestation time, we should be less prone to this type of bias in our results. Additionally, using children during a short period of time should also permit to capture exposure to ENSO events (El Niño-La Niña) more immediately.

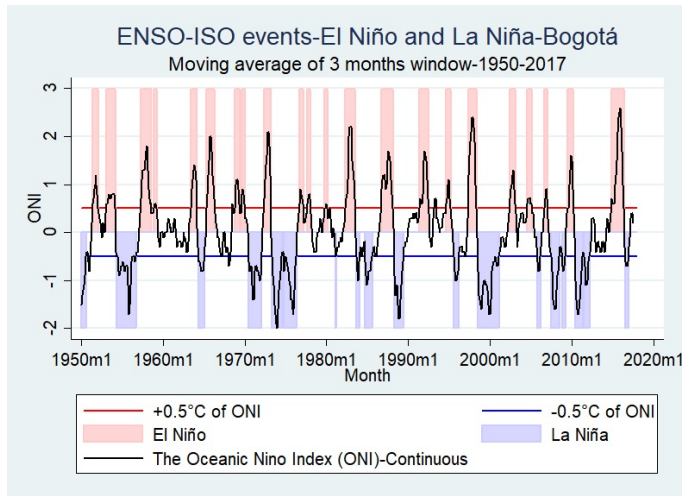
2.3 Data and stylized facts

We employ a rich database collected from several sources. Information on ENSO is taken from The National Oceanic and Atmospheric Administration (NOAA). El Niño and La Niña data are available from 1950 to 2017. Episodes are reported on a monthly basis in two ways. The first is a continuous variable named The Oceanic El Niño Index (ONI). It is calculated as a three-month moving average of the Sea Surface Temperature (SST) in the tropical Pacific. The second is a discrete variable that indicates if the ONI lies within specific bands that describe El Niño and La Niña events or normal status. Periods above 0.5 or below -0.5 of ONI for a minimum of 5 consecutive overlapping months (colored in red and blue) define El Niño or La Niña events, respectively. The past 30 years (i.e. 1986-2015) are used to compute the departure from the SST average.

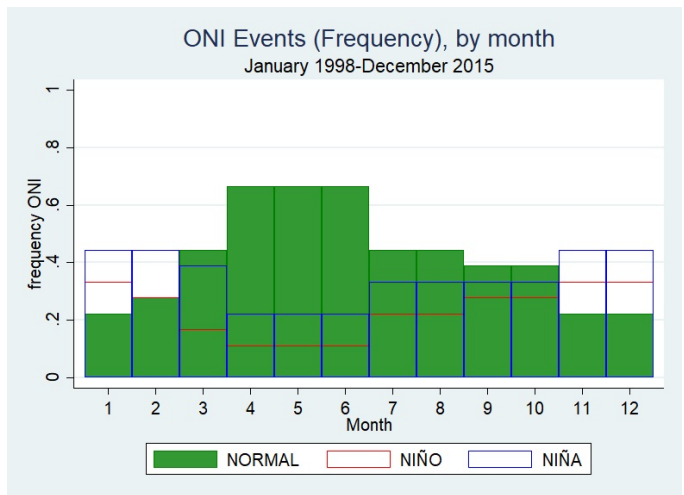
Figure 2.1 (a) depicts the ONI index since 1950. The graph shows two important characteristics. In the last 20 years, La Niña (in blue) is more frequent and El Niño (in red) reaches much higher values over time. The latter fact coincides with the climate model forecasts indicating that the extreme El Niño frequency will increase linearly with the Global Mean Temperature (GMT) towards a doubling of 1.5 Celsius degrees on warming (Wang et al. (2017)). Moreover, Figure 2.1 (b) displays the percentage of months exhibiting Normal, El Niño, and La Niña episodes between 1998 and 2015 (percentage with respect to last 18 years). For instance, 44% of the months of January between 1998 and 2015 faced La Niña, 33% El Niño, while 22% were normal months. It suggests that ENSO events usually occur between November and February, while normal months are frequent between April and June. This kind of seasonality is considered in our analysis when we assign exposure to children *in utero*.

Weather variables come from the Air Quality Monitoring Network of Bogotá (RMCAB). It covers from twelve to twenty one monitoring stations. Figure 2.2 shows the spatial distribution of

Figure 2.1: ONI Index-ENSO events and Frequency



(a)



(b)

Note: The National Oceanic and Atmospheric Administration (NOAA)

RMCAB in the city. Hourly data exist from August 1st of 1997 to December 31st of 2015.⁷ We exclude the stations of Usme, Vitelma and Bolivia because of lack of observations. Available variables are wind speed (*WS*), wind direction (*WD*), solar radiation (*SR*), temperature (*TMP*), barometric pressure (*BPR*), rainfall (*RAIN*), and relative humidity (*RH*). For our analysis, we use rainfall, temperature and wind variables to describe local weather conditions. In the case of wind direction, we convert degree units to a set of indicator variables that correspond to the standard eight-azimuth

⁷Although information may exhibit missing values due to blackouts or failures of the monitoring platform, the missing patterns seem to be random and reports for most stations have a large valid data representation.

bearings.⁸ Figure 2.3 displays the average intraday variation of local weather conditions such as rain and temperature under ENSO conditions.

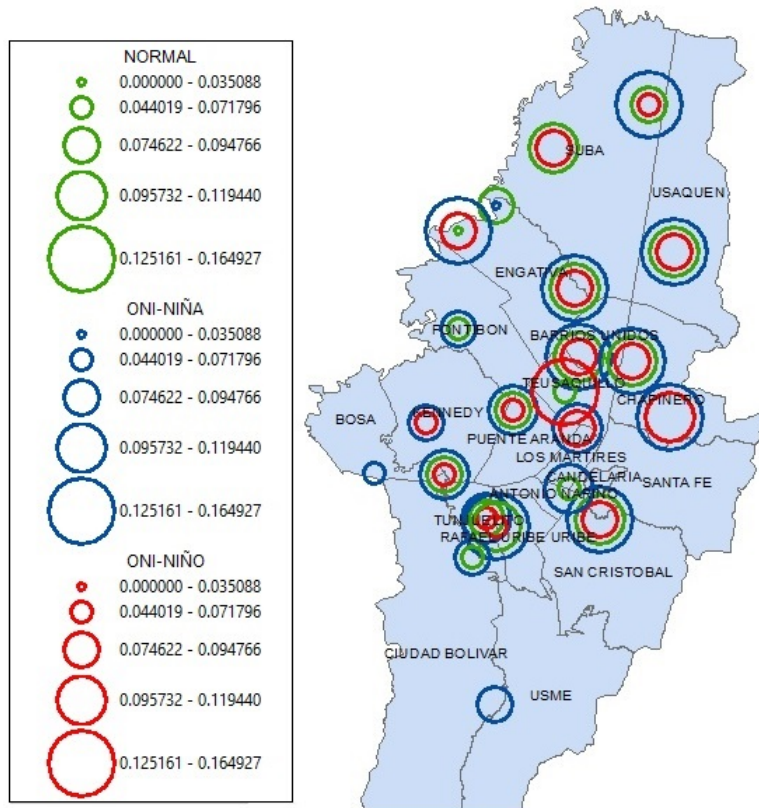
ENSO events also manifest in different ways, depending on the country. For Bogotá, we observe that El Niño is associated with less rain and higher temperature, while La Niña does the opposite, bringing more rain and lower temperature (see Figure 2.3). This pattern holds across the city. Figure 2.2 shows levels of rain for each monitoring station, where the size and color of the circles illustrate magnitude and ENSO condition. In addition, Figure 2.3 shows the rainfall (rain) and temperature (tmp) pattern for the 24 hours schedule for business days (Monday to Friday). It is evident that rainfall tends to be higher during La Niña episodes, particularly during the afternoon. These results are also confirmed in the Appendix 2B based on an additional regression analysis.

Air Quality information is also provided by RMCAB for similar monitoring stations of weather data.⁹ The network registers hourly readings and consists of four to ten monitoring stations depending on the air pollutant measured. The data spans the same period as the weather variables (August 1997 to December 2015). In many cases stations monitoring weather also measure pollutants. The list of pollutants includes particulate matter with a diameter less than 10 micrometers (PM_{10}), particulate matter with a diameter less than 2.5 micrometers ($PM_{2.5}$), carbon monoxide (CO), sulfur dioxide (SO_2), and nitrogen oxides (NO_x). As pollution may be affected by local weather and ENSO, in Appendix 2B we analyze how sensitive pollutant concentrations are to climatic conditions. This analysis suggests that not all the pollutants are affected in the same way. For instance, particles tend to be affected by extreme rain, whereas gases such as NO_x may respond to extreme temperature. It is consistent with what other studies suggest regarding the relationship between pollutants and weather, for instance rainfall helps to washout particles (see Guo et al. (2016)), while ambient temperature plays an important role in NO_x formation in urban environments (see Ko et al.

⁸Wind Direction Rose: 1 "WD: N 337.5-22.5" 2 "WD: NE 22.5-67.5" 3 "WD: E 67.5-112.5" 4 "WD: SE 112.5-157.5" 5 "WD: S 157.5-202.5" 6 "WD: SW 202.5-247.5" 7 "WD: W 247.5-292.5" 8 "WD: NW 292.5-337.5"

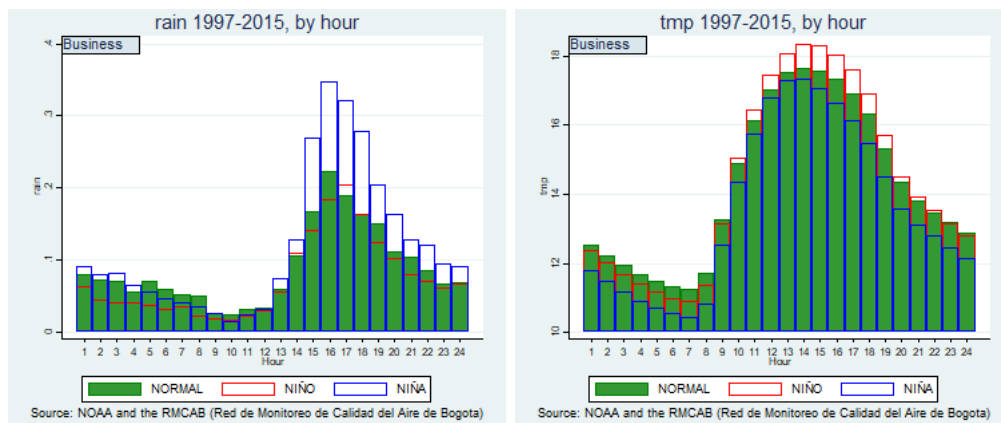
⁹List of monitoring stations for the RMCAB: Usaquén, Carvajal, Tunal, Simón Bolívar, Ferias, Cazuca, Guaymaral, Kennedy, Chicó-Lago, Suba, Cade- Energía, Puente Aranda, Fontibón, San Cristobal, Olaya y OPSIS, Universidad Nacional, Engativa and Central de Mezclas. All the stations capture traffic pollution. Background pollution is also captured in all the stations, except for Ferias, Guaymaral, Cade- Energía, Puente Aranda and Central de Mezclas. Industry stations are Carvajal, Cazuca, Puente Aranda and Fontibón. The classification given by type of monitoring stations are for year 2009 and they seem to capture similar type for air pollution (see [link document](#)).

Figure 2.2: Localities and Monitoring Stations in Bogotá



Note: Authors' elaboration based on data from the Air Quality Monitoring Network in Bogotá (RMCAB)

Figure 2.3: Weather variables-24hour schedule



Note: Authors' elaboration based on data from the Air Quality Monitoring Network in Bogotá (RMCAB)

(2019)). Those effects tend to occur in rush hours, periods where the population is more exposed to pollution because individuals are going to work or returning home. Therefore, we explore a set

of different intraday pollutant exposure measures in the subsequent analysis.

Information on health outcomes comes from the Vital Statistics Registrations gathered by the National Administrative Department of Statistics of Colombia (DANE). It provides data for each child born in the city over the period 1997-2015 on month of birth, birth weight, gestation length, gender, parents' characteristics, health care institution attending the birth, among others. Using birth weight (in grams) and gestation length (in weeks), we construct indicator variables for low birth weight and premature birth of the child. A child has low birth weight if its weight is less than 2500 grams. A gestation length lower than 38 weeks is considered as premature birth (see [Knittel et al. \(2016\)](#)).¹⁰

The databases described above are merged with information of geographical location of Health Service Providers (IPS). We follow an algorithm to match address and names of IPS listed in the Ministry of Health with health reports of DANE and computed all the variables at the IPS level. The procedure conducted to assign air pollutant concentrations to each child is explained in detail in the empirical strategy section.

2.3.1 Descriptive statistics

In [Table 2.1](#) and [Table 2.2](#), we present the descriptive statistics for the variables collected. In [Table 2.1](#), we observe that, on average, health centers attend to 18% of premature births and 11% of births with low weight. The average child's weight at birth in an IPS is 3016 grams. [Table 2.1](#) also shows high levels of pollutants during peak hours. The maximum average temperature is around 19° Celsius ([Table 2.2](#)). Although, [Table 2.10](#) in [Appendix 2A](#) presents all the eight categories of the Wind Direction Rose, only five are relevant for the sample ([Table 2.2](#)). Interestingly, wind blows mainly from south, south-east and south west, with differences across the city that play an important role in pollution transport. Mothers, on average, are 26 years old and have two children. 62% of mothers are married and 71% have secondary education. 66% of births are spontaneous deliveries and almost 50% of children are females ([Table 2.2](#)). The last panel of [Table 2.2](#) also shows

¹⁰Gestation length has four categories: 1=less than 22 weeks, 2=from 22 to 27 weeks, 3=from 28 to 37 weeks, 4=from 38 to 41 weeks and 5=from 42 or more. We used categories 1, 2 and 3 to define premature birth.

different transport measures implemented in the city during the last years.

Table 2.1: Descriptive Statistics for children (health outcomes and pollution)

	Observations	Mean	SD	Min	Max
ave-birth-weight (gr)	11655	3016.13	226.03	750.00	4750.00
ave-birth-height (cm)	11647	50.00	2.16	24.50	54.50
ave-weeks-ges (weeks)	11655	38.22	1.33	21.50	42.50
premat-ure (% per IPS)	11655	0.18	0.17	0.00	1.00
weeks-ges-premature	10426	32.35	0.56	21.50	32.50
low-birth-weight (% per IPS)	11655	0.11	0.13	0.00	1.00
premat-ure-low-birth-weight (% per IPS)	11655	0.08	0.12	0.00	1.00
weight-low-birth-weight	9898	2069.78	167.98	750.00	2250.00
apgar1 Index (1-10)	11651	5.43	2.61	1.00	10.00
apgar2 Index (1-10)	11650	6.05	3.21	1.00	10.00
Male to Female ratio in IPS	11110	1.05	0.51	0.00	9.78
pm10-peakm ($\mu\text{g}/\text{m}^3$)	15092	75.70	26.59	23.29	176.41
pm10-peaka ($\mu\text{g}/\text{m}^3$)	15092	54.91	17.88	15.55	131.24
pm10-max ($\mu\text{g}/\text{m}^3$)	15092	113.25	34.04	36.53	267.37
pm10-avma7 ($\mu\text{g}/\text{m}^3$)	15092	56.61	18.38	18.66	124.53
pm10-ave ($\mu\text{g}/\text{m}^3$)	15092	56.22	18.34	17.68	121.83
pm25-peakm ($\mu\text{g}/\text{m}^3$)	8139	36.16	12.48	7.43	68.83
pm25-peaka ($\mu\text{g}/\text{m}^3$)	8139	22.77	6.14	5.94	44.70
pm25-max ($\mu\text{g}/\text{m}^3$)	8152	53.41	16.11	12.96	92.03
pm25-avma7 ($\mu\text{g}/\text{m}^3$)	8144	26.31	8.62	5.42	48.84
pm25-ave ($\mu\text{g}/\text{m}^3$)	8152	26.06	8.69	5.67	48.99
o3-peakm (ppb)	14527	9.46	4.16	2.33	30.03
o3-peaka (ppb)	14526	12.91	3.95	2.30	43.76
o3-max (ppb)	14529	29.60	12.36	4.99	81.17
o3-avma7 (ppb)	14453	12.67	3.76	2.32	36.35
o3-ave (ppb)	14529	12.57	3.78	2.64	35.36
co-peakm (ppm)	14573	1.98	1.18	0.36	6.68
co-peaka (ppm)	14552	1.48	0.88	0.24	5.33
co-max (ppm)	14647	2.81	1.48	0.39	9.03
co-avma7 (ppm)	14329	1.42	0.85	0.22	4.77
co-ave (ppm)	14647	1.43	0.84	0.23	4.80
so2-peakm (ppb)	15092	9.84	6.92	0.66	47.24
so2-peaka (ppb)	15092	6.69	5.27	0.45	38.06
so2-max (ppb)	15092	13.91	10.93	0.93	94.08
so2-avma7 (ppb)	15080	7.12	5.16	0.47	34.01
so2-ave (ppb)	15092	7.06	5.14	0.47	32.71
no-peakm (ppb)	14577	45.51	20.02	3.30	120.15
no-peaka (ppb)	14577	18.24	11.46	2.03	135.68
no-max (ppb)	14578	71.70	30.36	3.70	166.48
no-avma7 (ppb)	14339	22.89	10.46	3.85	89.86
no-ave (ppb)	14578	22.70	10.82	3.23	91.81
no2-peakm (ppb)	14567	22.16	8.56	1.97	74.27
no2-peaka (ppb)	14567	18.20	5.85	1.36	49.77
no2-max (ppb)	14567	31.83	11.25	2.47	95.35
no2-avma7 (ppb)	14352	17.22	5.35	1.70	47.28
no2-ave (ppb)	14567	17.01	5.53	1.38	47.97
nox-peakm (ppb)	11521	64.85	27.02	4.83	169.26
nox-peaka (ppb)	11521	35.65	15.30	5.97	144.14
nox-max (ppb)	11522	91.65	36.94	6.54	213.54
nox-avma7 (ppb)	11283	38.90	14.48	8.52	109.04
nox-ave (ppb)	11522	38.43	15.20	5.24	115.47

Note: based on the Air Quality Monitoring Network Data for Bogotá (RMCAB) and Vital Statistics from DANE, aggregated by the Health Service Providers (IPS). Micrograms per cubic meter ($\mu\text{g}/\text{m}^3$), parts per million (*ppm*) and parts per billion (*ppb*). For particulate matter, sizes are expressed in micron or micrometer.

Table 2.2: Descriptive statistics for socio-economic variables, weather, ENSO events and other controls

	Observations	Mean	SD	Min	Max
rain-sum (ml)	15092	2.28	1.14	0.14	10.15
rain-suma7 (ml)	15092	2.30	1.14	0.08	8.34
rain-peakm (ml)	15092	0.03	0.02	0.00	0.29
rain-peaka (ml)	15092	0.17	0.12	0.01	1.19
rain-max (ml)	15092	1.30	0.63	0.08	4.35
ttmp-peakm (celsius)	15013	13.11	1.06	8.74	16.38
ttmp-peaka (celsius)	15013	15.30	1.13	11.39	18.59
ttmp-min (celsius)	15034	10.45	1.28	5.20	13.91
ttmp-max (celsius)	15034	18.91	1.26	14.23	22.92
ttmp-avma7 (celsius)	15031	14.10	1.04	10.35	16.98
ttmp-ave (celsius)	15034	14.11	1.05	10.34	17.05
ws-peakm (meters/second)	15092	1.15	0.42	0.18	3.44
ws-peaka (meters/second)	15092	2.06	0.69	0.37	4.60
ws-max (meters/second)	15092	3.52	1.08	1.07	9.00
ws-avma7 (meters/second)	15092	1.57	0.53	0.37	3.46
ws-ave (meters/second)	15092	1.57	0.54	0.37	4.23
wd-rose==NE: 22.5-67.5	15092	0.00	0.04	0.00	1.00
wd-rose==E: 67.5-112.5	15092	0.01	0.11	0.00	1.00
wd-rose==SE: 112.5-157.5	15092	0.24	0.43	0.00	1.00
wd-rose==S: 157.5-202.5	15092	0.59	0.49	0.00	1.00
wd-rose==SW: 202.5-247.5	15092	0.15	0.36	0.00	1.00
numconsul	11576	5.98	2.07	0.00	30.00
N. Pregnancies	11642	2.14	0.53	1.00	12.00
N. children	11639	1.94	0.49	1.00	11.00
father-age	11593	29.54	3.46	14.00	60.00
max-mother-age	11653	27.78	3.31	14.00	49.00
min-mother-age	11653	23.78	3.31	10.00	45.00
ave-mother-age	11653	25.78	3.31	12.00	47.00
Type delivery-spontaneous	11650	0.66	0.25	0.00	1.00
Type delivery-cesarea	11650	0.31	0.23	0.00	1.00
Type delivery-instrumented	11650	0.03	0.05	0.00	1.00
More deliveries (2-3-4)	11627	0.02	0.04	0.00	1.00
Partner's marital status	11628	0.62	0.27	0.00	1.00
seg-social-contributive	11611	0.48	0.46	0.00	1.00
seg-social-subsidized	11611	0.28	0.32	0.00	1.00
seg-social-others	11611	0.16	0.27	0.00	1.00
seg-social-uninsured	11611	0.07	0.17	0.00	1.00
Father's edu.-none	11568	0.00	0.04	0.00	1.00
Father's edu.-primary	11568	0.12	0.14	0.00	1.00
Father's edu.-secondary	11568	0.67	0.22	0.00	1.00
Father's edu.-terciary	11568	0.20	0.22	0.00	1.00
Mother's edu.-none	11627	0.00	0.03	0.00	1.00
Mother's edu.-primary	11627	0.11	0.14	0.00	1.00
Mother's edu.-secondary	11627	0.71	0.23	0.00	1.00
Mother's edu.-terciary	11627	0.18	0.24	0.00	1.00
female	11655	0.49	0.17	0.00	1.00
male	11655	0.51	0.17	0.00	1.00
Min dist. IPS to station	46872	1.97	1.30	0.06	9.41
No-car-day during the quarter	15198	0.08	0.15	0.00	0.67
Transmilenio Phase I	15198	0.16	0.36	0.00	1.00
Transmilenio Phase II	15198	0.51	0.50	0.00	1.00
Transmilenio Phase III	15198	0.18	0.38	0.00	1.00
Peak-and-plate change I	15198	0.61	0.48	0.00	1.00
Peak-and-plate change II	15198	0.19	0.38	0.00	1.00
Peak-and-plate change III	15198	0.18	0.38	0.00	1.00
nino	15198	0.22	0.39	0.00	1.00
nina	15198	0.35	0.45	0.00	1.00

Note: based on the Air Quality Monitoring Network Data for Bogotá (RMCAB) and Vital Statistics from DANE, aggregated by the Health Service Providers (IPS).

2.4 Method and empirical strategy

The identification procedure uses a system of equations to evaluate first, the effect of pollution on health using an instrumental variable approach and second, estimate the impact of ONI (Niño-Niña) on weather variables, on pollution and on health outcomes. We divide the analysis in two parts: 1) an hourly analysis (Appendix 2B) of the ONI index, as well as El Niño and La Niña dummy variables by monitoring station for weather and pollution variables; 2) a quarterly analysis (subsection 2.4.1) by health center, matching weather and pollution variables with health data (our main estimation). Although secondary, the hourly analysis provides guidance for the second part. Hence, the main focus of our study is the quarterly analysis, which is shown below in subsection 2.4.1.

2.4.1 Identification of health effects

The analysis uses the administrative registers of all births observed in the city from 1998 to 2015. This enables us to capture a long period of time and many different events (NIÑO-NIÑA). All the weather and pollution variables are aggregated at the level of health center and by quarter, grouping children that were born in the same month of the year. This allows us to separate the effects of the levels of pollution during each quarter of gestation of the child. We will proceed by first, describing how we assign pollution to each childbirth and then, discuss the identification we propose.

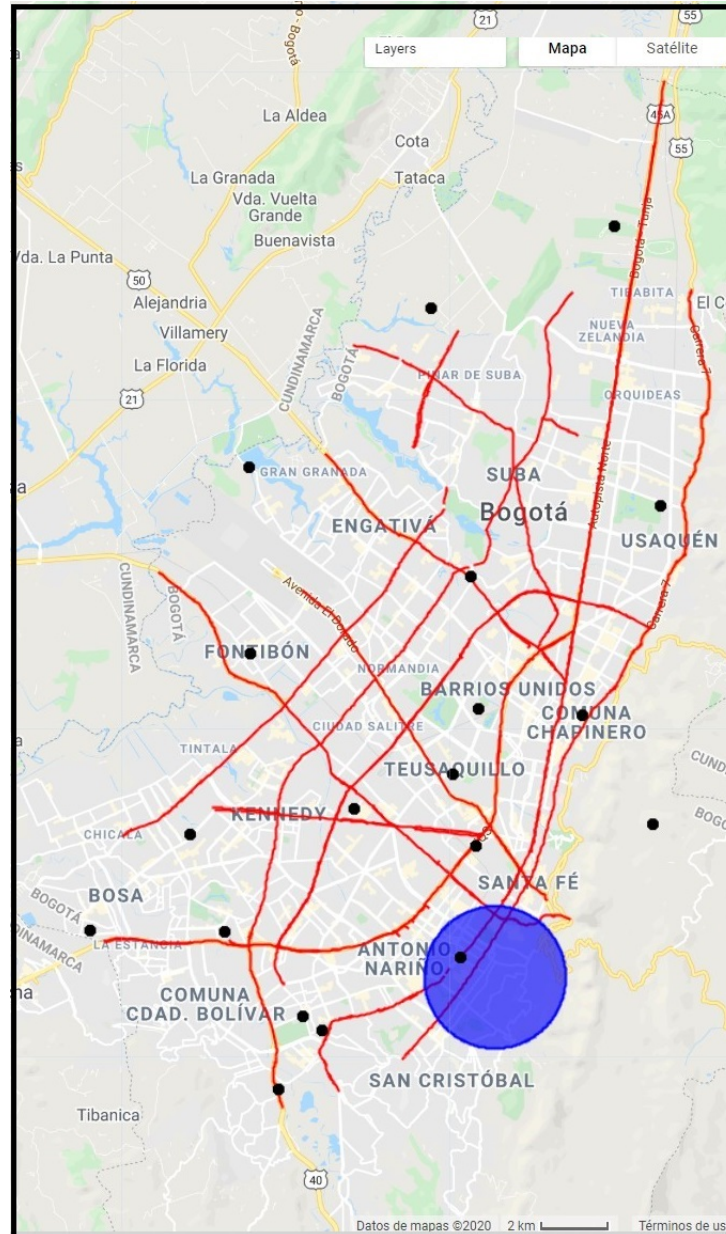
Assigning pollution and weather to each childbirth

Assigning the pollution and weather to health outcomes constitutes a challenge. To start, weather and pollution are aggregated for each monitoring station, and then by: average by day, average during the morning peak by day (7am-9am), average during afternoon peak by day (5pm-7pm), maximum during the day and average in the last seven days. The analysis of the effects of Niño-Niña on pollution and on weather in Appendix 2B allows us to understand in which period of the day the variables were more affected by the phenomena and aggregate the variables in similar way. For rainfall, we use the cumulative sum instead of the average by day.

We then have a daily panel of monitoring stations. The data-set is imported to Google Earth Engine and it uses the inverse distance weighting interpolation to estimate the levels of pollution and

weather for the whole area of Bogotá and for each month of the data-set. This data-intensive procedure creates for example, a map of PM_{10} in January 2000, another map for February 2000 and so on.

Figure 2.4: Map to assign weather and pollution to each health center



Note: 15 main roads (in red) in the city and health centers' addresses. This is an example for one health center (the blue dot), making a buffer of 2 km to give an idea about the area imputed. The black dots correspond to the location of the monitoring stations.

As observed by [Graff Zivin and Neidell \(2013\)](#), assigning climatic variables and air pollutants

to individuals' health outcomes according to the nearest monitoring station, could lead to measurement error, since it assumes that people are affected by pollution of that station, and that pollutants spread homogeneously over space. This justifies the use of the inverse distance weighting interpolation, since it allows us to use as much daily information as possible, in order to create monthly maps of the city.¹¹

As a next step, ideally, we should assign pollution and weather of the monthly maps to the child's mother's address. Unfortunately, access to this was denied and we have to use another strategy, by using the address of the IPS health center where the mother was giving birth.¹² As the average minimum distance of health centers to the monitoring stations is 2km (see Figure 2.4), we designed a buffer of 2km around each health center. For instance, we take the health center in the blue dot in the map and create a 2km buffer, and intersect it with the monthly maps of pollution and weather. This data-set is then matched to each child according to the health center, their respective month-year in which they were born and the specific months of gestation (time of gestation varies per child and it is accounted for).¹³ The data-set is then collapsed by running quarter of gestation and IPS, to create a quarterly panel of health centers from 1998 to 2015 (see the next section for the details about the running quarter).

Although this methodology assumes that individuals move or live inside the buffer of 2km, which

¹¹Kriging or other techniques of spatial forecasting based on the monitoring stations information are alternatives to the inverse distance weighting interpolation.

¹²We use the sample of IPS health centers from the Integrated System of information for the Social Protection (SISPRO) for the city. After cleaning the data-set for repeated names of health centers with the same address, etc., we ended up with 1268 health centers. As not all the health centers attend births, only 217 match with our sample of health centers of administrative registers of births in the city for which we have the location. However, the registers of birth in those 217 health centers correspond to 87.7% of the total sample of births from 1998 to 2015. As the data of health centers represent a census of all the births during the period in the city, attrition could come from two sources: attrition for reason one, for children not being delivered at the health center and attrition for reason two, for children born in a health center but for which we could not recover the address. Figure 2.10 in Appendix 2C shows the percentage of attrition of both sources by month. We observe first, that attrition for reason one is very low and second, the total number of attrition for reason two is stable by month and around 10%.

¹³As the exact day of childbirth was not provided (only month and year), the birth of the child it is set as the 15th of that month. Also, as the variable weeks of gestation is recorded in brackets (22-27 weeks, 28-37 weeks, 38-41 weeks, etc.), the average in the bracket is used to determine the months of gestation. For instance, a child was born on the 15th of September of 2004 with 38-41 weeks of gestation, so she had on average 39.5 weeks of gestation, or $39.5 * 7 = 276.5$ days of gestation. We can subtract that from the day of birth and infer the first day of gestation, which is the 13th of December of 2003. The period between the first day of gestation and the birthday gives approximately nine months of gestation in total. This allows us to take into account that some children can have less than nine months of gestation. As a next step, we create a panel for each child and her months of gestation, and match it with the respective month in which pollution and weather affected the child contemporaneously.

could be a strong assumption given the mobility of people in Bogotá, aggregating by health center can circumvent in some way this issue, by allowing us to express the health outcomes on average for the children that were born in that health center. We consider this as the best alternative, given the restriction of not having access to the exact address of the mother giving birth. We test the robustness of the results to this aspect by changing the buffer size to 4km in a later section.

Identification strategy

The identification relies on the variation in weather, pollution and NIÑO-NIÑA events at the quarterly level, which allows us to disentangle in which quarter of gestation the effects are more critical and can affect more the children. It also permits to capture concentration and time of exposure to the concentration of pollutants, in the sense that mothers might face similar levels of pollution and NIÑO-NIÑA events, but the time exposure (and doses of pollution) can affect more the baby during some particular quarters of gestation. As such, the system identifies not only the effects of ENSO events on pollution and health, but also the total effects of pollution on health during normal days, which is very relevant from a policy-maker's perspective. Hence, the system of equations for pollution ($P_{j,t-k}$) and health ($H_{j,t}$), aggregated by health center j is:

$$\begin{aligned}
P_{j,t-k} = & \alpha_0 + \sum_{q=0,3,6} \left(\alpha_{1,q} NI\tilde{N}O_{t-q} + \alpha_{2,q} NI\tilde{N}A_{t-q} + \alpha_{3,q} WR_{j,t-q} \right) \\
& + \sum_{q=0,3,6} \left(\alpha_{4,q} WS_{j,t-q} + \alpha_{5,q} WS_{j,t-q} \times WD_{t-q}^{rose} + \kappa_W D_{t-q}^{rose} \right) \\
& + \eta_1 X_{j,t-k}^{hh} + \eta_2 X_{t-k}^{fe} + \eta_{3,q} \sum_{q=0,3,6} T_{t-q} + \epsilon_{j,t-k} \quad \forall k = 0, 3, 6 \quad (2.1)
\end{aligned}$$

$$\begin{aligned}
H_{j,t} = & \theta_0 + \sum_{k=0,3,6} \theta_{1,k} P_{j,t-k} + \sum_{q=0,3,6} \left(\theta_{2,q} NI\tilde{N}O_{t-q} + \theta_{3,q} NI\tilde{N}A_{t-q} \right) \\
& + \sum_{q=0,3,6} \left(\theta_{4,q} WR_{j,t-q} + \theta_{5,q} WS_{j,t-q} \right) + \lambda_1 X_{j,t}^{hh} + \lambda_2 X_t^{fe} + \epsilon_{j,t}
\end{aligned}$$

where $P_{j,t-k}$ corresponds to the different variables of pollution averaged for each health center j at running quarter t , with $k = 0, 3, 6$, which corresponds to the third, second and first quarter of gestation, Q3, Q2, Q1, respectively.¹⁴ $WR_{j,t-q}$ are the different variables of rain and tempera-

¹⁴ t is constructed as a running quarter according to the month of birth of the children born in a specific month. If the children were born in March 2000, for instance, t corresponds to the average for January-February-March of 2000 (Q3 or last quarter of gestation), $t - 1$ stands for December 1999-January 2000-February 2000. In the same way, $t - 3$ corresponds to October-November-December of 1999 (Q2 or second quarter of gestation), and $t - 6$ stands

ture constructed previously at health center j and quarter q (rain and temperature included at the same time); $NI\tilde{N}O_{t-q}$ ($NI\tilde{N}A_{t-q}$) capture the percentage of months of *niño* (*niña*) during the quarter (3/3, 2/3, 1/3, 0); $WS_{j,t-q}$ is the wind speed for the quarter and $\kappa_{WD_{t-q}^{rose}}$ a rose of wind direction fixed effects, the instruments for the pollution equation;¹⁵ $X_{j,t-k}^{hh}$ is a vector of household characteristics (father’s and mother’s average age, father’s and mother’s dummies for education level, partner’s marital status, type of delivery, multiple deliveries, number of previous pregnancies, number of children, number of consultations, and type of social security of the mother); X_{t-k}^{fe} a set of fixed effects (*locality* \times *year*, month, health center); and T_{t-q} a set of variables for transportation changes (3/3, 2/3 1/3, 0 for having no-car-day during the months of the quarter, for Transmilenio phases or for changes in the peak and plate measures). These transportation variables also enter in the system as instruments. $H_{j,t}$ are the average of the different health outcomes of the children born in the running quarter t , which are only observed at the moment of the birth. The channel of pollution on health was already found relevant for childbirth outcomes in previous studies (Currie et al. (2009) and Lavaine and Neidell (2017), among others).

Notice that ENSO events might influence health outcomes through different channels in our context: (1) ENSO affects weather, which may affect health; (2) ENSO influences air pollution, which may affect health; (3) there may be a direct effect of ENSO on health; and (4) ENSO affects the economy through changes in food markets inducing variations in household income and consumption, or other economic variables, which also may affect health. In this sense, we consider that the system of equations we propose is able to identify several of those channels. The coefficient of ENSO events in the second stage of equation 2.1 might be interpreted as an overall estimate of the effects from channel (1), and partial effects from channels (3) and (4), while the coefficient of pollutant variables measure the impact of pollution exposure. We argue that ENSO events coefficients might reflect partial effects of economics shocks because we are not able to structurally separate such shocks from other channels. Furthermore, in order to capture the total compounded effect

for July-August-September of 1999 (Q1 or first quarter of gestation). The aggregation of the data-set by running quarter and health center takes into account the children who were alive based on the months of gestation; if four children were born in health center j in March 2000, but two had gestation time of six months and the two other nine months, the variables in Q1 take into account only the children with nine months of gestation, while Q2 and Q3 takes into account all the four children.

¹⁵Table 2.10 in Appendix 2A shows the categories of the Wind Direction Rose. Only the categories in Table 2.2 are relevant for the estimation and in all the estimations, category two is the baseline wind direction.

of the four different channels of ENSO events on health, we proceed by estimating a reduced-form version of equation 2.1 without including pollution, or weather factors. The section of results discusses further the reduced-form estimations, and then, moves towards the estimation of the different channels of ENSO events on health through the IV approach.

In practical terms, the system of equations 2.1 could be estimated through a model of simultaneous equations, by using a General Structural Equation Modelling (Eleftherios and Ozgur (2016)), by using mediation analysis (Laffan (2018), Barrett et al. (2017)), or by using an instrumental variables approach, with the rose of wind direction as the instrument for the endogenous variable of pollution (see Deryugina et al. (2019) and Schlenker and Walker (2016) for articles instrumenting pollution with wind direction). Our estimations follow the latter approach, and use standard errors robust to heteroscedasticity and autocorrelated cross-panel disturbances (see Driscoll and Kraay (1998)). Notice also that the variables of transportation changes T_{t-q} are only included in the pollution equation and as such, they also enter in the system as instrumental variables. However, our identification strategy relies more on the variation coming from wind direction, as it can be seen in Appendix 2D. Appendix 2D shows that there is not only cross-section variation in wind direction and wind speed but also variation over the years. In the map of Appendix 2D, we can see that the wind blows from the east (mountains), towards the west and the south.¹⁶

Importantly, the system captures the lagged effect on health of the exposure to pollution, $NI\tilde{N}O$ and $NI\tilde{N}A$, during the three quarters of gestation. The first stage of the IV approach is estimated for three endogenous variables $P_{j,t}$, $P_{j,t-3}$, $P_{j,t-6}$, one for the exposure during each quarter of gestation. According to the literature, it is important to consider pollution and weather during the different quarter of gestation, e.g., Currie et al. (2009) who use CO and PM_{10} and by Barreca et al. (2015) and Deschênes et al. (2009), who use the temperature of the different quarters. We use both in the same estimation and add the effect of ENSO during the different quarters of gestation.

The previous system also allows us to capture the effect of El Niño and La Niña events, di-

¹⁶An alternative to the rose of wind direction in some studies consist of using a dummy variable for upwind or downwind direction.

rectly in the equation of pollution and health, considering them as exogenous global phenomena (see Hsiang et al. (2011) and Dingel et al. (2020)). Although they could disturb weather conditions, affecting the levels of pollution, they could also affect, for instance, food prices or other economic variables, which might impact the health of pregnant mothers, and hence, their children. As such, ENSO events do not change with respect to location as they are considered as global phenomena, affecting several countries like Colombia. Therefore, there is no spatial variation across locations in Bogotá of the ENSO events captured here (see Figure 2.1). Once we control for the set of fixed effects ($locality \times year$, month, health center), the identification of ENSO events in our main equation relies more on the time variation of what we define as running-quarter. This is an advantage that we have in the data, varying over 1998-2015 by quarter of gestation of the children and by month of birth (cohort). In studies such as Dingel et al. (2020), the identification of ENSO events also relies on the variation over time but the impacts on trade depend on the latitude-longitude of the location and are spatially correlated.

The system also includes the effect of classical local weather on health as it could be that higher levels of rainfall or temperature affect negatively the health of the pregnant mother and hence, the infant (Knittel et al., 2016; Barreca and Schaller, 2020). They find an impact of high temperatures on delivery timing and gestational lengths at the level of county in the United States. The argument is that heat increases the levels of oxytocin, the hormone regulating the onset of delivery; alternatively, extreme heat might also cause earlier deliveries via cardiovascular stress (see for instance Hampel et al. (2011)). Some interesting heterogeneous effects are found by Barreca and Schaller (2020), so the cumulative effects are particularly higher in cold counties while hot counties might be more accustomed to extreme heat and can adapt better. In the context of India, high temperature should also affect directly and indirectly (via pollution) health outcomes. Although in Bogotá temperatures does not reach such extreme values, it might be that pregnant mothers are affected by extreme rainfall, which is more likely during La *Niña*.

2.4.2 Potential threats to identification

In the literature, studies such as Hanna and Oliva (2015) and Anderson (2014) analyze the effects of pollution on health by using a single shock to an emission source like an industry strike or a plant

closure. Also, [Knittel et al. \(2016\)](#) assign pollution of the nearest station to particular individuals, and solve the issue of measurement errors by instrumenting pollution with weekly shocks to traffic, including weather conditions as additional confounders. Our strategy relies more on the use of wind direction to instrument pollution and exploit NIÑO-NIÑA as exogenous sources of variation. The problem of using a single shock such as a plant closure or unexpected changes in traffic, is that it could affect multiple pollutants, putting some threats to the identification, unless some extra assumptions are considered (see [Graff Zivin and Neidell \(2013\)](#)). This justifies the use of weather variables as additional confounders that help in identifying the effects of different air pollutants.

[Currie et al. \(2009\)](#) study the effect of pollution on infant health in New Jersey. They assign the closest air quality monitoring station variables to the mother's location and finds a negative impact of exposure to carbon monoxide (CO) during and after birth, with larger effects for the group of smokers and older mothers. By analyzing infants, they can capture a causal effect more immediately; we follow their approach in this article.

Residential sorting could be a potential problem (see [Graff Zivin and Neidell \(2013\)](#)) as families might try to locate to areas where air quality is less contaminated, with better education and health care access, leading us to underestimate the effects of pollution. [Currie et al. \(2009\)](#) propose to solve this issue through fixed effects for mothers and station. For this reason, our system of equations also control for a rich set of parents' characteristics. However, individuals move in Bogotá mainly after experiencing income shocks, and hence, moving to areas with less pollution should be uniquely a secondary aspect to consider in relocation decisions after an income shock (see [Table 2.22](#) in [Appendix 2G](#)). Locations can also differ in terms of access to Primary Care Physician, income, access to transport, etc., and to that aim, the system of equations [2.1](#) also includes health center fixed effects and locality \times year fixed effects. In addition, we explore the residential sorting issue in more detail in subsection [2.5.3](#).

A more important challenge for the empirical strategy is the assignment of pollution and weather variables to the children based on the health center's location using a radius of 2km. Unfortunately, mothers' addresses were not provided and we consider much better to aggregate the data-set at

the level of health center. We explore some robustness checks using a larger radius, or aggregating the data-set at the level of locality instead of health center in Appendix 2I and Appendix 2N, respectively. By using a larger buffer or aggregating at locality level, we give a decreasing weight to the farthest measurement stations, and increase the likelihood that the measure better characterizes the air pollution level in the area where the mother is supposed to live/work/sleep during pregnancy.

Another point with respect to the implicit assumption that women live close to the health centers is how mothers choose the place of birth. Mothers can choose the health center to give birth in two ways: by planning and by not planning. In the first case, mothers can organize the delivery in advance. As this is a risky situation for both, mother and baby, they might try to avoid choosing a health center very far from where they live. It is well known that the traffic in Bogotá is very complicated and they might want to avoid risking the baby's life on the day of delivery by choosing something not too far, and closer to where they live. Although the second situation is possible, it would be surprising to see that this happens in the majority of the deliveries in the city, but still, the mother could try to go to the planned health center of delivery. In both cases, mothers try to avoid unnecessary movements around the city in the last weeks of gestation and might try to stay in an area not far from where they live, and probably, closer to the planned place of delivery. Another potential threat is that the mother's exposure could be different at home, or work, etc.. However, this is less a concern in our case as pregnant women tend to move less in the city. It would have been a more important threat in the case of estimating effects of pollution on individuals that move more.

Finally, there is another issue with respect to the temporal scale. [Graff Zivin and Neidell \(2013\)](#) mention that some air pollutants like ozone have nearly immediate effects after 1-2 hours, some have incubation periods and others could have both immediate and delayed effects. For this reason, we aggregate the pollution variables as a running quarter, split the effects by quarter and analyse several different pollutants.

2.5 Results and robustness checks

The section of results is organized in the following way. We start by estimating the total compounded effect of ENSO on health in a reduced-form style and describe the mechanisms we aim to capture. This will give a better understanding of the main contribution of the article, the estimate of the El Niño/La Niña impacts on birth outcomes. Then, we proceed with the instrumental variable approach, starting with the first stage. As air pollution is endogenous and affects health outcomes at birth, we instrument it with wind variables. This is followed by the discussion of the main effects we capture on health in the second stage, coming directly from ENSO events and from pollution. The next subsections present some robustness checks, discuss some threats to the identification, show some additional tests, heterogeneous effects, and check effects on additional health outcomes.

2.5.1 Total ENSO effect on health

In order to assess the total compounded effect of ENSO events on health, we proceed by making a reduced-form estimation of equation 2.1 without including pollution or weather factors to capture the total composed effect of ENSO. As we argue, ENSO might affect health outcomes through different channels (in our context) and the reduced-form estimations capture the four different channels as a compounded effect of ENSO on health.

Table 2.3 summarizes the results of these estimations. It shows that ENSO events increase the percentage of low weight at birth and decrease the average weight of the children at birth in the health centers. The results are different depending on the quarter, with statistically significant effects at conventional levels from El Niño during the first and third quarter of gestation, and from La Niña during the second quarter. However, there is no effect of ENSO on the other health outcomes considered. Using the rows NINO and NINA (cumulative effect) in Table 2.3, we can say that being exposed to El Niño (or La Niña) during pregnancy increases the percentage of children with low weight at birth in the health centers by 1 percentage points. This is an increase of 9.1% of low birth weight children delivered on average (from a base of 11 percentage points of low birth weight in the IPS-health centers). The result is bigger compared to [Brando and Santos \(2015\)](#) who use the Colombian panel survey ELCA. In their article, an average of nine more days of La Niña

rainfall exposure *in utero* increases the probability of being a low birth weight baby by 2.92% (one standard deviation of 16.2 more days of exposure *in utero* increases the probability of being a low birth weight baby by 4.7%). Consequently, being exposed to El Niño during pregnancy (cumulative effect) decreases the average weight of children in the health centers by 23 grams, which is a decrease of 0.8% from a baseline average of 3016.1 grams (see Table 2.1). The cumulative effect of La Niña on birth weight is not statistically significant at conventional levels, but statistically significant at the 5% significance level during the second quarter.

Table 2.3: REDUCED FORM: ENSO EFFECTS ON HEALTH

	RF (ENSO)			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)
Niño-Q3	0.008 (0.005)	0.009* (0.005)	-0.061 (0.039)	-23.443*** (7.658)
Niño-Q2	-0.005 (0.006)	-0.004 (0.004)	0.034 (0.045)	8.008 (7.114)
Niño-Q1	0.003 (0.006)	0.008* (0.005)	0.016 (0.044)	-7.670 (7.977)
Niña-Q3	0.004 (0.006)	-0.000 (0.004)	-0.004 (0.043)	2.060 (7.684)
Niña-Q2	0.001 (0.006)	0.011*** (0.004)	-0.020 (0.042)	-20.784** (8.758)
Niña-Q1	-0.006 (0.008)	-0.002 (0.005)	0.061 (0.058)	5.253 (8.816)
Observations	11193	11193	11193	11193
NINO	0.01	0.01	-0.01	-23.10
P_NINO	0.36	0.01	0.84	0.01
NINA	-0.00	0.01	0.04	-13.47
P_NINA	0.92	0.21	0.63	0.14
r2_a	0.37	0.31	0.36	0.34

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

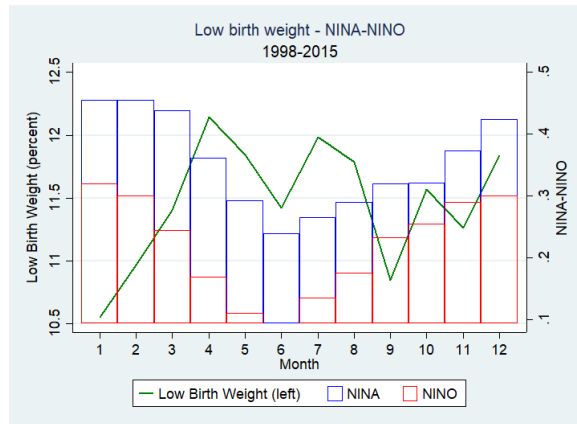
Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the IPS, Weeks gest. is the average of weeks of gestation for all the children born in that IPS and Birth Weight is the average weight for all the children born in that IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics mentioned in section 2.4.1. *NINO*, *NINA* correspond to the sum of the three coefficients (the cumulative effect) of the variables for El Niño and for La Niña, while *P_NINO*, *P_NINA* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

In Figure 2.5, we show the relationship between ENSO and the different pollutants with health. Panel a) shows the relation of ENSO with low weight at birth, aggregating by the different months of the year. ENSO events are more frequently observed at the end and beginning of the years. We also notice a lagged effects in the sense that once ENSO events become more frequent after the months of September, there is an increase in the children born with low weight. This also happens

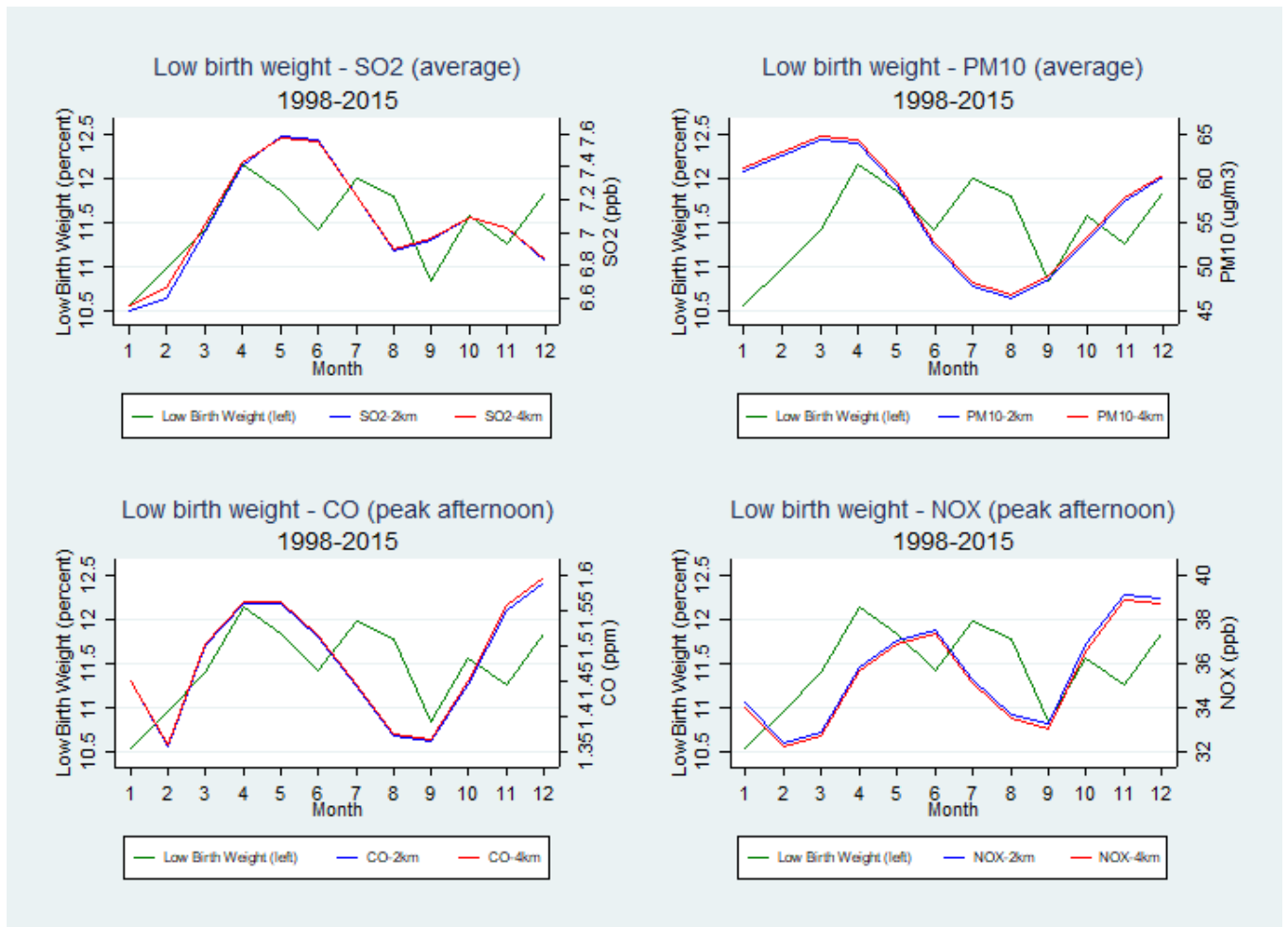
from January to March. In addition, ENSO events are less persistent from April to September, which coincides with a decrease of the low birth weight variable (with increases in some months). It is worthwhile to notice that ENSO can have not only an effect manifested as extreme climatic shocks on health, but also, an effect mediated through pollution. The relationship described in panel a) of the figure includes both aspects. In fact, the reduced-form estimations presented allow to understand better the total compounded effect of ENSO events on health outcomes (the sum of the four different channels discussed).

We can see the important effect of air pollution on health by relating the different pollutants with the variable of low birth weight in panel b) of Figure 2.5. It shows a positive correlation of air pollution with the low birth weight variable across the different months. The relationship is much clearer for CO , so months with higher emissions of CO such as April, go together with increases in the low weight at birth, and reductions of CO emissions in September coincide with lower values for the variables of low weight at birth. This pattern is also observed for PM_{10} and to a lesser extent for NO_X and SO_2 . Low birth weight tends to move more contemporaneously with air pollution, but more in a lagged manner with respect to ENSO events. This confirms that lagged terms should be included in the estimations.

Figure 2.5: ENSO and Pollutants effect on Low Birth Weight



(a) ENSO



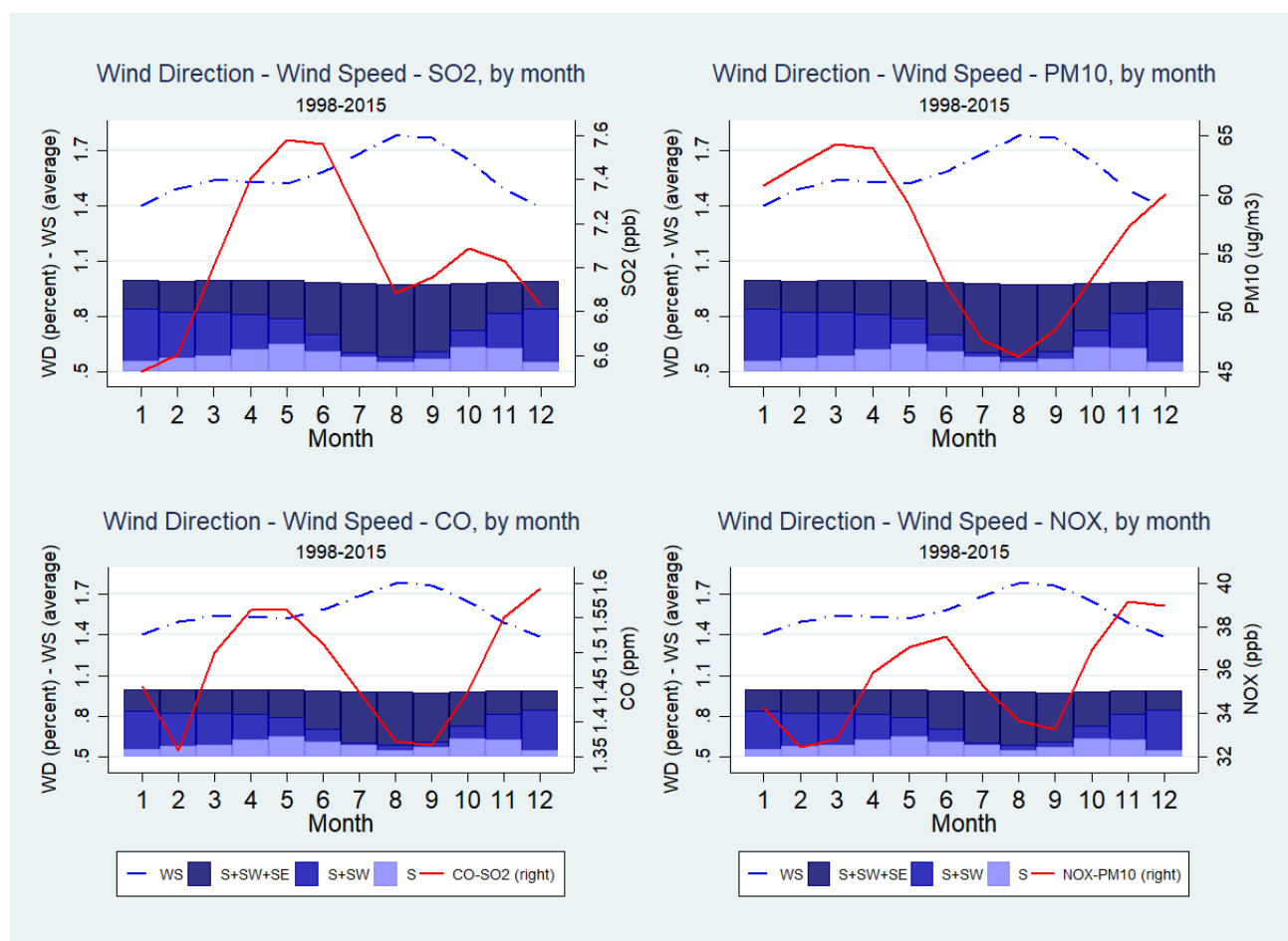
(b) Pollutants

Note: based on the Air Quality Monitoring Network Data for Bogotá (RMCAB) .

2.5.2 The Instrumental Variable approach: how ENSO affects health

After discussing the total compounded effect of ENSO events on health in the previous subsection, we move to the instrumental variable approach for the system of equations 2.1 using wind variables as instruments. Figure 2.6 describes this relationship between wind variables and four of the pollutants used, over the year.

Figure 2.6: Wind Direction (WD) and Wind Speed (WS) on Pollutants



Note: based on the Air Quality Monitoring Network Data for Bogotá (RMCAB). S: percentage of wind blowing from the south; S+SW: percentage of wind blowing from the south and the south-west; S+SW+SE: percentage of wind blowing from the south, the south-west and from the south-east.

During periods of high wind speed (WS) like July, August and September, the levels of pollution

tend to decrease across the four pollutants in the figure.¹⁷ For PM_{10} , and to a lesser extent for CO and NO_x , lower levels of wind speed in April to May and October to December come in hand with higher levels of pollution.

Figure 2.6 also shows the cumulative percentage when wind blows from the south (S), from the south and the south-west (S+SW) and from the south, the south-west and the south-east (S+SW+SE) between 1998 to 2015. These are the main directions from which wind blows in the city. As observed, when wind direction is predominantly from the south and the south-west, the levels of pollution tend to be lower. Similarly to wind speed, for the period between July to September, wind blows less from the south and the south-west. Once wind blows more from these directions, the levels of pollution tend to increase (April to May and October to December).

This discussion shows the importance of using the interactions between wind speed and the wind direction variables for the identification strategy, in order to have a richer set of instruments for pollution. We proceed by estimating the system of equations 2.1 using the instrumental variable approach. Tables 2.4 and 2.5 present the first-stage estimates that support the statistical correlation between the wind direction instruments with pollutant concentrations (interactions of the wind direction rose with wind speed not shown to save space). We estimate several exposure measures of pollution, but for ease of presentation we select exposure measures grouped by pollutants for which the results are statistically relevant. In the case of Table 2.4, we use average daily pollution to analyze impacts of SO_2 , PM_{10} and $PM_{2.5}$, while in Table 2.5, we employ pollutant concentrations during evening peak hours for CO , and NO_x . The second stage estimations presented later on are consistent with the same aggregation of pollutant that are used in the first stage. Almost all coefficients of the set of instruments are statistically different from zero at conventional significance levels. Additionally, the adjusted R-squared of the first stage accounts for the importance of the pollution variability explained by the wind variables and the exogenous variables (from 6% to 34%). The F-statistics across the majority of the first stage results are larger than ten, exceeding the rule of thumb cutoff for weak instruments proposed by Staiger and Stock (1997).

¹⁷August in fact, is the most preferred month for flying kites for the Bogotanos (demonym for the people from Bogotá) because of the very strong winds and better weather conditions. This is a common practice with several kites festival during the month.

Note that each regression of the health outcomes (second-stage) has three equations in the first stage, one for each quarter of gestation, and corresponds to the set of three columns shown in the tables of the first stage for each pollutant. In fact, the three equations estimated in the first stage are the same for the four health outcomes presented in the second stage. For instance, the first four columns of Table 2.6 (for SO_2) share the same equations in the first stage, and we present only once the results of the first stage (first three columns of Table 2.4). The same holds for the other pollutants.

Table 2.4: IV FIRST STAGE: IV ON SO_2 PM_{10} $PM_{2.5}$ (average)

	SO2 (average)			PM10 (average)			PM2.5 (average)		
	Q3-gest. (1)	Q2-gest. (2)	Q1-gest. (3)	Q3-gest. (4)	Q2-gest. (5)	Q1-gest. (6)	Q3-gest. (7)	Q2-gest. (8)	Q1-gest. (9)
Niño-Q3	0.119 (0.352)	0.868 (0.536)	-0.384 (0.368)	-3.260 (2.012)	0.528 (1.247)	1.552 (1.446)	0.718 (1.662)	-3.239*** (0.992)	-1.117 (1.741)
Niño-Q2	-1.438*** (0.384)	0.194 (0.578)	1.211*** (0.409)	-3.341* (1.838)	-5.029*** (1.597)	-1.261 (1.054)	2.091 (1.646)	-3.801*** (1.219)	2.538* (1.302)
Niño-Q1	0.074 (0.452)	-1.377** (0.538)	-0.580 (0.464)	2.100 (1.663)	0.044 (2.009)	-2.789 (2.278)	-0.814 (1.273)	-0.268 (1.291)	2.013 (1.230)
Niña-Q3	0.443 (0.381)	-0.507 (0.543)	0.147 (0.336)	-2.280 (1.984)	0.141 (1.566)	0.615 (1.948)	0.301 (1.745)	-2.297 (2.180)	9.584*** (2.033)
Niña-Q2	-0.038 (0.562)	0.127 (0.565)	-0.233 (0.609)	-4.828*** (1.798)	-2.146 (2.855)	-0.572 (1.917)	-2.892 (2.000)	-1.780 (1.792)	-4.144 (2.561)
Niña-Q1	1.769*** (0.471)	-0.617 (0.627)	1.459*** (0.511)	-0.429 (1.966)	-5.473*** (1.735)	-3.067* (1.856)	0.870 (1.844)	-5.882*** (1.401)	5.408** (2.272)
ws_ave-Q3	-4.957* (2.605)	1.195 (1.079)	-1.350 (1.061)	2.333 (5.475)	14.378*** (4.422)	2.177 (8.037)	15.899 (10.518)	-4.300 (10.620)	-18.063 (14.735)
ws_ave-Q2	-3.922** (1.889)	-3.605 (2.315)	1.114 (1.143)	1.991 (9.466)	3.743 (6.526)	17.605*** (3.123)	-13.113 (9.182)	12.678 (10.061)	24.149 (17.825)
ws_ave-Q1	1.772 (1.337)	-3.231 (2.617)	-4.848** (2.268)	-16.790** (6.749)	2.981 (5.190)	3.737 (10.034)	5.767* (3.416)	-3.998 (3.202)	-1.329 (2.869)
wd_rose_Cat3-Q3	-3.003 (2.510)	1.743 (1.357)	-1.891 (1.263)	1.295 (5.124)	6.169 (5.259)	-9.134 (8.182)	15.832 (12.189)	-8.628 (12.356)	-15.790 (17.510)
wd_rose_Cat4-Q3	-4.922* (2.629)	1.650 (1.441)	-2.041 (1.345)	0.883 (4.947)	10.141* (5.913)	-5.355 (8.360)	15.675 (12.310)	-5.879 (12.401)	-17.229 (17.253)
wd_rose_Cat5-Q3	-4.998* (2.678)	2.608** (1.329)	-1.528 (1.307)	4.738 (4.993)	11.408* (5.836)	-2.852 (8.283)	17.731 (12.309)	-5.484 (12.200)	-18.523 (16.864)
wd_rose_Cat6-Q3	-6.195** (2.754)	2.421 (1.508)	-1.818 (1.533)	11.906** (5.188)	17.877*** (6.246)	-0.919 (8.584)	17.223 (12.360)	-4.814 (12.198)	-18.439 (16.797)
wd_rose_Cat3-Q2	-4.513** (2.070)	-0.309 (2.354)	1.825 (1.379)	6.067 (10.611)	2.004 (6.374)	8.100** (3.667)	-12.853 (10.769)	10.488 (11.502)	24.630 (20.658)
wd_rose_Cat4-Q2	-4.067* (2.303)	-3.086 (2.557)	1.991 (1.426)	5.112 (10.522)	1.589 (5.979)	9.899*** (3.757)	-14.137 (10.598)	11.614 (11.295)	30.021 (20.796)
wd_rose_Cat5-Q2	-3.343 (2.225)	-2.438 (2.498)	3.132** (1.393)	8.649 (10.484)	4.587 (6.337)	8.358** (3.628)	-14.965 (10.494)	14.576 (11.447)	28.003 (20.676)
wd_rose_Cat6-Q2	-2.729 (2.310)	-3.330 (2.610)	2.778* (1.454)	6.356 (10.631)	7.748 (6.448)	13.019*** (4.183)	-17.704* (10.669)	13.193 (11.032)	29.860 (20.864)
wd_rose_Cat3-Q1	2.111* (1.150)	-1.803 (2.990)	-2.493 (2.710)	-20.433** (8.850)	1.097 (6.349)	-2.094 (9.701)	8.834** (3.795)	-3.140 (4.411)	-1.998 (3.162)
wd_rose_Cat4-Q1	2.332* (1.273)	-1.301 (3.291)	-5.172* (2.786)	-22.081** (8.608)	-0.374 (6.332)	-1.442 (9.757)	9.957** (4.234)	-3.816 (4.301)	-1.369 (3.293)
wd_rose_Cat5-Q1	2.808** (1.287)	-0.251 (3.225)	-4.567* (2.766)	-22.666** (8.840)	2.198 (6.399)	2.150 (9.785)	9.593** (4.310)	-3.843 (4.278)	0.495 (3.318)
wd_rose_Cat6-Q1	2.263 (1.418)	0.182 (3.310)	-5.762** (2.797)	-20.574** (9.245)	0.137 (6.496)	4.978 (10.026)	11.836** (4.684)	-7.329 (4.606)	1.070 (3.708)
<i>N</i>	11027	11027	11027	11027	11027	11027	4413	4413	4413
<i>R2_first_a</i>	0.16	0.14	0.08	0.11	0.09	0.06	0.20	0.22	0.34
<i>F_first_a</i>	27.88	19.79	9.71	14.27	10.61	6.62	11.12	20.32	42.68

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Pollution is averaged per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. Each regression of the health outcomes has three equations in the first stage, one for each quarter of gestation. In this case, the three equations are the same for each set of four health outcomes presented in the second stage, for instance, the first four columns of Table 2.6 (for SO_2) share the same equations in the first stage, so we will only present once the results. The same holds for the other pollutants. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay). The interaction of the wind direction rose with wind speed not shown to save space.

Table 2.5: IV FIRST STAGE: IV ON CO NO_X (peak afternoon)

	CO (peak afternoon)			NOX (peak afternoon)		
	Q3-gest. (1)	Q2-gest. (2)	Q1-gest. (3)	Q3-gest. (4)	Q2-gest. (5)	Q1-gest. (6)
Niño-Q3	0.064 (0.074)	-0.389*** (0.105)	0.163 (0.106)	-5.086* (3.051)	-5.279*** (1.663)	-2.075 (1.813)
Niño-Q2	-0.203* (0.118)	0.049 (0.111)	-0.357*** (0.131)	2.015 (3.153)	-4.439** (1.991)	-4.981*** (1.732)
Niño-Q1	-0.102 (0.082)	-0.310 (0.247)	0.276** (0.123)	0.553 (3.370)	-0.021 (3.278)	-6.844** (2.855)
Niña-Q3	-0.166* (0.089)	-0.252** (0.099)	0.347*** (0.116)	0.134 (2.148)	-4.478** (1.895)	5.413** (2.506)
Niña-Q2	-0.058 (0.071)	-0.222 (0.139)	-0.115 (0.109)	4.218 (2.775)	-1.425 (1.953)	-2.630 (2.818)
Niña-Q1	-0.487*** (0.096)	-0.124 (0.092)	-0.152 (0.096)	-6.105** (2.766)	6.502** (3.093)	2.797 (2.259)
ws_peaka-Q3	-0.183 (0.329)	0.591 (0.644)	-0.207 (0.262)	8.137 (6.451)	-5.280 (3.456)	4.886 (3.213)
ws_peaka-Q2	-1.098*** (0.356)	0.147 (0.313)	0.626* (0.369)	7.704 (4.775)	-3.593 (5.598)	-8.537*** (2.614)
ws_peaka-Q1	0.339 (0.253)	-0.447 (0.465)	0.587** (0.292)	-6.140 (4.434)	-3.234 (3.794)	-10.109* (5.424)
wd_rose_Cat3-Q3	-0.126 (0.524)	0.279 (0.840)	0.026 (0.338)	8.104 (9.230)	-11.673 (7.115)	9.575* (5.178)
wd_rose_Cat4-Q3	0.289 (0.490)	0.571 (1.016)	-0.478 (0.370)	17.147* (9.004)	-7.897 (6.761)	9.170* (5.126)
wd_rose_Cat5-Q3	0.245 (0.506)	0.559 (1.015)	-0.415 (0.383)	23.280** (9.551)	-6.840 (6.963)	10.439* (5.380)
wd_rose_Cat6-Q3	0.213 (0.515)	0.698 (1.027)	-0.350 (0.411)	25.464*** (9.450)	-5.638 (7.202)	12.001* (6.489)
wd_rose_Cat3-Q2	-1.636*** (0.612)	0.681 (0.465)	0.382 (0.516)	13.423* (7.245)	-1.398 (6.967)	-3.455 (4.898)
wd_rose_Cat4-Q2	-1.575*** (0.549)	0.539 (0.432)	0.911 (0.619)	9.247 (7.162)	4.239 (7.372)	-9.334** (4.696)
wd_rose_Cat5-Q2	-1.746*** (0.536)	0.297 (0.409)	1.014 (0.635)	20.063*** (7.498)	6.942 (7.526)	-12.856** (5.284)
wd_rose_Cat6-Q2	-1.599*** (0.542)	0.362 (0.419)	1.204* (0.635)	20.111*** (7.512)	7.605 (7.884)	-13.894** (6.104)
wd_rose_Cat3-Q1	0.012 (0.520)	-0.779 (0.694)	0.798* (0.420)	-3.922 (5.908)	-0.979 (4.973)	2.114 (7.162)
wd_rose_Cat4-Q1	0.506 (0.387)	-0.544 (0.729)	0.654* (0.369)	-8.577 (5.751)	-4.279 (4.675)	1.483 (7.051)
wd_rose_Cat5-Q1	0.725* (0.374)	-0.813 (0.748)	0.478 (0.358)	-4.343 (5.832)	1.131 (5.049)	0.690 (7.329)
wd_rose_Cat6-Q1	0.641* (0.384)	-0.599 (0.748)	0.353 (0.369)	-6.334 (6.533)	2.975 (5.478)	-0.835 (7.597)
<i>N</i>	9833	9833	9833	7683	7683	7683
<i>R</i> ² _first_a	0.26	0.30	0.34	0.18	0.21	0.24
<i>F</i> _first_a	32.94	33.08	19.38	10.97	29.68	34.43

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Peak afternoon is the average pollution during the peak in the afternoon per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. Each regression of the health outcomes has three equations in the first stage, one for each quarter of gestation. In this case, the three equations are the same for each set of four health outcomes presented in the second stage, for instance, the first four columns of Table 2.6 (for SO_2) share the same equations in the first stage, so we will only present once the results. The same holds for the other pollutants. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay). The interaction of the wind direction rose with wind speed not shown to save space.

The first stage estimations show not only the importance of using the wind variables lagged according to the quarter of gestation of the children born in the health center, but also, the complexity of the relationships between pollution with wind variables and their interactions in our setting. This can be noticed in the results of the first stage where some signs of wind direction on pollutants might change according to the quarter of gestation. This can be due to the fact that the instruments of wind direction and wind speed present time variation during the period analysed. As shown in Figure 2.11 and Figure 2.12 of Appendix 2D, wind direction and wind speed change during the different quarters and years. For the city, the wind blows mainly from the south, the south-east and the south west, with some years like 2008 where wind blew more from the south-west (see Figure 2.11, panel a) of Appendix 2D). The time variation of wind direction and wind speed variables is more evident in Figure 2.12, panel a) of Appendix 2D. The figure uses quarterly variation, much closer to the variation we exploit for the identification. As such, the data-set was constructed as a panel of health centers and running quarters according to the specific month in which the children were born (see subsection 2.4.1). Therefore, some children might be more or less affected by the levels of pollution, which in fact, depends on the delivery month. While a child born in December 2005 could have benefited from the lower levels of pollution and higher levels of wind speed and wind from the south and the south-west during the second quarter of gestation (July-August-September 2005), a child born in March 2006 might have being exposed to higher levels of pollution and less wind speed and wind from the south during the second quarter of gestation (October-November-December 2005).

Similar studies disaggregating the effect of pollution on newborns by quarter of gestation have also found a change in the sign of the effects of pollution and weather depending on the quarter (Lavaine and Neidell, 2017; Currie et al., 2009), and variability with respect to the pollutant. For this reason, we present in the second stage the cumulative effects (sum of the coefficients of the three quarters) of pollution and ENSO on health, to have a better approximation of total effects during pregnancy.

Tables 2.6, 2.7 and 2.8 show the main results of the second stage of the effects of ENSO and pollution on health outcomes. They are consistent with the estimations of the first stage presented

before. In the case of Table 2.6 and 2.7, we use average daily pollution to analyze impacts of SO_2 , PM_{10} and $PM_{2.5}$, while in Table 2.8, we employ pollutant concentrations during evening peak hours for CO , and NO_x . Reduced form specifications of health outcomes on weather and ENSO events are shown in Tables 2.14, 2.15 and 2.16 of Appendix 2D. Appendix 2F, also displays the results using exposure measures of pollution such as the mean of the moving average of the last 7 days and the mean of the daily maximum concentrations. Those results are qualitatively similar to those presented in Tables 2.6, 2.7 and 2.8, and constitute a robustness check for the estimations.

Table 2.6: IV SECOND STAGE: SO_2 (average) ON HEALTH

	SO2 (average)			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)
Niño-Q3	0.010*	0.014***	-0.081*	-31.838***
	(0.006)	(0.005)	(0.049)	(8.321)
Niño-Q2	-0.007	-0.012**	0.050	15.396*
	(0.007)	(0.005)	(0.053)	(8.353)
Niño-Q1	0.007	0.011*	-0.016	-14.291
	(0.006)	(0.006)	(0.044)	(8.929)
Niña-Q3	0.003	0.007	0.004	2.196
	(0.006)	(0.004)	(0.046)	(7.555)
Niña-Q2	-0.002	0.012***	0.004	-19.659***
	(0.006)	(0.004)	(0.045)	(7.370)
Niña-Q1	-0.013	-0.004	0.111*	8.564
	(0.008)	(0.006)	(0.059)	(10.535)
so2_ave-Q3	-0.000	-0.002	-0.000	-0.412
	(0.001)	(0.001)	(0.012)	(2.473)
so2_ave-Q2	-0.002	0.001	0.017	-2.306
	(0.002)	(0.002)	(0.013)	(3.138)
so2_ave-Q1	0.005**	0.003**	-0.038**	-6.386**
	(0.002)	(0.001)	(0.017)	(2.706)
Observations	11027	11027	11027	11027
NINO	0.01	0.01	-0.05	-30.73
P_NINO	0.08	0.01	0.29	0.00
NINA	-0.01	0.01	0.12	-8.90
P_NINA	0.27	0.07	0.12	0.48
POL	0.00	0.00	-0.02	-9.10
P_POL	0.25	0.07	0.21	0.00
r2_a	0.36	0.30	0.35	0.34

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the IPS, Weeks gest. is the average of weeks of gestation for all the children born in that IPS and Birth Weight is the average weight for all the children born in that IPS. Pollution is averaged per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the three coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

Regarding ENSO impacts, for specifications when pollutants SO_2 , PM_{10} and $PM_{2.5}$ are analyzed, El Niño and La Niña episodes reduce birth weight when exposure takes place in the last

quarter (Q3-gestation), when the mother is close to give birth. The effect ranges from 19 to 45 grams relative to normal conditions. These results manifest that climatic shocks via ENSO are not negligible. Cumulative impacts over the entire gestational length tend to be statistically significant different from zero for El Niño but not significant for La Niña (rows *NINO*, *NINA*, *P_NINO* and *P_NINA* of Table 2.6 and Table 2.7).

When it comes to pollutant effects, our findings indicate that exposure *in utero* to average SO_2 concentrations in the first and third quarter of gestation influences health outcomes at birth. An increase of 10 *ppb* in SO_2 in the first quarter of gestation raises the probability of premature birth by 5 percentage points, while 1 *ppb* extra of SO_2 may cause a loss of 6.4 grams in weight at birth. In a similar study, Lavaine and Neidell (2017) find that a refinery strike in France decreased the SO_2 emissions, which could have increased the birth weight of infants living close to the refineries, particularly during the first and the third quarter of gestation. In their article, a decrease of 1 $\mu g/m^3$ of SO_2 ($2.62 \mu g/m^3 = 1 \text{ ppb}$) for one month, increases the birth weight in the third quarter by 6.62 grams. Qualitatively, the results are similar to our findings, but larger.

In the case of particles, $PM_{2.5}$ exposure during the third quarter of gestation also increases the probability of being premature. The effect of an increase of 10 $\mu g/m^3$ of $PM_{2.5}$ is almost equivalent to a 6 percentage points increase in the likelihood of being premature. PM_{10} appears to decrease birth weight when exposure occurs in the second quarter of gestation. For the low weight outcome variable the sign of the PM_{10} effects differs across quarters, but when we sum the coefficients of the three quarters, the cumulative impact is not significant and statistically equal to zero (rows *POL* and *P_POL* of Table 2.7) (see Currie et al. (2009) and Lavaine and Neidell (2017) for references about the differences of the effect of air pollution on health by quarter of gestation). In fact, among this set of specifications, SO_2 is the only pollutant that has a negative cumulative effect during total gestational length on birth weight at conventional significance levels. Our results show a negative impact of pollution on low birth weight and premature birth. Such effects could be long-lasting, as Knittel et al. (2016) show that premature and low birth weight infants could be at higher risk of mortality in later stages of their development, when exposed to air pollution.

Although there is no statistical evidence that CO concentrations during evening peak hours

Table 2.7: IV SECOND STAGE: PM_{10} - $PM_{2.5}$ (average) ON HEALTH

	PM10 (average)				PM2.5 (average)			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)	Premature (5)	Low Weight (6)	Weeks Gest. (7)	Birth Weight (8)
Niño-Q3	0.005 (0.006)	0.003 (0.005)	-0.040 (0.042)	-20.068** (7.977)	0.005 (0.010)	0.000 (0.006)	-0.028 (0.078)	-10.391 (11.916)
Niño-Q2	0.002 (0.006)	-0.002 (0.005)	-0.011 (0.046)	-0.264 (8.199)	0.004 (0.010)	0.003 (0.007)	-0.025 (0.071)	-7.072 (9.997)
Niño-Q1	0.007 (0.006)	0.011* (0.006)	-0.017 (0.046)	-12.449 (9.983)	0.014 (0.010)	0.016** (0.006)	-0.065 (0.070)	-20.121* (11.107)
Niña-Q3	0.001 (0.006)	-0.000 (0.005)	0.017 (0.042)	7.275 (9.285)	-0.034** (0.016)	-0.020** (0.009)	0.272** (0.122)	19.773 (17.348)
Niña-Q2	-0.000 (0.006)	0.008 (0.006)	-0.002 (0.045)	-19.390* (10.742)	0.039*** (0.012)	0.033*** (0.008)	-0.260*** (0.094)	-44.379*** (11.263)
Niña-Q1	-0.003 (0.007)	-0.000 (0.006)	0.044 (0.057)	-3.750 (10.756)	-0.001 (0.012)	0.002 (0.009)	0.029 (0.094)	-0.052 (13.184)
pm10_ave-Q3	0.000 (0.001)	-0.001*** (0.001)	-0.001 (0.005)	1.195 (0.804)				
pm10_ave-Q2	0.001 (0.001)	0.001* (0.001)	-0.003 (0.005)	-1.871* (0.971)				
pm10_ave-Q1	0.000 (0.001)	0.001 (0.001)	-0.002 (0.006)	-0.915 (1.235)				
pm25_ave-Q3					0.006*** (0.002)	0.002 (0.001)	-0.039** (0.015)	-1.632 (1.992)
pm25_ave-Q2					-0.002 (0.002)	0.002 (0.001)	0.011 (0.014)	-2.172 (2.205)
pm25_ave-Q1					0.001 (0.001)	0.001 (0.001)	-0.004 (0.008)	-0.421 (1.307)
Observations	11027	11027	11027	11027	4413	4413	4413	4413
NINO	0.01	0.01	-0.07	-32.78	0.02	0.02	-0.12	-37.58
P_NINO	0.05	0.07	0.20	0.01	0.13	0.04	0.31	0.02
NINA	-0.00	0.01	0.06	-15.87	0.00	0.02	0.04	-24.66
P_NINA	0.83	0.38	0.43	0.22	0.86	0.23	0.79	0.21
POL	0.00	0.00	-0.01	-1.59	0.00	0.00	-0.03	-4.23
P_POL	0.25	0.55	0.40	0.27	0.32	0.09	0.29	0.29
r2_a	0.36	0.30	0.35	0.34	0.35	0.32	0.35	0.36

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the IPS, Weeks gest. is the average of weeks of gestation for all the children born in that IPS and Birth Weight is the average weight for all the children born in that IPS. Pollution is averaged per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the three coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

affect birth weight, *CO* exposure seems to increase the probability of having low birth weight, being premature and decrease the weeks of gestation. This impact particularly arises if the exposure occurs in the third quarter of gestation. 1 *ppm* increase in *CO* raises the probability of low weight at birth by 1.5 percentage points. This is an increase of 13.6% of low birth weight children delivered on average (from a base of 11 percentage points of low birth weight in the IPS-health centers). The results are high but in line with effects of previous studies. For example, Currie et al. (2009) find that one unit change in *CO* would lead to an increase in low birth weight of 0.0083 (from a base of 0.106) for the second quarter of gestation, an 8% increase of the incidence of low birth weight.

Aggregate coefficients (rows *POL* and *P_POL* of Table 2.8) of *CO* impacts on low birth weight are statistically different from zero at the 5% significance level. In the case of *NO_X*, the results indicate that this pollutant does not affect birth weight, although the coefficient is estimated with low precision.

In the set of specifications for *CO* and *NO_X*, the separate effects of ENSO show that El Niño and La Niña episodes increase the likelihood of having low weight at birth, and reduce birth weight. For El Niño events, exposure in the first and third quarter are relevant, whereas for La Niña episodes only exposure in the second quarter seems to matter. Quarter impacts of ENSO may decrease a child's birth weight from 18 to 26 grams compared to normal conditions. Cumulative ENSO effects across gestational quarters are more often statistically different from zero at the 5% of significance level for El Niño, but less for La Niña events (rows *NINO*, *NINA*, *P_NINO* and *P_NINA* of Table 2.8). Although some studies have already shown an effect of extreme temperature on health, we consider the impact of climatic shocks via ENSO events on health as a new channel that has not been explored in the literature yet. For example, Deschênes et al. (2009) show that exposure to extreme temperatures (particularly hot) during pregnancy leads to lower birth weight, especially during the second and third quarter of gestation (see also Barreca et al. (2015) for the effects on child mortality). For the case of Colombia, El Niño brings higher temperatures, while La Niña brings more rain. As local weather factors are already controlled for in our specifications, the ENSO events capture a more global distortion such as extreme climatic shocks.

In addition, the reduced-form estimations in Appendix 2E show that ENSO shocks, weather factors and the instruments used, such as wind direction, explain around 30 to 40% (using the adjusted R^2) of the variation in the health outcomes used. As our estimations use a large set of instruments with their lags, we use the R^2 and adjusted R^2 as a measure of the total significance of the exogenous variables on health outcomes. The reduced-form results for *PM_{2.5}* (Table 2.15 in Appendix 2E) give larger effects for the different instruments used and the effects of the instruments and exogenous variables are more often statistically significant at 10% level for birth weight. In many of the reduced-form estimations, there is a significant and negative effect of El Niño (Q3-gestation) on the birth weight of children, at the 5% significance level.

Table 2.8: IV SECOND STAGE: $CO-NO_X$ (peak afternoon) ON HEALTH

	CO (peak afternoon)				NOX (peak afternoon)			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)	Premature (5)	Low Weight (6)	Weeks Gest. (7)	Birth Weight (8)
Niño-Q3	0.003 (0.006)	0.011** (0.006)	-0.025 (0.042)	-22.746*** (8.249)	-0.000 (0.006)	0.008 (0.006)	-0.009 (0.048)	-25.653*** (9.667)
Niño-Q2	-0.007 (0.008)	-0.002 (0.005)	0.045 (0.062)	10.171 (9.183)	-0.005 (0.007)	-0.008 (0.005)	0.042 (0.051)	8.255 (9.045)
Niño-Q1	0.006 (0.007)	0.013** (0.005)	-0.013 (0.049)	-14.216 (9.109)	0.007 (0.006)	0.019*** (0.005)	-0.016 (0.048)	-22.639** (9.421)
Niña-Q3	0.003 (0.006)	0.007* (0.004)	0.001 (0.042)	0.861 (8.362)	-0.002 (0.006)	-0.001 (0.005)	0.049 (0.048)	15.045 (9.826)
Niña-Q2	-0.004 (0.005)	0.012*** (0.004)	0.020 (0.038)	-19.123*** (8.218)	-0.009 (0.007)	0.011** (0.005)	0.052 (0.053)	-17.768* (9.129)
Niña-Q1	-0.003 (0.007)	0.002 (0.006)	0.043 (0.054)	5.939 (9.782)	-0.009 (0.007)	-0.008 (0.007)	0.089* (0.054)	11.571 (10.476)
co_peaka-Q3	0.014* (0.008)	0.015** (0.007)	-0.103* (0.060)	2.462 (12.977)				
co_peaka-Q2	-0.004 (0.006)	0.010 (0.006)	0.010 (0.047)	9.545 (10.793)				
co_peaka-Q1	-0.001 (0.005)	0.003 (0.004)	-0.002 (0.036)	6.531 (8.242)				
nox_peaka-Q3					0.000 (0.000)	0.000 (0.000)	0.000 (0.003)	-0.293 (0.559)
nox_peaka-Q2					-0.000 (0.000)	0.000 (0.000)	0.000 (0.002)	-0.753 (0.505)
nox_peaka-Q1					-0.000 (0.000)	0.000 (0.000)	0.002 (0.002)	-0.620 (0.429)
Observations	9833	9833	9833	9833	7683	7683	7683	7683
NINO	0.00	0.02	0.01	-26.79	0.00	0.02	0.02	-40.04
P_NINO	0.81	0.00	0.93	0.02	0.85	0.01	0.79	0.01
NINA	-0.00	0.02	0.06	-12.32	-0.02	0.00	0.19	8.85
P_NINA	0.67	0.00	0.38	0.31	0.07	0.88	0.02	0.53
POL	0.01	0.03	-0.10	18.54	-0.00	0.00	0.00	-1.67
P_POL	0.51	0.02	0.38	0.46	0.64	0.44	0.66	0.19
r2_a	0.37	0.30	0.35	0.34	0.38	0.33	0.36	0.37

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the IPS, Weeks gest. is the average of weeks of gestation for all the children born in that IPS and Birth Weight is the average weight for all the children born in that IPS. Peak afternoon is the average pollution during the peak in the afternoon per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the three coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

Overall, the compelling message of our results is that across all specifications ENSO affects birth weight and the probability of low birth weight after separating pollution and classical local weather impacts. In addition, the semi-elasticity computations using column 4) of Table 2.6 shows that being exposed to a Niño event during gestation decreases birth weight by 1.02%, while an increase of one unit of *ppb* of SO_2 , on average, decreases birth weight in the health centers by 0.3% (being exposed to a Niña event during gestation decreases birth weight by 0.3%).¹⁸ Similar semi-elasticities

¹⁸The semi-elasticities are calculated for the cumulative effect (sum of the effects) for the three quarters of gestation.

calculated using column 8) of Table 2.7 show that being exposed to a Niño event during gestation decreases birth weight by 1.26%, while an increase in one $\mu\text{g}/\text{m}^3$ of $PM_{2.5}$ (average) decreases birth weight in the health centers by 0.14% (being exposed to a Niña event during gestation decreases birth weight by 0.82%). These results lead us to conclude that the ENSO effects on birth weight are several times larger than the impacts coming from pollution. It is very relevant from the policy point of view, because regardless of the measure of pollution that we employ, the amount of the impacts exhibited by climatic shocks via ENSO events dominate. If ENSO conditions may be accurately predicted by meteorologists, our article suggests that policymakers might anticipate and consider in their planning strategies to avoid health impacts caused by extreme climatic variability.

2.5.3 Spatial variation and residential sorting

Some potential threats to identification were discussed in subsection 2.4.2 and we investigate them here in more detail. Table 2.20 of Appendix 2G calculates the Moran Index of spatial correlation for the main weather variables and pollutants, year by year and for a spatial distance of 2km (the distance used to construct the buffers around each health center). The table shows that pollutants such as SO_2 , for instance, have a positive spatial correlation but the correlations tend to change over time, with some values much smaller than in other years. This pattern holds for the different pollutants analyzed. Similar results can be seen for rainfall, temperature and wind speed. There is not only cross-section but also yearly variation for the different locations around the city. There may be a concern if the variation was driven mainly from the cross-section, and some areas tend to have more pollution than others. However, we also observe variation along time for the different locations of the IPS health center.

Another issue was that changes in air quality may result in residential sorting, for instance, if wealthier or more educated parents decide to move to cleaner neighborhoods, generating a bias in our results. To this aim, Hanna and Oliva (2015) analyze the effects of pollution on health by using the closure of a large refinery in Mexico city as a single shock, and check the problem of residential sorting by testing if migration and other demographic characteristics change in the areas analyzed before and after the shock (the instrument). We follow this approach in three steps, first by checking if our instruments affect the demographic characteristics, second, using another data-set to verify if

some socioeconomic characteristics vary by locality during the years, and third, by restricting the sample to the South-West of the city.

For the first step, Table 2.21 of Appendix 2G shows the coefficient estimates of wind direction and wind speed (our instrumental variables for pollution) and NIÑO-NIÑA on some key socioeconomic characteristics of the running-quarter analysis. All variables are averaged at health center by running-quarter. Many variables such as parents' age, marital status, the percentage of female births in the health center and the mother's education are not affected by our instruments. In addition, the effect on the type of social security of the mother and father's education is small. Although this confirms that those factors should be controlled for in the identification, it is less likely that individuals decide to locate based on whether the wind blows pollution to some areas more than others.

For the second step, the economic factors tend to be more important in location decisions than the fact that the wind blows pollution from the south, as we observe in part b) of Table 2.22. Part a) of the table shows the movement of individuals between localities in the city, which tends to be stable for the years of the Multipurpose Survey for Bogotá. The last two columns of part a) of the table shows the test comparing the proportion of migration in each locality between 2011 and 2014. We fail to reject the null hypothesis that the difference in the percentage of migration between 2011 and 2014 was equal to zero (except for the locality of Engativa). Part b) of Table 2.22 shows the different reasons for moving between localities in the city. It shows that economic, family factors and education are the main reasons why individuals decide to move inside the city and health factors account only for around 2% of them. Part b) of the table shows the test comparing if the proportion of individuals moving for economic reasons is larger than the proportion of those moving for health reasons. We fail to reject the null hypothesis that the economic reasons for moving is larger than the health reasons for moving at the conventional levels, and by the different years.

Additionally, Table 2.23 presents how some socioeconomic characteristics vary for the years of the survey and by locality. We do not see large changes in the socioeconomic variables between the different years. Migration is very low by locality and the percentage of women by locality is rather stable. Although there is an increasing trend in the years of education and age, this is a general as-

pect similar along the localities. By including year times locality fixed effects in our estimations, the general trend should be controlled for already. The results of part b) in Table 2.23 show that there is no statistically significant differences in the average of the socioeconomic characteristics along the years. At least for the period we have data on, we do not have evidence to assert the presence of a residential sorting problem in terms of these variables, as they do not change drastically during the years of the sample.

With respect to the third step, residential sorting may occur from the south to the north in the city. It is common knowledge that individuals who have the economic possibilities tend to locate more towards the north and less to the south, in order to have access to better economic conditions, live in more secure areas with better amenities and probably, less pollution. For this reason, Appendix 2H and the next section show the results restricting the sample for the South-West of the city, which tends to be poorer and more polluted. The results are qualitatively similar to the ones found for the entire city with some coefficients slightly higher. This shows that the potential residential sorting from the south to the north might be less important here. If it exists, we should expect very different coefficients in the estimations using only the South-West of the city.

Another issue of residential sorting could arise if parents caring for their children may prefer some less polluted areas and also pay more attention to the mother's health during pregnancy. That is to say, a self-selection of less polluted places of residence positively correlated with the quality of the attention paid to health during pregnancy. This should be less a concern here for three reasons: 1) wind direction is exogenous, it should control for this kind of sorting problem; also, wind direction tends to change during the years, there is not only cross-section but also time variation of wind, which helps in the identification (see Appendix 2D); 2) Panel b) of Table 2.22 in Appendix 2G shows that individuals in fact move inside the city mainly for economic reasons, then for family or education reasons, while health aspect accounts for just 2% of those permanent movements; 3) even if some mothers pay attention to the levels of pollution during pregnancy, it would not be economically feasible for many mothers to relocate to areas of less pollution only during the specific time of pregnancy in an urban city like Bogotá; this re-settlement would be very costly for just the months of the pregnancy.

2.5.4 Restricting the sample to the South-West of the city, aggregating by localities and using a bigger buffer

Appendix 2H shows a robustness check by restricting the sample only to the localities of Bosa, Ciudad Bolivar, Engativa, Fontibon, Kennedy, Puente Aranda and Tunjuelito, located in the South-West of the city. These localities tend to be poorer and to present higher levels of pollution than other areas of the city. Also, as it is shown in the map of wind direction and wind speed in Appendix 2D, the wind blows towards the south-west of the city. The second stage tables (Table 2.25 and Table 2.26) are qualitatively similar to the ones presented in subsection 2.5.2, and give evidence that the results are not driven only by localities with higher pollution and lower income. However, some of the coefficients for El Niño as well as for PM_{10} and $PM_{2.5}$ are slightly larger than before. For the south-west of the city, we find that the cumulative effect (sum of the three coefficients) is statistically significant at the 10% level less often than with the baseline sample for PM_{10} and $PM_{2.5}$. In line with the baseline sample, the first-stage estimates of Appendix 2H show the presence of statistical correlation between the instruments of wind direction with pollutant concentrations (Table 2.27 and Table 2.28).

As an additional robustness check, Appendix 2I presents the results aggregating the sample by locality. Although we lose precision in assigning pollution and weather variables to a whole locality, the results follow a similar pattern for some of the pollutants presented in subsection 2.5.2. Table 2.29, Table 2.30 and Table 2.31 give the results for the second stage estimations. Regarding ENSO events, the quarterly coefficients and the cumulative effects follow a similar pattern as in the baseline estimation for CO , PM_{10} and $PM_{2.5}$ in general and for NO_X , but mainly for birth weight. The results are less intuitive in the estimation involving SO_2 , which could be due to the lack of precision while aggregating the sample at the locality level. The effects of PM_{10} and CO by quarter are qualitatively similar to those obtained when aggregating by IPS health centers. Interestingly, the quarterly coefficients are less often statistically significant at conventional levels in comparison to the main baseline results. However, the cumulative effect is very similar for CO , PM_{10} and $PM_{2.5}$ with more statistically significant results and in the expected direction for NO_X (not for SO_2). The first-stage estimates of Appendix 2I confirm the statistical correlation between the instruments of

wind direction with pollutant concentrations (Table 2.32 and Table 2.33).

We carry out an additional robustness check by changing the size of the buffer around each health center to investigate the sensitivity of the results. As the spatial correlation between the place of residence and the health center location is certainly positive (but not necessarily high) and we do not know where the mother actually live, we prefer to use a larger buffer of 4km. By increasing the buffer and giving a decreasing weight to the farthest measurement stations, we increase the likelihood that the measure better characterizes the air pollution level in the area where the mother is supposed to live / work / sleep during pregnancy. The results of ENSO and pollution on health outcomes using the buffer of 4km (see Appendix 2J) are qualitatively and quantitatively very similar to the baseline results using a buffer of 2km (see Table 2.6, Table 2.7 and Table 2.7), with a few coefficients of the ENSO effect on health outcomes with the 4km buffer being slightly smaller. In this regard, the results are very robust with respect to changes in the buffer size and the concern that air pollution might not be correctly assigned to where the mother lives seems to be less of an issue here.

Another potential concern is whether the pollution monitoring stations that underlie the computation of the weighted average concentration around each health center are of the same type (roadside, urban, industrial, near city background), and could affect the results. Although using health center fixed effects should control for this, the robustness check changing the 2km buffer when assigning pollution data to the health centers helps also to deal with the differences in the type of monitoring station. As mentioned previously, the results are very robust to the use of a bigger buffer of 4km and the differences in the type of monitoring station should be of less concern in the estimations.

2.5.5 Direction of the OLS bias

Appendix 2K has the OLS estimations of the effects of ENSO and pollution on the main variables of health without the instruments. This exercise allows us to understand better the direction of the bias, by comparing the IV results of the 2nd stage with this table. Across the majority of the estimations, the negative effect of ENSO and pollution on health is underestimated. For example, columns 3) and 4) of Table 2.6 and Table 2.7 show a higher negative effect of ENSO on weeks

of gestation and on birth weight, compared to the coefficients in Table 2.37, columns 3, 4, 7 and 8. With respect to pollution, the majority of the coefficients in Table 2.37 are not statistically significant, but once we adjust the estimations in the IV regressions, we find the negative effect of pollution on the health variables used. Not solving the endogeneity problem in the estimations could lead us to neglect the negative effect of the pollutants on health and underestimate the negative impact of ENSO.

2.5.6 Heterogeneous effects, additional health outcomes and air pollution index

We explore heterogeneity with respect to the educational level of the mother. In the estimations, the variables representing the mother's education variables are defined as the percentage of mothers delivering in the IPS-health center with primary education, secondary education or tertiary education. We use the interaction of the tertiary education variable with ENSO events as we are more interested in the effects of these shocks on human capital and consumption, particularly for mothers with higher education (see Appendix 2L). Stunningly, the negative effect of ENSO events on health outcomes at birth is driven by mothers without tertiary education (primary or secondary education), with a positive effect from ENSO for the mothers with tertiary education. This pattern is found across all the different estimations using the different pollutants (Table 2.38, Table 2.39 and Table 2.40). For instance, the effect of having an El Niño episode during the three quarters of gestation if the mother had primary or secondary education is a reduction in the birth weight of the child of 19.42 grams, while a La Niña episode entails a reduction in the birth weight of the child of 30.81 grams (cumulative impact, rows *NINO-NINA* of column 4, Table 2.38 in Appendix 2L). However, the cumulative effect of having an El Niño episode for a child with a mother with tertiary education is an increase in the birth weight of 32.9 grams, while having a La Niña episode for a child with a mother with tertiary education is an increase in the birth weight of 106.05 grams (cumulative impact, rows *NINO_EDU-NINA_EDU* of column 4), Table 2.38 in Appendix 2L). It is also compelling to find that the effects of pollution on health outcomes for the estimations interacting ENSO events with tertiary education are robust and similar to the ones found in Table 2.6, Table 2.7 and Table 2.8.

In Appendix 2M, we explore if ENSO events and pollution can also affect three additional health

outcomes. First, the average Apgar score at one minute of birth (index 1-10) of the children born in the health center, which can help to assess the health of the newborn at birth. Second, we combine low weight at birth and prematurity in a single indicator to better characterize the most adverse health effects. And third, we construct the sex-ratio as the percentage of boys over the percentage of girls for the children born in that health center (see [Lichtenfels et al. \(2007\)](#) for evidence of the change of male to female ratio due to pollution). With respect to the second additional outcome, it is possible that low birth weight is observed at term, or that premature newborns have a reasonable weight, which could weaken the detection of adverse health effects from air pollution exposure. Despite that, a newborn with low weight and prematurity is more likely to reflect the effects of exposure to harmful conditions during pregnancy. Economically speaking, this would also involve higher medical costs, educational costs during childhood or opportunity losses in future earnings (see for example [Almond and Currie \(2011a\)](#) and [Almond et al. \(2005\)](#)).

We do not find statistically significant effects of ENSO on the Apgar score in [Table 2.41](#) and [Table 2.42](#) in [Appendix 2M](#), with some mixed results from pollutants on the score (positive effect from quarter two and quarter three from SO_2 and negative impact from CO in the first quarter). Nonetheless, the separate effects of ENSO show that El Niño episodes (but not La Niña) increase the likelihood of having low weight and being premature at birth. Using the aggregate coefficients (rows $NINO$ and P_NINO of [Table 2.41](#) and [Table 2.42](#)) the El Niño impact on having low weight and being premature at birth are statistically different from zero at the 5% significance level in the estimations using SO_2 , CO and NO_X and at 10% significance level for PM_{10} . With respect to pollutants, only SO_2 and CO have a statistically significant effect at reasonable levels on the low birth weight and premature variable. In terms of magnitude, being exposed to El Niño during the three quarters of gestation increases the percentage of children with low weight and premature by 2 percentage points (baseline of 8% in [Table 2.1](#)). An increase of 1 *ppm* of CO also rises the percentage of children with low weight and prematurity by 2 percentage points during the three quarters.

With respect to the last outcome variable, being exposed to El Niño decreases the sex ratio (on average less boys are born), particularly during the second quarter of gestation (columns 3 and

6 of Table 2.41 and Table 2.42 in Appendix 2M), but the effect is the opposite during the first and third quarter of gestation. However, the net cumulative effect during the three quarters of gestation is null, and only statistically significant for El Niño at the 10% of significance level in one specification. There is no statistically significant effects from the cumulative effect from La Niña at reasonable levels across all the estimations and during the different quarters of gestation. Regarding the pollutant variables, only SO_2 has a statistically significant effect at the 5% significance level on the sex ratio variable.

Finally, we construct a quarterly index of pollution which gathers the level of exposure of the five main pollutants. Following Arceo et al. (2015), we apply principal component analysis (PCA) on the standardized version of the pollutants of the main specifications (SO_2 , PM_{10} , $PM_{2.5}$ (average), CO , NO_X (peak afternoon)). The estimations use the first principal component, which has an eigenvalue larger than one. As $PM_{2.5}$ has many missing values, PCA version A includes $PM_{2.5}$ (explained variation of 48.9%), while PCA version B does not include $PM_{2.5}$ (explained variation of 44.1%). Both indices are scaled to be between zero and one for an easier interpretation. Table 2.43 in Appendix 2N shows that PCA version A uses a very small sample of observations, which might not be large enough to capture the time variation of the effects of ENSO events. For a better interpretation, we use PCA version B of Table 2.43 in Appendix 2N. With this composite measure of pollution, being exposed to El Niño increases the percentage of children with low weight at birth in the IPS by 3 percentage points across the three quarters of gestation (cumulative effect), while an increase of 0.1 units in the index of pollution increases the percentage of children with low weight at birth in the IPS by 2 percentage points across the three quarters of gestation (cumulative effect). The effect of El Niño on weight at birth follows the same pattern, so being exposed to the phenomenon decreases the weight at birth by 40.1 grams (cumulative effect). The results for the cumulative effect on birth weight are not statistically significant at reasonable levels for the index of pollution, but qualitatively similar to the ones found using the pollutants separately.

2.6 Conclusions

This article contributes to the literature of air pollution and health by adding an additional channel, the effect of ENSO events (El Niño and La Niña) on health. Although there is a vast literature of the effects of air pollution on health (Graff Zivin and Neidell, 2013; Lavaine and Neidell, 2017; Arceo et al., 2015), our research, unlike other studies, investigates at the same time the effects of ENSO events, air pollution and local weather on health. Our argument is that ENSO not only manifests itself as an extreme climatic shock that influences weather, but that it also has an impact on agriculture inducing changes in food markets and household income and consumption, which influence health. Therefore, the impact of pollution on health can be seen as separate effects from the other shocks mediated through ENSO events. Although the effects of air pollution on health have already been studied in the literature using similar instrumental variables like the ones used here (wind direction), the channel of ENSO events on health has not been explored yet. Our approach relies on studies like Hsiang et al. (2011) and Dingel et al. (2020), where ENSO events are considered in a wide perspective as global effects, affecting not only weather factors, but also creating disturbances on economic variables and, in that manner, pregnant mothers and infant health. Our article links the literature on air pollution and health and the growing number of studies that assesses the impact of extreme temperatures on health (see Barreca et al. (2015), Deschênes et al. (2009), Deschênes and Greenstone (2011)).

We conduct the analysis using data for Bogotá from 1998 to 2015, a city with high levels of pollution. Our approach consists of creating a running-quarter panel of health centers, where the quarter corresponds to the quarter of gestation for the infants born at a specific time in the health center. The identification strategy relies on a two stage approach, where the first stage instruments the pollution equation by wind direction; the second stage establishes the relationship of pollution and ENSO on health, controlling for weather factors, among other controls. Importantly, the system has the advantage of capturing the lagged effect on health of the exposure to pollution and ENSO during the three quarters of gestation.

Regarding ENSO impacts, when estimations involve SO_2 , PM_{10} and $PM_{2.5}$, El Niño and La

Niña events reduce birth weight when exposure happens in the last quarter of gestation, the period in which the mother is close to deliver. In the estimations with CO and NO_X , the separate effects of ENSO show that El Niño (in the third and first quarter) and La Niña (in the second quarter) increase the likelihood of having low weight at birth, and reduce the birth weight. Independently of the pollutant used, the cumulative effect of ENSO across gestational quarters is more often statistically different from zero at conventional significance levels for El Niño, but less for La Niña episodes.

For the effects of pollution on health, we find that exposure *in utero* to average SO_2 concentrations in the first and third quarter of gestation influences several health outcomes at birth, while $PM_{2.5}$ exposure during the third quarter of gestation increases the probability of being premature. In addition, PM_{10} appears to decrease birth weight when exposure occurs in the second quarter of gestation, while CO exposure (in the second and particularly, in the third quarter) increases the probability of having low birth weight. When analysing the cumulative effect during the total gestational length, SO_2 and CO are the only pollutants that have a negative impact on birth weight at conventional significance levels.

In a nutshell, across all specifications our results indicate that after including the channels of pollution and classical local weather, during some gestational quarters ENSO affects birth weight. Moreover, the size of this impact is much larger than the consequences of pollution exposure alone. Our article thus sheds light on the magnitude of the estimate of the overall effect of extreme climatic shocks on health at birth via ENSO. To the extent that accurate weather and ENSO forecasts are available, policymakers could design ex-ante measures to mitigate the impact of these shocks and implement policies to reduce the health effects induced by extreme climatic variability. As the *in utero* period is one of the most important stages for children's later development and a crucial component of human capital development (see [Almond and Currie \(2011a\)](#)), policy interventions should also be targeted during pregnancy.

Appendices

2A Appendix Summary Traffic Policies - Wind Direction

Table 2.9: Traffic Policies

No Car Days		
February 4, 2000	February 1, 2001	February 7, 2002
February 6, 2003	September 22, 2003	February 5, 2004
February 3, 2005	February 2, 2006	February 1, 2007
February 7, 2008	February 5, 2009	February 4, 2010
February 3, 2011	February 2, 2012	February 7, 2013
February 6, 2014	February 5, 2015	April 22, 2015

Transmilenio Phases		
	Starting in	Ending in
Phase I	December 4, 2000	September 26, 2003
Phase II	September 27, 2003	June 8, 2012
Phase III	June 9, 2012	Still in place

Peak and Plate Changes		
	Starting in	Ending in
Peak and Plate 1	August 18, 1998	February 5, 2009
Peak and Plate 2	February 6, 2009	July 2, 2012
Peak and Plate 3	July 3, 2012	Still in place

Table 2.10: Wind Direction (Categories)

Categories	Wind Direction	Degrees Range
1	North	337.5° - 22.5°
2	North-East	22.5° - 67.5°
3	East	67.5° - 112.5°
4	South-East	112.5° - 157.5°
5	South	157.5° - 202.5°
6	South-West	202.5° - 247.5°
7	West	247.5° - 292.5°
8	North-West	292.5° - 337.5°

2B Identification of the effects on Weather and Pollution Hourly Analysis

2B.1 ONI Index

As ENSO events bring more extreme weather events, the first equation to estimate is:

$$\begin{aligned} Weather_{i,h,t} = & \beta_0 + \beta_1 \times ONI_t + \gamma_i + \theta_i \times t \\ & + \nu_{month} + \varepsilon_{i,h,t} \text{ for } h = 1, 2, \dots, 24 \end{aligned} \quad (2)$$

where ONI is the continuous monthly index as described previously. γ_i and $\theta_i \times t$ are station-specific fixed effects and station trends to account for unobserved time-invariant differences and trends across i . Equation 2 also includes ν_{month} as fixed effects for the month of the year, which should control for seasonal effects on weather.

The coefficient of interest is β_1 .¹⁹ It measures the effect of a unit increase in the monthly ONI index on weather in time t for each specific hour $h = 1, 2, \dots, 24$. This allows to capture easily the effect of a unit increase in the ONI index on weather for the first hour, for the second hour, and so on. $\varepsilon_{i,h,t}$ are the standard errors, which are clustered at the i station level in the estimation. In a second step, we estimate the effects of ONI on pollutants:

$$\begin{aligned} Pollution_{i,h,t} = & \alpha_0 + \alpha_1 \times ONI_t + \Gamma_i + \omega_i \times t + \eta_{dow} + \mu_{month} \\ & + \alpha_2 \times WS + \alpha_3 \times WS^2 + \kappa_{wd} + \psi_{no-car} + \phi_{Transmilenio} \\ & + \delta_{peak-plate} + \epsilon_{i,h,t} \text{ for } h = 1, 2, \dots, 24 \end{aligned} \quad (3)$$

This relates the monthly effect of the ONI index on pollutants for station i and hourly time t . As previously, the effect is evaluated hour by hour so it creates a set of twenty-four regressions per pollutant. Equation 3 follows a similar structure as Equation 2, except that it includes η_{dow} as fixed effects of the day of the week. It is expected that traffic patterns change depending on the day of

¹⁹As an alternative to equation 2 that captures the ONI effect for each hour separately (24 regressions), we estimate a model of the ONI effect for all the 24 hours (by interacting the ONI index with each hour in a more parsimonious way), controlling for hourly fixed effects, and the same set of fixed effects as in equation 2. However, the article only presents the empirical strategy hour by hour, that we consider more conservative. For illustrative purposes, the results of this alternative model are shown in the Appendix 2B.5 and Appendix 2B.6 as MODEL B. MODEL J is the baseline presented here.

the week, particularly during Saturdays and Sundays.

Equation 3 also includes other weather factors such as wind speed (WS) in a quadratic way and a set of dummy variables of wind direction (8 categories, see Appendix 2A, and Appendix 2D for a detailed description). The inclusion of additional weather factors such as wind speed and wind direction have been well documented in the literature and not including them would lead to omitted variable bias (see Knittel et al. (2016) or Hanna and Oliva (2015)). Weather factors should have a relation with pollution, but it might be that during El Niño or La Niña events, the interaction changes and affects the levels of pollution. This could increase the levels of pollution or even reduce them, depending on the type of pollutant.

In terms of traffic policies, equation 3 includes a dummy for the no-car-day measure, a set of dummies for three different Transmilenio²⁰ Phases implemented in the city and a set of dummies for three changes in the peak-and-plate measures to reduce traffic congestion (See Appendix 2A). $\epsilon_{i,h,t}$ are the standard errors as usual. In the estimations, the errors are clustered at the monitoring station or i level.

2B.2 Separating the effects of NIÑO and NIÑA

To complement the analysis and understand better the effects of El Niño versus La Niña, we estimate a variant of equation 2 in which we replace the ONI index by dummies of El Niño and La Niña. This allows to compare the effect of both events with normal months (or months with no events). As before, the first equation to estimate is:

$$\begin{aligned} Weather_{i,h,t} = & \beta'_0 + \beta'_1 \times NI\tilde{N}O_t + \beta'_2 \times NI\tilde{N}A_t + \gamma'_i + \theta'_i \times t \\ & + \nu'_{month} + \epsilon'_{i,h,t} \quad for \quad h = 1, 2, \dots, 24 \end{aligned} \quad (4)$$

The coefficients of interest are β_1 and β_2 .²¹ Here, NIÑO and NIÑA are discrete dummies to

²⁰Transmilenio is the massive transportation system currently working in the city and it was implemented in three phases.

²¹Like for equation 2, equation 4 is also estimated for an alternative model of the ONI effect for all the 24 hours (by interacting the NIÑO and NIÑA dummies with each hour in a more parsimonious way), controlling for hourly fixed effects, the NIÑO and NIÑA dummies, and the same set of fixed effects as in equation 4. This section only presents the empirical strategy of the 24 hour by hour regressions, that we consider more conservative. For illustrative purposes,

account for periods when the events happened. As in the case of ONI (equation 2), γ'_i and $\theta'_i \times t$ are station-specific fixed effects and station trends to account for unobserved time-invariant differences and trends across i stations and ν'_{month} are fixed effects for the month of the year, to control for seasonal effects on weather. Equation 4 estimates the monthly effect of NIÑO and NIÑA (with respect to normal months) on weather in time t for each specific hour $h = 1, 2, \dots, 24$. $\epsilon'_{i,h,t}$ are the standard errors, which are clustered at the i or station level in the estimation. The second step will be to estimate the effects of NIÑO and NIÑA on pollutants:

$$\begin{aligned}
Pollution_{i,h,t} = & \alpha'_0 + \alpha'_1 \times NI\tilde{N}O_t + \alpha'_2 \times NI\tilde{N}A_t + \Gamma'_i + \omega'_i \times t + \eta'_{dow} + \mu'_{month} \\
& + \alpha'_3 \times WS + \alpha'_4 \times WS^2 + \kappa'_{wd} + \psi'_{no-car} + \phi'_{Transmilenio} \\
& + \delta'_{peak-plate} + \epsilon'_{i,h,t} \quad for \quad h = 1, 2, \dots, 24
\end{aligned} \tag{5}$$

This relates the monthly effect of NIÑO and NIÑA (with respect to normal months) on pollutants for station i and hourly time t . As previously, the effect is evaluated hour by hour so it creates a set of twenty-four regressions per pollutant. Equation 5 includes identical controls as those used in equation 3. $\epsilon'_{i,h,t}$ are the standard errors, which are clustered at the i or station level in the estimations.

An alternative model was estimated including ONI events (and NIÑO-NIÑA) in a contemporaneous and in a lagged way (a month back). As the effects do not change a lot, they are not presented here.²² In the case of NIÑO-NIÑA, a potential explanation is that by definition, the events are determined when the ONI index reaches five consecutive months above or below 0.5, hence, it already captures the impact from previous periods and there is no need to include lags for the ONI index or lags for NIÑO-NIÑA in the respective equations.

2B.3 Results of ONI-NIÑO-NIÑA on weather - Hourly Analysis

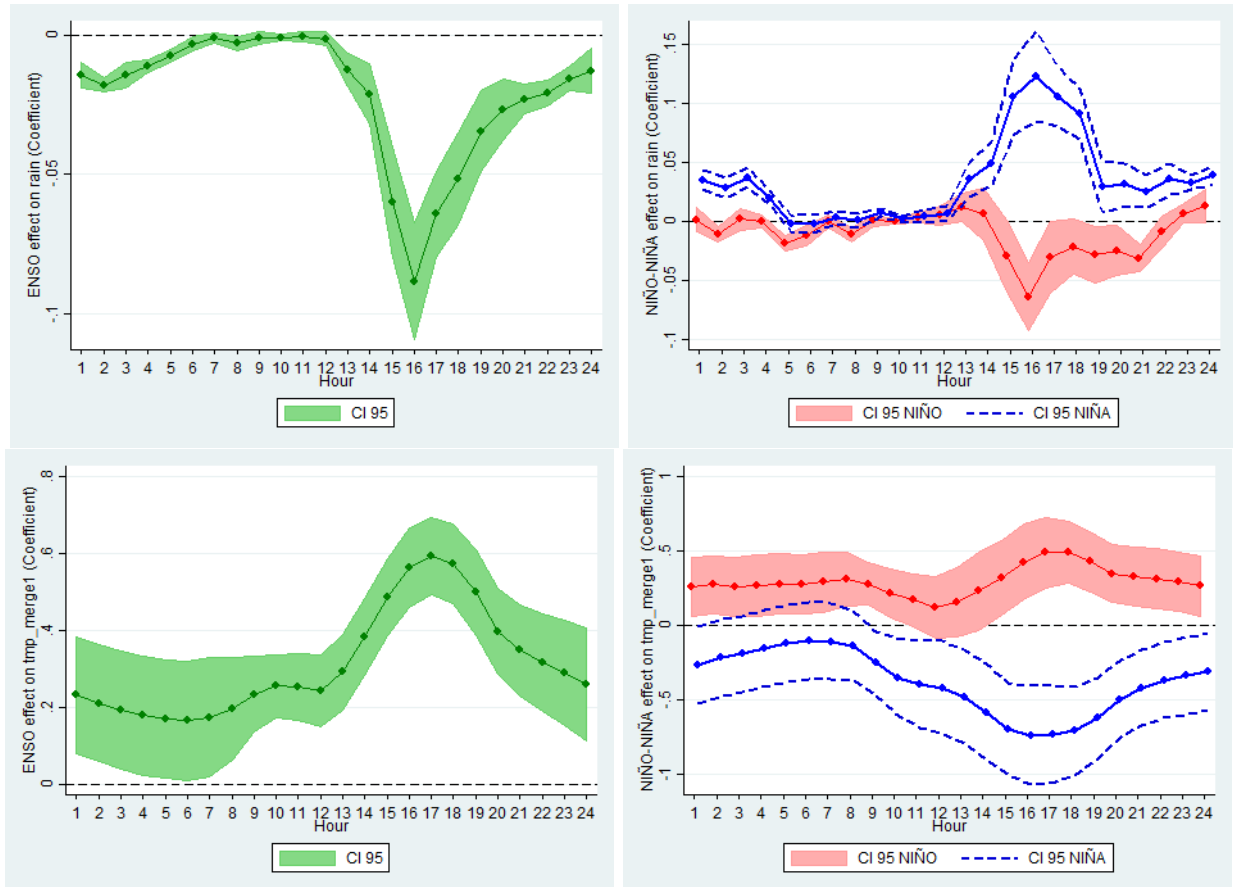
Figure 2.7 shows the ONI effect and NIÑO-NIÑA on rain and temperature (tmp), as described in the equations 2 and 4. The effects are estimated hour by hour and graphed with 95% confidence

the results of this alternative model are shown in subsection 2B.5 and subsection 2B.6 as MODEL B. MODEL J is the baseline presented here.

²²Results available upon request.

intervals. The ONI index tends to decrease rainfall, particularly after 1pm, reaching a minimum at 4pm, while it increases the temperature during all the 24 hours, reaching a maximum around 5pm.

Figure 2.7: ONI-NIÑO-NIÑA on weather-24hours schedule



Note: based on the Air Quality Monitoring Network Data for Bogotá (RMCAB)
 ENSO(ONI) in green and NIÑO-NIÑA in red and blue respectively

As discussed in the introduction, El Niño is associated with less rain (in comparison with normal months), while La Niña comes with more (in comparison with normal months). In terms of temperature, it increases during the majority of the hours of El Niño, while it decreases during La Niña. Importantly, rain and temperature tend to be more affected by La Niña than by El Niño. Rain and temperature also reach the maximum and minimum around 4pm-5pm, time during which individuals are moving in the city. Also, temperature increases in 0.5 degrees Celsius during El Niño, while it decreases 0.7 degrees Celsius during La Niña. This shows the differences in magnitude that

both ENSO events have in the city.

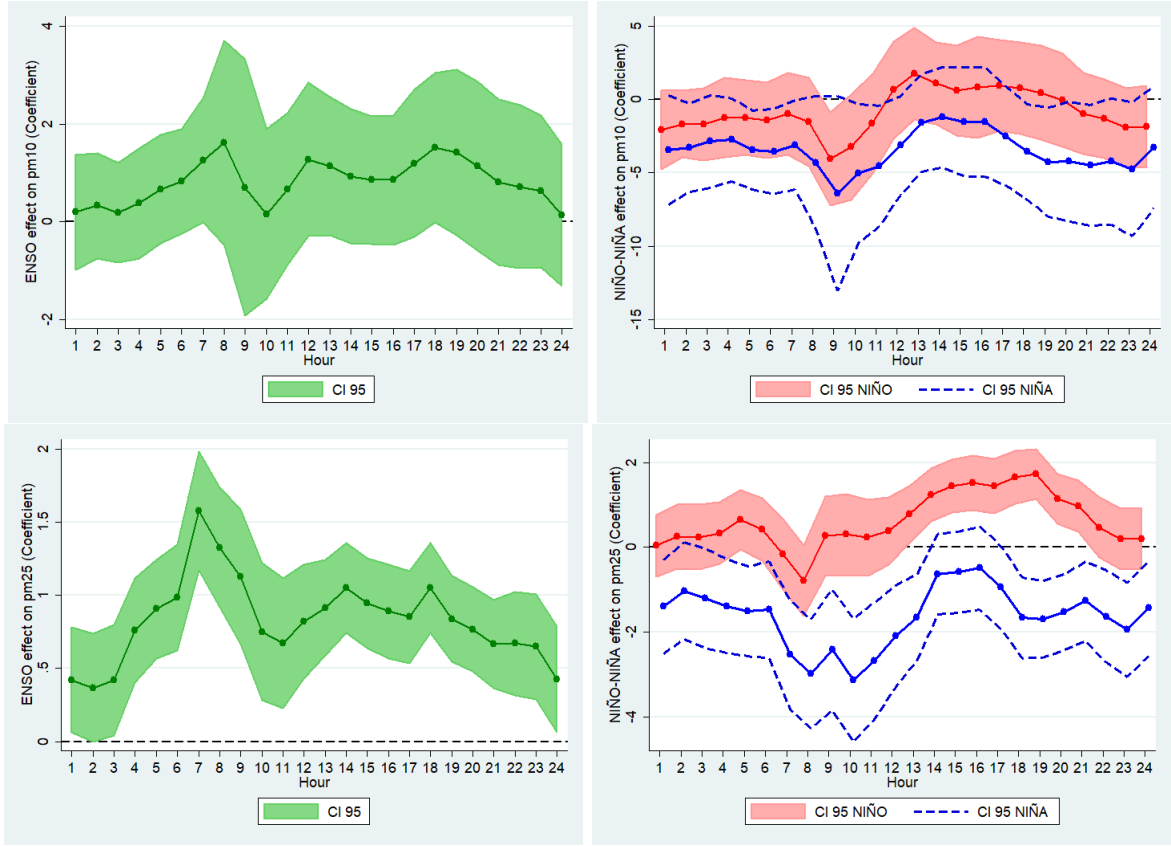
Additionally, the effect on weather is higher during the afternoon peak than during the morning peak. The results of these estimations can be found in Table 2.11 of Appendix 2B.5 as MODEL J; as explained, an alternative MODEL B was estimated, but it is shown only in the appendices for illustrative purposes. In terms of weather, the Institute of Hydrology, Meteorology and Environmental Studies of Colombia (known by its acronym in Spanish, IDEAM) has conducted some studies of the effects of El Niño-La Niña on weather, but only on a monthly base. In this sense, this study provides a deeper analysis on an hourly basis for the city of Bogotá.

2B.4 Results of ONI-NIÑO-NIÑA on pollution - Hourly Analysis

Although information is available for different pollutants, the analysis will focus on four contaminants, particulate matters or PM_{10} and $PM_{2.5}$, Nitrogen Oxide NO_x and Carbon Monoxide CO . In the same way as for weather factors, figure 2.8 gives the estimated coefficients of ONI-NIÑO-NIÑA on particulate matters PM_{10} and $PM_{2.5}$ hour by hour that were described in equation 3 and equation 5. The figure shows that the ONI index affects more PM_{10} than $PM_{2.5}$, with significant results for $PM_{2.5}$. In spite of that, both particulate matters follow similar patterns during the hourly schedule, increasing early in the morning and reaching a maximum peak at 8am-9am, decreasing and fluctuating around the same range of values later on. Interestingly, the peak in the afternoon is less pronounced.

If the ENSO events are isolated using equation 5, effects are found more of La Niña than El Niño on PM_{10} , but they are very small. La Niña effects are bigger for $PM_{2.5}$, reducing the concentration levels during the morning peak. It can reach a reduction of 0.3 micrograms per cubic meter on particulate matter at 10am with respect to an average of 31.38 micrograms per cubic meter during normal months at 10am (0.95% less). Although small during a particular hour, this reduction is not negligible when it is accumulated during a whole day or month. On the other hand, $PM_{2.5}$ concentrations tend to increase during the afternoon of Niño events, being higher around 5pm-7pm when individuals are returning home. Compared with normal days at the same hours, this corresponds to an increase of 1.1%-0.98%.

Figure 2.8: ONI-NIÑO-NIÑA on Pollutants ($PM_{10} - PM_{2.5}$)-24hours schedule

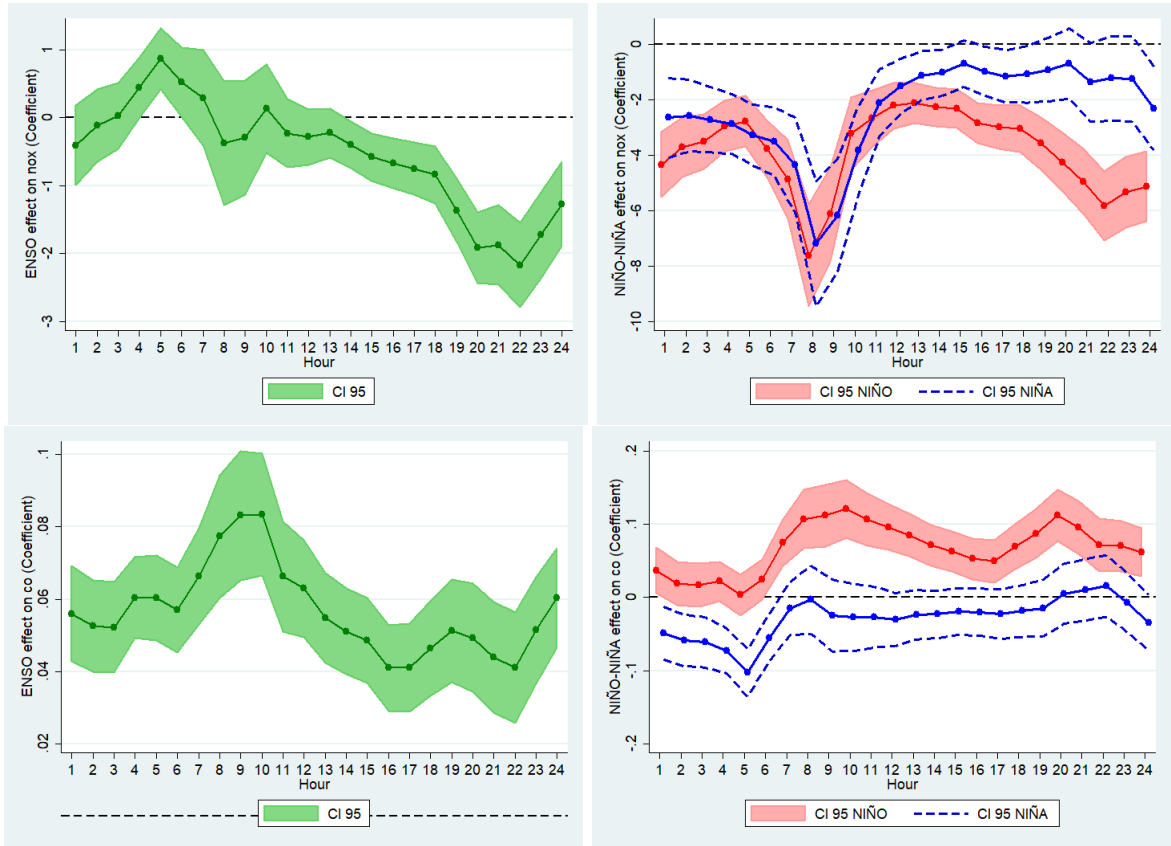


Note: based on the Air Quality Monitoring Network Data for Bogotá (RMCAB)
 ENSO(ONI) in green and NIÑO-NIÑA in red and blue respectively

At first, we argue that ENSO events could affect pollution. Nonetheless, not all pollutants are affected in the same way and in some cases, they could decrease such as it happens during La Niña. A potential explanation could be the interaction with weather factors. For instance, NIÑA brings more rain as was shown before and this could help to reduce particulate matters. On the contrary, El Niño brings higher temperature and less rain, increasing particulate matters $PM_{2.5}$.

Figure 2.9 shows the estimated coefficients of ONI-NIÑO-NIÑA on Nitrogen Oxide NO_x and Carbon Monoxide CO hour by hour. With respect to the ONI index, it decreases NO_x concentrations. CO follows a similar pattern as the particulate matters, being more important in the morning peak and much less in the afternoon. In terms of the ENSO events, NO_x decreases less in La Niña than El Niño (in comparison to normal months), and in almost all the hours with significant results. Also, the peak of concentration is reached around 5am, with a large reduction after that hour for

Figure 2.9: ONI-NIÑO-NIÑA on Pollutants ($NO_x - CO$)-24hour schedule



Note: based on the Air Quality Monitoring Network Data for Bogotá (RMCAB) ENSO(ONI) in green and NIÑO-NIÑA in red and blue respectively

both ENSO events, reaching a maximum reduction at 8am. CO also has higher levels for El Niño than La Niña, but only significant during El Niño. In the last case, the concentration levels of CO tend to increase during the peaks of city’s mobility.

To interpret these results, weather factors can help to give some hints. We find that pollution correlates positively with temperature and negatively with rain. As El Niño brings higher temperatures, pollutants such as particulate matter and CO increase during this event, in particular during peak hours. In addition, La Niña brings more rain which could help to wash out the levels of pollution. In terms of traffic, the city historically tends to exceed more the limits of particulate matters than the levels of Carbon Monoxide (CO). An explanation of this comes from the fact that vehicles are an important source of CO . For the case of Bogotá, the private car fleet is quite large. As new cars have been purchased in recent years, we should expect that they pollute less as they are

better equipped with catalytic converters. The same does not hold for particulate matters, which are emitted mainly by buses and public transport that use diesel and tend to be older. It is then not a surprise that unfortunately, particulate matters exceed more frequently the recommended limits for Bogotá.

2B.5 Results: ONI-Niño-Niña on weather

Table 2.11: OLS Estimations on Weather Variables (24 hours schedule)

HOUR:	ONI INDEX				NINO				NINA			
	RAIN		TEMPERATURE		RAIN		TEMPERATURE		RAIN		TEMPERATURE	
	Model B (1)	Model J (2)	Model B (3)	Model J (4)	Model B (5)	Model J (6)	Model B (7)	Model J (8)	Model B (9)	Model J (10)	Model B (11)	Model J (12)
h1	-0.0146*** (0.00220)	-0.0146*** (0.00217)	0.220** (0.0801)	0.232** (0.0722)	-0.0134* (0.00541)	0.000951 (0.00497)	0.0747 (0.110)	0.259** (0.0947)	0.0210*** (0.00373)	0.0350*** (0.00401)	-0.384* (0.160)	-0.266* (0.122)
h2	-0.0196*** (0.00111)	-0.0180*** (0.00125)	0.204* (0.0805)	0.211** (0.0720)	-0.0252*** (0.00285)	-0.0108*** (0.00323)	0.0657 (0.109)	0.271** (0.0931)	0.0173*** (0.00364)	0.0282*** (0.00411)	-0.361* (0.160)	-0.219 (0.122)
h3	-0.0152*** (0.00229)	-0.0145*** (0.00224)	0.189* (0.0814)	0.192** (0.0729)	-0.0120* (0.00505)	0.00149 (0.00430)	0.0352 (0.110)	0.257** (0.0946)	0.0245*** (0.00448)	0.0371*** (0.00400)	-0.355* (0.159)	-0.193 (0.121)
h4	-0.0122*** (0.00131)	-0.0111*** (0.00116)	0.176* (0.0827)	0.178* (0.0738)	-0.00992** (0.00371)	-0.000606 (0.00254)	0.0161 (0.113)	0.265** (0.0972)	0.0127*** (0.00283)	0.0199*** (0.00271)	-0.338* (0.158)	-0.156 (0.121)
h5	-0.0103*** (0.00122)	-0.00735*** (0.00111)	0.169* (0.0828)	0.169* (0.0728)	-0.0288*** (0.00347)	-0.0186*** (0.00328)	-0.00120 (0.111)	0.276** (0.0960)	-0.00523 (0.00272)	-0.00238 (0.00340)	-0.331* (0.158)	-0.123 (0.121)
h6	-0.00569*** (0.00129)	-0.00320** (0.00111)	0.167* (0.0840)	0.165* (0.0733)	-0.0235*** (0.00437)	-0.0125** (0.00417)	-0.0253 (0.111)	0.277** (0.0947)	-0.00834* (0.00336)	-0.00234 (0.00351)	-0.336* (0.160)	-0.104 (0.122)
h7	-0.00317** (0.00101)	-0.00121 (0.000845)	0.180* (0.0852)	0.174* (0.0739)	-0.00874* (0.00355)	0.0000377 (0.00254)	-0.0390 (0.111)	0.290** (0.0952)	-0.00217 (0.00311)	0.00323 (0.00258)	-0.369* (0.162)	-0.108 (0.122)
h8	-0.00476*** (0.00138)	-0.00315* (0.00126)	0.224** (0.0755)	0.197** (0.0635)	-0.0189*** (0.00359)	-0.0114*** (0.00283)	-0.0260 (0.0960)	0.311*** (0.0857)	-0.00439 (0.00313)	0.000540 (0.00299)	-0.441** (0.147)	-0.137 (0.111)
h9	-0.00297* (0.00121)	-0.00107 (0.00110)	0.279*** (0.0522)	0.234*** (0.0463)	-0.00243 (0.00298)	0.000391 (0.00215)	0.185* (0.0725)	0.278*** (0.0678)	0.00691*** (0.00198)	0.00705*** (0.00164)	-0.371*** (0.105)	-0.252* (0.100)
h10	-0.00242*** (0.000733)	-0.000923 (0.000586)	0.295*** (0.0408)	0.256*** (0.0387)	-0.00113 (0.00202)	-0.000444 (0.00117)	0.372*** (0.0980)	0.216** (0.0768)	0.00268 (0.00226)	0.00160 (0.00127)	-0.248* (0.106)	-0.350** (0.122)
h11	-0.00271* (0.00109)	-0.000602 (0.000933)	0.282*** (0.0399)	0.254*** (0.0408)	-0.00151 (0.00271)	0.00320 (0.00222)	0.454*** (0.102)	0.169* (0.0835)	0.00348 (0.00349)	0.00439 (0.00259)	-0.162 (0.121)	-0.395** (0.137)
h12	-0.00364** (0.00116)	-0.00139 (0.00113)	0.261*** (0.0449)	0.243*** (0.0445)	-0.00227 (0.00399)	0.00485 (0.00392)	0.493*** (0.114)	0.117 (0.0984)	0.00406 (0.00391)	0.00663* (0.00297)	-0.0974 (0.139)	-0.419** (0.148)
h13	-0.0128*** (0.00288)	-0.0125*** (0.00290)	0.302*** (0.0487)	0.291*** (0.0473)	-0.00694 (0.00608)	0.0116* (0.00568)	0.589*** (0.123)	0.152 (0.110)	0.0180* (0.00770)	0.0351*** (0.00704)	-0.102 (0.142)	-0.480** (0.148)
h14	-0.0190*** (0.00515)	-0.0212*** (0.00517)	0.392*** (0.0494)	0.382*** (0.0477)	-0.00910 (0.0115)	0.00603 (0.0107)	0.662*** (0.129)	0.228 (0.124)	0.0321*** (0.00713)	0.0485*** (0.00888)	-0.218 (0.130)	-0.582*** (0.152)
h15	-0.0527*** (0.00878)	-0.0602*** (0.00956)	0.499*** (0.0502)	0.487*** (0.0474)	-0.0240* (0.0122)	-0.0296* (0.0144)	0.673*** (0.121)	0.315** (0.118)	0.0965*** (0.0163)	0.105*** (0.0154)	-0.413*** (0.118)	-0.697*** (0.145)
h16	-0.0811*** (0.0103)	-0.0885*** (0.0100)	0.591*** (0.0531)	0.563*** (0.0487)	-0.0225 (0.0152)	-0.0638*** (0.0139)	0.666*** (0.124)	0.423*** (0.119)	0.147*** (0.0228)	0.123*** (0.0183)	-0.588*** (0.130)	-0.736*** (0.157)
h17	-0.0614*** (0.00778)	-0.0645*** (0.00747)	0.628*** (0.0532)	0.592*** (0.0475)	0.00730 (0.0133)	-0.0303* (0.0145)	0.639*** (0.117)	0.488*** (0.113)	0.134*** (0.0151)	0.105*** (0.0121)	-0.683*** (0.129)	-0.731*** (0.154)
h18	-0.0519*** (0.00803)	-0.0518*** (0.00789)	0.612*** (0.0539)	0.573*** (0.0488)	0.0119 (0.00831)	-0.0215 (0.0113)	0.565*** (0.108)	0.487*** (0.0976)	0.122*** (0.0113)	0.0904*** (0.0100)	-0.730*** (0.119)	-0.709*** (0.139)
h19	-0.0355*** (0.00708)	-0.0347*** (0.00697)	0.527*** (0.0531)	0.498*** (0.0526)	-0.000157 (0.00913)	-0.0288* (0.0115)	0.416*** (0.107)	0.428*** (0.0959)	0.0628*** (0.0120)	0.0293** (0.0107)	-0.703*** (0.119)	-0.624*** (0.129)
h20	-0.0248*** (0.00536)	-0.0268*** (0.00532)	0.403*** (0.0548)	0.397*** (0.0529)	-0.00831 (0.0111)	-0.0247* (0.00997)	0.296** (0.0999)	0.347*** (0.0934)	0.0451*** (0.00927)	0.0307*** (0.00875)	-0.562*** (0.132)	-0.498*** (0.122)
h21	-0.0236*** (0.00241)	-0.0231*** (0.00251)	0.341*** (0.0595)	0.348*** (0.0559)	-0.0259*** (0.00643)	-0.0313*** (0.00547)	0.248* (0.101)	0.325*** (0.0950)	0.0318*** (0.00709)	0.0250*** (0.00668)	-0.472*** (0.141)	-0.418*** (0.120)
h22	-0.0213*** (0.00215)	-0.0210*** (0.00226)	0.300*** (0.0652)	0.315*** (0.0599)	-0.0136* (0.00569)	-0.00927 (0.00649)	0.206* (0.103)	0.313*** (0.0945)	0.0301*** (0.00645)	0.0352*** (0.00652)	-0.427** (0.149)	-0.367** (0.120)
h23	-0.0168*** (0.00193)	-0.0158*** (0.00213)	0.272*** (0.0713)	0.288*** (0.0648)	-0.00306 (0.00443)	0.00648 (0.00391)	0.158 (0.105)	0.292** (0.0940)	0.0236*** (0.00406)	0.0326*** (0.00293)	-0.415** (0.158)	-0.339** (0.123)
h24	-0.0139*** (0.00377)	-0.0129*** (0.00380)	0.244** (0.0762)	0.259*** (0.0690)	-0.00236 (0.00750)	0.0125 (0.00677)	0.102 (0.110)	0.262** (0.0957)	0.0245*** (0.00362)	0.0386*** (0.00401)	-0.408* (0.160)	-0.312* (0.122)

Standard errors in parentheses
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note:

Model B: Weather = $f(\text{ONI}, \text{HOUR fixed effects}(h=1, \dots, 24), \text{ONI} \times \text{HOUR}(1, \dots, 24))$ taking $h=1$ as base. Hence, effect of $h=1$: ONI; effect $h=2$: ONI+ONI*HOUR($h=2$); effect $h=3$: ONI+ONI*HOUR($h=3$);...

Model J: Weather = $f(\text{ONI}(h))$ if $h=1, \dots, 24$

For Niño and Niña, both are included at the same time

Model B and Model J include Station fixed effects, Station-Trend fixed effects and month fixed effects

Standard Errors clustered at station level

2B.6 Results: ONI-NIÑO-NIÑA on pollution

Table 2.12: OLS Estimations on Pollutants (PM_{10} - $PM_{2.5}$) (24 hours schedule)

HOUR:	ONI INDEX				NINO				NINA			
	PM10		PM2.5		PM10		PM2.5		PM10		PM2.5	
	Model B (1)	Model J (2)	Model B (3)	Model J (4)	Model B (5)	Model J (6)	Model B (7)	Model J (8)	Model B (9)	Model J (10)	Model B (11)	Model J (12)
h1	0.322 (0.561)	0.196 (0.557)	0.580*** (0.127)	0.423* (0.183)	-2.411 (1.379)	-2.091 (1.277)	0.745** (0.278)	0.0395 (0.371)	-4.139* (1.607)	-3.444 (1.760)	-0.237 (0.394)	-1.402* (0.567)
h2	0.510 (0.550)	0.325 (0.508)	0.837*** (0.128)	0.367 (0.188)	-1.910 (1.263)	-1.676 (1.085)	1.110*** (0.282)	0.257 (0.384)	-4.055** (1.415)	-3.275* (1.430)	-0.830* (0.398)	-1.025 (0.582)
h3	0.528 (0.579)	0.186 (0.485)	0.507*** (0.129)	0.422* (0.193)	-1.923 (1.311)	-1.688 (1.168)	0.952*** (0.280)	0.242 (0.389)	-4.145** (1.472)	-2.858 (1.495)	-0.126 (0.402)	-1.209* (0.597)
h4	0.711 (0.634)	0.380 (0.533)	0.820*** (0.128)	0.759*** (0.182)	-1.384 (1.350)	-1.234 (1.281)	1.171*** (0.280)	0.335 (0.372)	-3.957** (1.400)	-2.753* (1.346)	-0.485 (0.398)	-1.385* (0.568)
h5	0.957 (0.621)	0.665 (0.528)	0.767*** (0.128)	0.907*** (0.172)	-1.455 (1.396)	-1.241 (1.206)	1.220*** (0.280)	0.650 (0.353)	-4.591*** (1.373)	-3.456** (1.275)	-0.189 (0.397)	-1.510** (0.539)
h6	1.233* (0.558)	0.825 (0.509)	0.718*** (0.128)	0.984*** (0.185)	-1.253 (1.343)	-1.417 (1.226)	0.877** (0.280)	0.426 (0.379)	-4.642*** (1.405)	-3.529* (1.380)	-0.0676 (0.396)	-1.476* (0.581)
h7	1.376* (0.601)	1.258* (0.603)	1.243*** (0.128)	1.578*** (0.207)	-0.726 (1.425)	-0.969 (1.319)	0.698* (0.278)	-0.171 (0.425)	-3.914** (1.433)	-3.096* (1.440)	-0.986* (0.396)	-2.536*** (0.655)
h8	1.416 (0.914)	1.614 (0.987)	0.482*** (0.127)	1.328*** (0.208)	-0.858 (1.583)	-1.530 (1.429)	-0.979*** (0.277)	-0.780 (0.424)	-4.064 (2.318)	-4.326* (2.149)	-1.572*** (0.395)	-2.986*** (0.657)
h9	0.370 (1.289)	0.698 (1.243)	-1.066*** (0.128)	1.127*** (0.234)	-2.680 (1.700)	-4.044** (1.514)	-2.391*** (0.280)	0.266 (0.478)	-4.323 (3.614)	-6.423* (3.146)	0.260 (0.395)	-2.420*** (0.728)
h10	-0.487 (0.938)	0.154 (0.823)	-1.641*** (0.128)	0.750** (0.239)	-1.592 (1.749)	-3.251 (1.712)	-2.551*** (0.282)	0.304 (0.488)	-1.773 (2.869)	-5.058* (2.252)	1.242** (0.396)	-3.138*** (0.740)
h11	-0.239 (0.750)	0.666 (0.735)	-1.021*** (0.128)	0.672** (0.227)	0.328 (1.474)	-1.628 (1.615)	-1.268*** (0.280)	0.227 (0.461)	0.128 (2.037)	-4.535* (1.958)	1.444*** (0.397)	-2.683*** (0.704)
h12	0.426 (0.712)	1.277 (0.741)	0.589*** (0.129)	0.819*** (0.198)	1.445 (1.221)	0.640 (1.563)	0.499 (0.282)	0.385 (0.401)	0.00357 (1.635)	-3.132 (1.600)	-0.522 (0.400)	-2.100*** (0.613)
h13	0.836 (0.715)	1.139 (0.669)	1.179*** (0.129)	0.916*** (0.167)	0.976 (1.172)	1.748 (1.485)	0.386 (0.283)	0.774* (0.340)	-1.508 (1.602)	-1.608 (1.579)	-2.929*** (0.400)	-1.667** (0.519)
h14	0.739 (0.701)	0.930 (0.650)	1.285*** (0.129)	1.052*** (0.157)	-0.436 (1.051)	1.105 (1.331)	0.446 (0.285)	1.237*** (0.319)	-2.495 (1.688)	-1.199 (1.620)	-3.353*** (0.400)	-0.634 (0.486)
h15	0.409 (0.648)	0.851 (0.622)	1.305*** (0.129)	0.946*** (0.158)	-0.861 (1.233)	0.595 (1.453)	0.687* (0.284)	1.442*** (0.320)	-2.349 (1.772)	-1.514 (1.757)	-3.589*** (0.400)	-0.582 (0.486)
h16	0.637 (0.656)	0.854 (0.629)	1.606*** (0.129)	0.890*** (0.163)	0.289 (1.445)	0.834 (1.629)	1.269*** (0.283)	1.523*** (0.330)	-2.008 (1.778)	-1.512 (1.768)	-4.034*** (0.399)	-0.490 (0.501)
h17	1.224 (0.749)	1.191 (0.714)	1.835*** (0.129)	0.852*** (0.161)	0.962 (1.409)	0.957 (1.475)	1.649*** (0.282)	1.446*** (0.326)	-2.751 (1.728)	-2.511 (1.597)	-4.372*** (0.398)	-0.944 (0.500)
h18	1.590* (0.809)	1.519* (0.725)	2.317*** (0.128)	1.049*** (0.157)	0.312 (1.452)	0.772 (1.478)	1.988*** (0.281)	1.654*** (0.318)	-4.157* (1.835)	-3.557* (1.544)	-5.938*** (0.396)	-1.665*** (0.485)
h19	1.647 (0.877)	1.420 (0.804)	2.109*** (0.128)	0.837*** (0.149)	-0.283 (1.598)	0.469 (1.528)	1.859*** (0.279)	1.726*** (0.302)	-5.141* (2.046)	-4.264* (1.757)	-5.957*** (0.394)	-1.697*** (0.463)
h20	1.235 (0.875)	1.138 (0.820)	1.745*** (0.128)	0.767*** (0.147)	-0.886 (1.557)	-0.0313 (1.506)	1.379*** (0.277)	1.142*** (0.295)	-4.888* (2.099)	-4.233* (1.932)	-4.963*** (0.394)	-1.530*** (0.455)
h21	0.860 (0.819)	0.812 (0.805)	1.269*** (0.128)	0.667*** (0.154)	-1.751 (1.383)	-0.967 (1.313)	1.228*** (0.278)	0.976** (0.310)	-4.825* (1.984)	-4.496* (1.949)	-2.761*** (0.393)	-1.269** (0.476)
h22	0.737 (0.730)	0.718 (0.788)	0.877*** (0.128)	0.670*** (0.180)	-1.584 (1.293)	-1.303 (1.265)	0.995*** (0.277)	0.466 (0.362)	-4.093* (1.807)	-4.226* (2.036)	-1.112** (0.392)	-1.635** (0.556)
h23	0.594 (0.630)	0.620 (0.736)	0.685*** (0.127)	0.649*** (0.183)	-2.097 (1.301)	-1.923 (1.272)	0.534 (0.277)	0.197 (0.368)	-4.390* (1.763)	-4.756* (2.170)	-1.264** (0.392)	-1.951*** (0.565)
h24	0.217 (0.608)	0.139 (0.685)	0.586*** (0.128)	0.427* (0.184)	-2.288 (1.384)	-1.864 (1.318)	0.730** (0.277)	0.200 (0.372)	-3.844* (1.708)	-3.257 (1.971)	-0.416 (0.393)	-1.438* (0.571)

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note:

Model B: Pollution = $f(\text{ONI}, \text{HOUR fixed effects}(h=1, \dots, 24), \text{ONI} * \text{HOUR}(1, \dots, 24))$ taking $h=1$ as base. Hence, effect of $h=1$: ONI; effect $h=2$: ONI+ONI*HOUR($h=2$); effect $h=3$: ONI+ONI*HOUR($h=3$);...

Model J: pollution = $f(\text{ONI}(h))$ if $h=1, \dots, 24$

For NIÑO and NIÑA, both are included at the same time

Model B and Model J include Station fixed effects, Station-Trend fixed effects, Day-of-the-week fixed effects, month fixed effects, Wind Speed, Wind Speed(squared), Dummies for 8 categories of Wind Direction, a Dummy for NO Car Day, Dummies for 3 phases of Transmilenio implementation and dummies for 3 changes in the Peak-and-Plate policy

Standard Errors clustered at station level

Table 2.13: OLS Estimations on Pollutants (NO_x -CO) (24 hours schedule)

HOUR:	ONI INDEX				NINO				NINA			
	NOx		CO		NOx		CO		NOx		CO	
	Model B (1)	Model J (2)	Model B (3)	Model J (4)	Model B (5)	Model J (6)	Model B (7)	Model J (8)	Model B (9)	Model J (10)	Model B (11)	Model J (12)
h1	-0.249 (0.185)	-0.409 (0.299)	0.0239*** (0.00578)	0.0560*** (0.00677)	-2.687*** (0.475)	-4.331*** (0.606)	-0.0138 (0.0146)	0.0367* (0.0159)	-1.578*** (0.475)	-2.642*** (0.731)	-0.0389** (0.0134)	-0.0490** (0.0186)
h2	-0.511** (0.185)	-0.114 (0.269)	0.0373*** (0.00580)	0.0526*** (0.00646)	-2.897*** (0.477)	-3.715*** (0.547)	-0.0514*** (0.0146)	0.0189 (0.0152)	-1.230** (0.477)	-2.568*** (0.658)	-0.122*** (0.0135)	-0.0588*** (0.0179)
h3	-0.571** (0.186)	0.0291 (0.248)	0.0509*** (0.00583)	0.0522*** (0.00636)	-3.043*** (0.479)	-3.497*** (0.504)	-0.0586*** (0.0147)	0.0169 (0.0149)	-1.236** (0.478)	-2.704*** (0.605)	-0.172*** (0.0136)	-0.0615*** (0.0175)
h4	-0.357 (0.186)	0.436 (0.227)	0.0627*** (0.00582)	0.0604*** (0.00574)	-2.812*** (0.480)	-2.938*** (0.462)	-0.0608*** (0.0147)	0.0218 (0.0135)	-1.592*** (0.478)	-2.862*** (0.554)	-0.204*** (0.0136)	-0.0728*** (0.0158)
h5	0.797*** (0.187)	0.871*** (0.229)	0.0656*** (0.00584)	0.0603*** (0.00600)	-1.343** (0.481)	-2.768*** (0.466)	-0.0808*** (0.0147)	0.00332 (0.0142)	-3.034*** (0.480)	-3.265*** (0.558)	-0.237*** (0.0136)	-0.103*** (0.0165)
h6	2.395*** (0.188)	0.523* (0.259)	0.0726*** (0.00583)	0.0570*** (0.00600)	0.339 (0.480)	-3.766*** (0.527)	-0.0544*** (0.0147)	0.0246 (0.0142)	-5.491*** (0.481)	-3.496*** (0.631)	-0.219*** (0.0136)	-0.0553*** (0.0165)
h7	4.261*** (0.186)	0.293 (0.358)	0.0759*** (0.00579)	0.0664*** (0.00676)	3.735*** (0.479)	-4.861*** (0.726)	0.0495*** (0.0146)	0.0748*** (0.0160)	-6.779*** (0.477)	-4.344*** (0.875)	-0.0865*** (0.0135)	-0.0150 (0.0187)
h8	3.213*** (0.186)	-0.376 (0.467)	0.0484** (0.00578)	0.0773*** (0.00860)	0.825 (0.478)	-7.609*** (0.943)	0.174*** (0.0204)	0.107*** (0.0146)	-6.971*** (0.204)	-7.193*** (1.143)	0.154*** (0.0134)	-0.00334 (0.0237)
h9	0.0188 (0.186)	-0.296 (0.428)	0.0626*** (0.00578)	0.0831*** (0.00911)	-2.646*** (0.477)	-6.117*** (0.865)	0.289*** (0.0145)	0.112*** (0.0215)	-2.278*** (0.477)	-6.179*** (1.046)	0.242*** (0.0134)	-0.0245 (0.0250)
h10	-1.248*** (0.186)	0.131 (0.331)	0.0869*** (0.00578)	0.0834*** (0.00859)	-4.572*** (0.477)	-3.219*** (0.669)	-0.272*** (0.0146)	0.121*** (0.0203)	-1.471** (0.477)	-3.822*** (0.810)	0.135*** (0.0134)	-0.0272 (0.0236)
h11	-1.854*** (0.187)	-0.230 (0.256)	0.0858*** (0.00581)	0.0662*** (0.00772)	-6.187*** (0.478)	-2.670*** (0.518)	0.171*** (0.0146)	0.106*** (0.0181)	-1.690*** (0.478)	-2.113*** (0.625)	0.00437 (0.0135)	-0.0270 (0.0210)
h12	-1.376*** (0.188)	-0.281 (0.208)	0.0924*** (0.00585)	0.0630*** (0.00682)	-5.949*** (0.480)	-2.208*** (0.419)	0.104*** (0.0147)	0.0962*** (0.0159)	-2.634*** (0.480)	-1.500** (0.506)	-0.0861*** (0.0135)	-0.0299 (0.0185)
h13	-1.337*** (0.188)	-0.226 (0.185)	0.0846*** (0.00586)	0.0548*** (0.00634)	-6.155*** (0.481)	-2.104*** (0.373)	0.0731*** (0.0147)	0.0853*** (0.0148)	-2.757*** (0.481)	-1.121* (0.451)	-0.100*** (0.0135)	-0.0236 (0.0171)
h14	-1.714*** (0.188)	-0.395* (0.175)	0.0740*** (0.00585)	0.0511*** (0.00608)	-6.283*** (0.481)	-2.256*** (0.353)	0.0469** (0.0147)	0.0715*** (0.0142)	-2.017*** (0.480)	-1.017* (0.426)	-0.105*** (0.0135)	-0.0231 (0.0165)
h15	-1.879*** (0.188)	-0.581** (0.177)	0.0691*** (0.00584)	0.0487*** (0.00600)	-5.945*** (0.480)	-2.323*** (0.356)	0.0480** (0.0147)	0.0631*** (0.0140)	-1.255** (0.480)	-0.683 (0.429)	-0.0909*** (0.0135)	-0.0189 (0.0163)
h16	-1.900*** (0.188)	-0.668*** (0.186)	0.0633*** (0.00583)	0.0410*** (0.00603)	-6.020*** (0.479)	-2.831*** (0.374)	0.0487*** (0.0147)	0.0527*** (0.0142)	-1.189* (0.480)	-0.965* (0.453)	-0.0782*** (0.0135)	-0.0200 (0.0164)
h17	-1.746*** (0.187)	-0.748*** (0.199)	0.0579*** (0.00581)	0.0411*** (0.00627)	-6.018*** (0.478)	-2.979*** (0.399)	0.0531*** (0.0147)	0.0496*** (0.0148)	-1.560** (0.478)	-1.136* (0.484)	-0.0554*** (0.0134)	-0.0226 (0.0171)
h18	-1.533*** (0.186)	-0.840*** (0.213)	0.0568*** (0.00580)	0.0464*** (0.00662)	-5.535*** (0.476)	-3.038*** (0.430)	0.0781*** (0.0146)	0.0690*** (0.0156)	-1.510** (0.476)	-1.073* (0.520)	-0.0272* (0.0134)	-0.0186 (0.0180)
h19	-1.527*** (0.185)	-1.367*** (0.239)	0.0576*** (0.00578)	0.0511*** (0.00723)	-5.448*** (0.475)	-3.573*** (0.481)	0.109*** (0.0146)	0.0874*** (0.0171)	-1.477*** (0.475)	-0.905 (0.582)	0.00899 (0.0134)	-0.0145 (0.0197)
h20	-1.319*** (0.185)	-1.916*** (0.266)	0.0609*** (0.00576)	0.0493*** (0.00763)	-5.027*** (0.474)	-4.247*** (0.537)	0.159*** (0.0145)	0.112*** (0.0180)	-1.686*** (0.474)	-0.677 (0.649)	0.0567*** (0.0133)	0.00533 (0.0209)
h21	-1.104*** (0.185)	-1.874*** (0.297)	0.0418** (0.00576)	0.0439*** (0.00782)	-4.985*** (0.473)	-4.951*** (0.600)	0.135*** (0.0145)	0.0960*** (0.0184)	-2.035*** (0.473)	-1.366 (0.726)	0.0963*** (0.0133)	0.0102 (0.0214)
h22	-1.217*** (0.185)	-2.170*** (0.316)	0.0102 (0.00576)	0.0410*** (0.00785)	-4.617*** (0.473)	-5.825*** (0.639)	0.0787*** (0.0145)	0.0718*** (0.0184)	-1.134* (0.473)	-1.219 (0.773)	0.128*** (0.0133)	0.0155 (0.0215)
h23	-0.913*** (0.185)	-1.728*** (0.322)	0.0192*** (0.00576)	0.0514*** (0.00751)	-3.676*** (0.474)	-5.329*** (0.650)	0.0661*** (0.0145)	0.0700*** (0.0176)	-0.831 (0.474)	-1.245 (0.787)	0.0877*** (0.0133)	-0.00703 (0.0206)
h24	-0.665*** (0.185)	-1.269*** (0.316)	0.0256*** (0.00577)	0.0603*** (0.00706)	-3.407*** (0.474)	-5.112*** (0.639)	0.0336* (0.0145)	0.0618*** (0.0166)	-1.403** (0.474)	-2.307** (0.773)	0.0248 (0.0134)	-0.0348 (0.0194)

Standard errors in parentheses
 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note:

Model B: Pollution = f(ONI, HOUR fixed effects(h=1,...,24), ONI*HOUR(1,...,24)) taking h=1 as base. Hence, effect of h=1: ONI; effect h=2: ONI+ONI*HOUR(h=2); effect h=3: ONI+ONI*HOUR(h=3);...

Model J: pollution = f(ONI(h)) if h=1,...,24

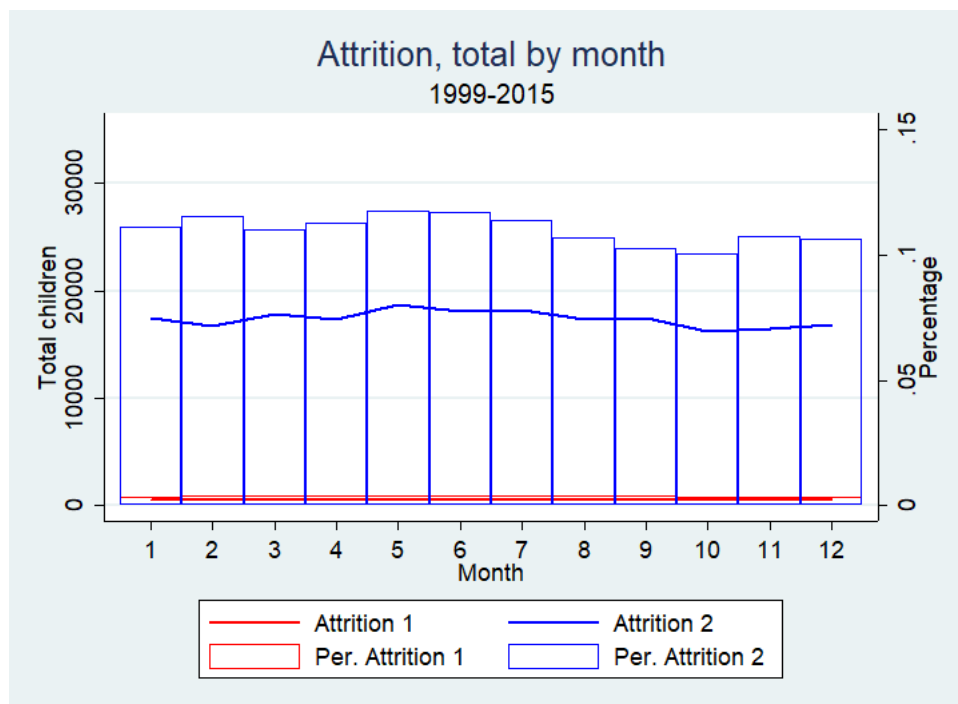
For NIÑO and NIÑA, both are included at the same time

Model B and Model J include Station fixed effects, Station-Trend fixed effects, Day-of-the-week fixed effects, month fixed effects, Wind Speed, Wind Speed(squared), Dummies for 8 categories of Wind Direction, a Dummy for NO Car Day, Dummies for 3 phases of Transmilenio implementation and dummies for 3 changes in the Peak-and-Plate policy

Standard Errors clustered at station level

2C Attrition in the sample

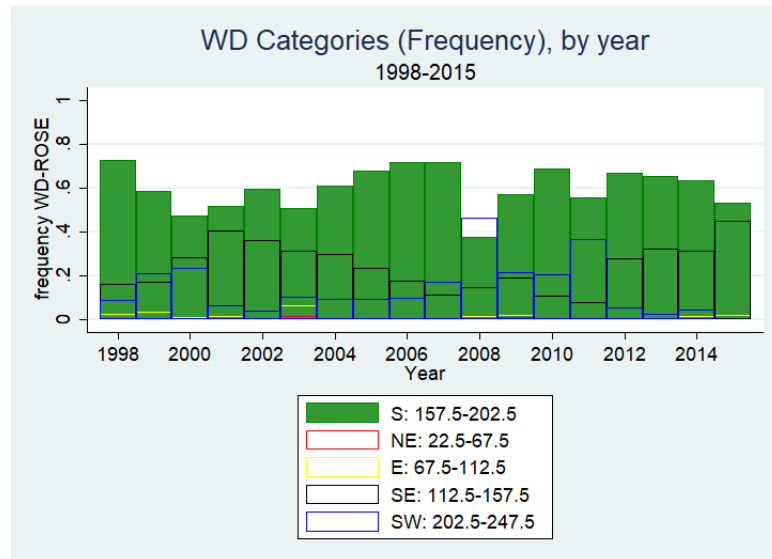
Figure 2.10: Type of attrition by month and source
Percentage and Total



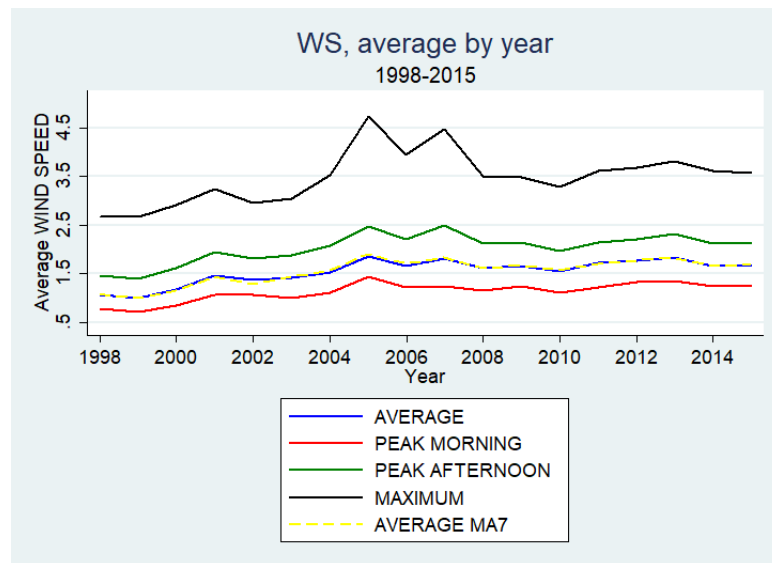
Note: based on the Air Quality Monitoring Network Data for Bogotá (RMCAB). Attrition could come from two sources: attrition for reason one, for children not being delivered at the health center and attrition for reason two, for children born in a health center but for which we could not recover the address.

2D Wind-direction over year

Figure 2.11: Categories of wind direction and wind speed per year



(a) Wind-direction

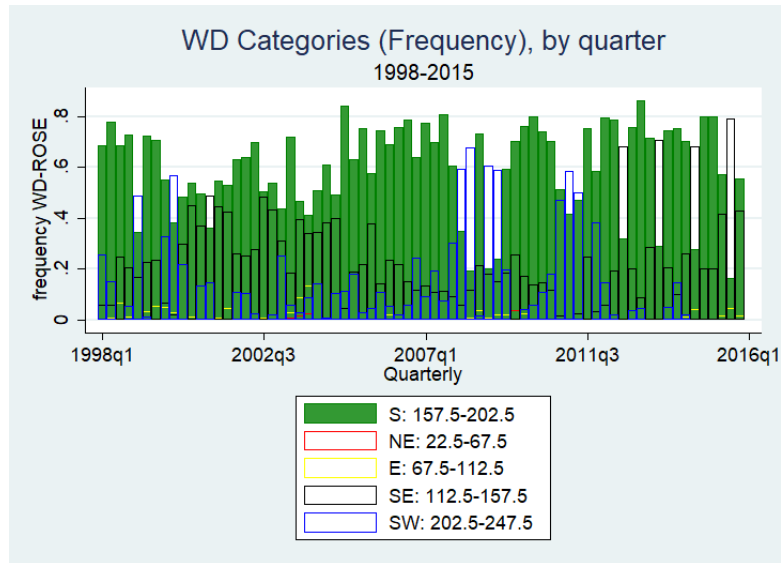


(b) Wind-speed

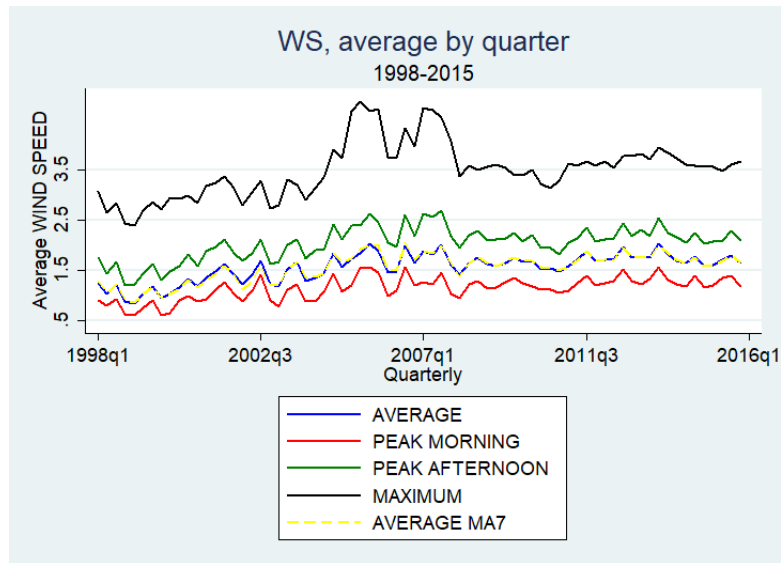
Note: based on the Air Quality Monitoring Network Data for Bogotá (RMCAB) .

In figure 2.11 and figure 2.12, part a), we can see the yearly and quarterly frequency of the different wind direction categories, with the wind blowing from the south in almost 60% of the cases, followed by wind blowing to the south-east (SE) and south-west (SW). It is important to notice that there is not only cross-section variation, but also variation along the years. Part b)

Figure 2.12: Categories of wind direction and wind speed per quarter



(a) Wind-direction

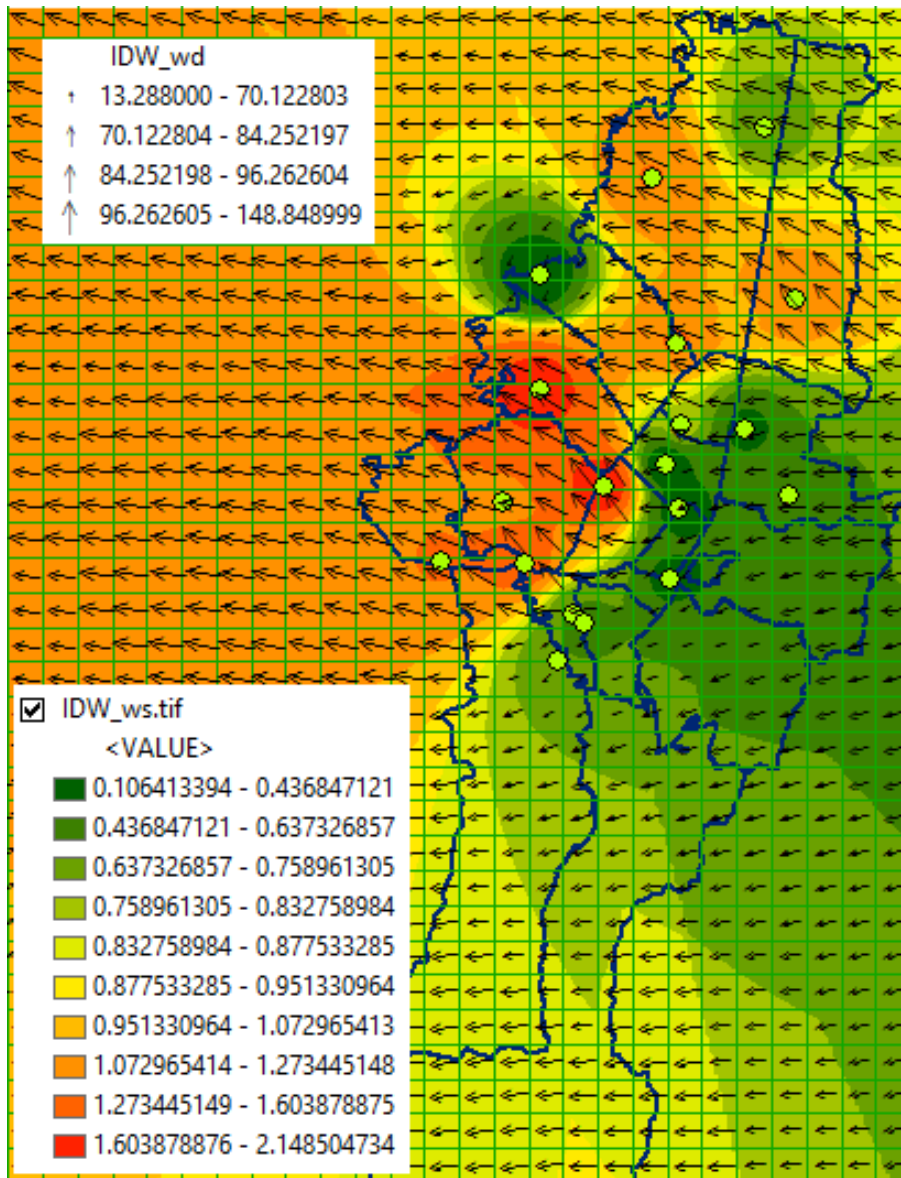


(b) Wind-speed

Note: based on the Air Quality Monitoring Network Data for Bogotá (RM CAB).

shows the average of the different variables used of wind speed by year and for all the health centers of the data-set. Here, we can also observe some variation along the years, with higher values around 2005-2007. The different time and cross-section variation in the data help in the identification strategy.

Figure 2.13: Map of wind direction (WD) and wind speed (WS) in the city



Note: based on the Air Quality Monitoring Network Data for Bogotá (RMCAB). WD in degrees and WS in meters/seconds. Map constructed using inverse distance weighting for the monitoring stations of the network. For WD, a value of 90 degrees corresponds to wind blowing from the east to the west. As can be seen in the map, the wind blows from the north and from the mountains that limit the city in the east, towards the south and west.

2E Reduced form results IV on health

Table 2.14: REDUCED FORM: IV ON HEALTH SO_2 (average)

	SO2 (average)			
	Premature	Low Weight	Weeks Gest.	Birth Weight
	(1)	(2)	(3)	(4)
Niño-Q3	0.004 (0.006)	0.007 (0.005)	-0.028 (0.046)	-25.230*** (7.726)
Niño-Q2	0.005 (0.006)	-0.001 (0.004)	-0.038 (0.044)	2.211 (7.459)
Niño-Q1	0.004 (0.007)	0.002 (0.005)	0.012 (0.051)	-4.597 (9.274)
Niña-Q3	0.003 (0.007)	0.003 (0.004)	0.014 (0.051)	10.898 (8.416)
Niña-Q2	0.000 (0.008)	-0.001 (0.006)	0.002 (0.060)	-10.686 (9.286)
Niña-Q1	-0.009 (0.009)	-0.011 (0.008)	0.094 (0.072)	8.040 (13.297)
ws_ave-Q3	-0.046 (0.042)	0.000 (0.020)	0.409 (0.322)	-11.770 (34.705)
ws_ave-Q2	-0.028 (0.046)	0.005 (0.029)	0.204 (0.338)	-37.445 (88.949)
ws_ave-Q1	-0.061 (0.047)	0.006 (0.046)	0.458 (0.360)	-90.533 (80.426)
wd_rose_Cat3-Q3	0.004 (0.036)	-0.020 (0.029)	0.044 (0.284)	17.173 (50.170)
wd_rose_Cat4-Q3	-0.036 (0.041)	0.014 (0.020)	0.331 (0.322)	23.530 (34.725)
wd_rose_Cat5-Q3	-0.039 (0.041)	0.003 (0.021)	0.382 (0.321)	27.916 (35.730)
wd_rose_Cat6-Q3	-0.042 (0.043)	-0.001 (0.026)	0.390 (0.341)	43.269 (44.856)
wd_rose_Cat3-Q2	-0.035 (0.055)	0.001 (0.029)	0.283 (0.390)	-7.503 (96.731)
wd_rose_Cat4-Q2	-0.029 (0.059)	0.010 (0.026)	0.211 (0.421)	-14.646 (95.165)
wd_rose_Cat5-Q2	-0.036 (0.062)	0.013 (0.028)	0.266 (0.440)	-12.909 (96.336)
wd_rose_Cat6-Q2	-0.039 (0.064)	0.009 (0.033)	0.321 (0.455)	-16.121 (95.670)
wd_rose_Cat3-Q1	-0.076 (0.047)	-0.020 (0.048)	0.582 (0.391)	-67.897 (88.900)
wd_rose_Cat4-Q1	-0.070 (0.049)	0.026 (0.042)	0.558 (0.404)	-66.923 (82.829)
wd_rose_Cat5-Q1	-0.070 (0.052)	0.031 (0.046)	0.566 (0.421)	-78.409 (86.025)
wd_rose_Cat6-Q1	-0.076 (0.053)	0.037 (0.047)	0.617 (0.430)	-68.770 (86.696)
<i>N</i>	11027	11027	11027	11027
R2a_RF	0.37	0.31	0.35	0.34

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. Each regression of the health outcomes has three equations in the first stage, one for each quarter of gestation. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay). The interaction of the wind direction rose with wind speed not shown to save space.

Table 2.15: REDUCED FORM: IV ON HEALTH PM_{10} $PM_{2.5}$ (average)

	PM10 (average)				PM2.5 (average)			
	Premature	Low Weight	Weeks Gest.	Birth Weight	Premature	Low Weight	Weeks Gest.	Birth Weight
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Niño-Q3	0.004 (0.006)	0.007 (0.005)	-0.028 (0.046)	-25.230*** (7.726)	0.016 (0.020)	-0.024* (0.013)	-0.129 (0.153)	-10.128 (17.885)
Niño-Q2	0.005 (0.006)	-0.001 (0.004)	-0.038 (0.044)	2.211 (7.459)	0.046*** (0.014)	0.001 (0.010)	-0.291*** (0.103)	-21.603 (17.364)
Niño-Q1	0.004 (0.007)	0.002 (0.005)	0.012 (0.051)	-4.597 (9.274)	0.012 (0.015)	0.015 (0.012)	-0.049 (0.115)	-28.714 (18.102)
Niña-Q3	0.003 (0.007)	0.003 (0.004)	0.014 (0.051)	10.898 (8.416)	-0.026 (0.028)	-0.015 (0.019)	0.283 (0.220)	25.935 (30.873)
Niña-Q2	0.000 (0.008)	-0.001 (0.006)	0.002 (0.060)	-10.686 (9.286)	0.039 (0.032)	0.005 (0.017)	-0.274 (0.237)	-40.906 (26.000)
Niña-Q1	-0.009 (0.009)	-0.011 (0.008)	0.094 (0.072)	8.040 (13.297)	0.003 (0.022)	-0.009 (0.019)	0.082 (0.186)	-3.457 (32.556)
ws_ave-Q3	-0.046 (0.042)	0.000 (0.020)	0.409 (0.322)	-11.770 (34.705)	-0.056 (0.488)	0.350 (0.451)	0.249 (3.353)	92.891 (576.529)
ws_ave-Q2	-0.028 (0.046)	0.005 (0.029)	0.204 (0.338)	-37.445 (88.949)	-0.779* (0.469)	-0.567 (0.353)	5.993* (3.264)	1306.477** (552.255)
ws_ave-Q1	-0.061 (0.047)	0.006 (0.046)	0.458 (0.360)	-90.533 (80.426)	-0.244 (0.325)	-0.137 (0.638)	1.331 (2.247)	-771.656 (564.553)
wd_rose_Cat3-Q3	0.004 (0.036)	-0.020 (0.029)	0.044 (0.284)	17.173 (50.170)	0.009 (0.572)	0.460 (0.543)	-0.206 (3.929)	69.150 (693.632)
wd_rose_Cat4-Q3	-0.036 (0.041)	0.014 (0.020)	0.331 (0.322)	23.530 (34.725)	-0.098 (0.584)	0.521 (0.538)	0.461 (4.005)	10.843 (672.912)
wd_rose_Cat5-Q3	-0.039 (0.041)	0.003 (0.021)	0.382 (0.321)	27.916 (35.730)	-0.089 (0.586)	0.502 (0.531)	0.416 (4.026)	28.201 (678.438)
wd_rose_Cat6-Q3	-0.042 (0.043)	-0.001 (0.026)	0.390 (0.341)	43.269 (44.856)	-0.075 (0.589)	0.537 (0.532)	0.316 (4.045)	-15.186 (683.301)
wd_rose_Cat3-Q2	-0.035 (0.055)	0.001 (0.029)	0.283 (0.390)	-7.503 (96.731)	-0.872 (0.554)	-0.630 (0.393)	6.769* (3.866)	1471.141** (633.713)
wd_rose_Cat4-Q2	-0.029 (0.059)	0.010 (0.026)	0.211 (0.421)	-14.646 (95.165)	-0.903* (0.543)	-0.580 (0.393)	6.975* (3.806)	1471.396** (644.099)
wd_rose_Cat5-Q2	-0.036 (0.062)	0.013 (0.028)	0.266 (0.440)	-12.909 (96.336)	-0.978* (0.542)	-0.601 (0.396)	7.430* (3.798)	1533.656** (650.463)
wd_rose_Cat6-Q2	-0.039 (0.064)	0.009 (0.033)	0.321 (0.455)	-16.121 (95.670)	-0.979* (0.549)	-0.603 (0.391)	7.477* (3.860)	1527.237** (645.125)
wd_rose_Cat3-Q1	-0.076 (0.047)	-0.020 (0.048)	0.582 (0.391)	-67.897 (88.900)	-0.326 (0.397)	-0.173 (0.767)	1.762 (2.721)	-880.841 (701.816)
wd_rose_Cat4-Q1	-0.070 (0.049)	0.026 (0.042)	0.558 (0.404)	-66.923 (82.829)	-0.359 (0.429)	-0.140 (0.786)	2.044 (2.945)	-781.602 (695.192)
wd_rose_Cat5-Q1	-0.070 (0.052)	0.031 (0.046)	0.566 (0.421)	-78.409 (86.025)	-0.403 (0.443)	-0.173 (0.788)	2.401 (3.042)	-759.094 (705.716)
wd_rose_Cat6-Q1	-0.076 (0.053)	0.037 (0.047)	0.617 (0.430)	-68.770 (86.696)	-0.399 (0.451)	-0.166 (0.794)	2.435 (3.096)	-741.730 (712.152)
<i>N</i>	11027	11027	11027	11027	4413	4413	4413	4413
R2a_RF	0.37	0.31	0.35	0.34	0.36	0.32	0.35	0.36

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: the estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. Each regression of the health outcomes has three equations in the first stage, one for each quarter of gestation. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay). The interaction of the wind direction rose with wind speed not shown to save space.

Table 2.16: REDUCED FORM: IV ON HEALTH $CO-NO_X$ (peak afternoon)

	CO (peak afternoon)				NOX (peak afternoon)			
	Premature	Low Weight	Weeks Gest.	Birth Weight	Premature	Low Weight	Weeks Gest.	Birth Weight
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Niño-Q3	-0.000 (0.008)	0.001 (0.006)	0.012 (0.058)	-14.769* (8.687)	0.002 (0.007)	0.005 (0.005)	-0.027 (0.051)	-24.453*** (8.910)
Niño-Q2	0.002 (0.010)	0.001 (0.006)	-0.008 (0.074)	-2.292 (10.511)	0.002 (0.007)	-0.004 (0.005)	-0.006 (0.055)	9.777 (9.009)
Niño-Q1	-0.001 (0.008)	0.003 (0.006)	0.048 (0.057)	-3.242 (10.691)	0.006 (0.007)	0.012** (0.006)	-0.002 (0.055)	-9.449 (9.992)
Niña-Q3	0.002 (0.008)	-0.001 (0.005)	0.031 (0.061)	15.286 (11.193)	-0.002 (0.008)	0.000 (0.005)	0.067 (0.064)	24.126** (10.141)
Niña-Q2	-0.005 (0.007)	-0.006 (0.005)	0.053 (0.057)	-2.561 (9.631)	-0.000 (0.008)	-0.003 (0.005)	-0.001 (0.065)	-9.733 (9.772)
Niña-Q1	-0.013 (0.010)	-0.013 (0.008)	0.133* (0.080)	13.798 (13.769)	-0.012 (0.009)	-0.014 (0.010)	0.125* (0.073)	13.809 (15.121)
ws_peaka-Q3	0.003 (0.076)	-0.031 (0.043)	0.069 (0.547)	88.708 (100.629)	-0.040 (0.026)	-0.004 (0.017)	0.338* (0.194)	-0.102 (30.092)
ws_peaka-Q2	-0.007 (0.042)	0.006 (0.035)	0.013 (0.335)	-68.466 (96.029)	-0.009 (0.030)	-0.002 (0.024)	0.063 (0.212)	-34.533 (66.402)
ws_peaka-Q1	-0.051 (0.041)	-0.007 (0.039)	0.398 (0.326)	-113.239* (68.392)	-0.013 (0.030)	0.028 (0.027)	0.091 (0.242)	-82.804 (60.878)
wd_rose_Cat3-Q3	0.033 (0.085)	-0.050 (0.056)	-0.151 (0.622)	148.113 (137.056)	-0.001 (0.030)	-0.018 (0.027)	0.051 (0.216)	25.528 (51.164)
wd_rose_Cat4-Q3	0.007 (0.081)	-0.015 (0.046)	0.079 (0.592)	159.136 (121.855)	-0.025 (0.037)	0.015 (0.018)	0.207 (0.269)	19.261 (37.467)
wd_rose_Cat5-Q3	0.014 (0.080)	-0.020 (0.047)	0.072 (0.597)	154.441 (122.599)	-0.032 (0.037)	0.009 (0.021)	0.282 (0.270)	17.363 (39.956)
wd_rose_Cat6-Q3	0.004 (0.084)	-0.021 (0.049)	0.103 (0.626)	165.091 (125.806)	-0.040 (0.040)	-0.005 (0.026)	0.341 (0.291)	44.138 (51.269)
wd_rose_Cat3-Q2	-0.036 (0.062)	0.013 (0.047)	0.239 (0.476)	-46.815 (125.521)	-0.008 (0.046)	0.001 (0.028)	0.144 (0.313)	-24.645 (88.320)
wd_rose_Cat4-Q2	-0.024 (0.065)	0.015 (0.047)	0.139 (0.503)	-46.427 (125.748)	-0.004 (0.049)	0.004 (0.026)	0.037 (0.337)	-34.288 (83.038)
wd_rose_Cat5-Q2	-0.021 (0.068)	0.025 (0.048)	0.107 (0.519)	-50.914 (127.515)	-0.007 (0.052)	0.009 (0.027)	0.082 (0.358)	-34.892 (84.557)
wd_rose_Cat6-Q2	-0.025 (0.071)	0.016 (0.052)	0.170 (0.534)	-36.637 (126.148)	-0.012 (0.057)	-0.003 (0.034)	0.168 (0.392)	-39.904 (86.578)
wd_rose_Cat3-Q1	-0.082* (0.047)	-0.032 (0.045)	0.632 (0.388)	-122.398 (93.935)	-0.038 (0.042)	0.002 (0.043)	0.273 (0.348)	-73.553 (84.268)
wd_rose_Cat4-Q1	-0.066 (0.046)	0.017 (0.040)	0.524 (0.381)	-134.205* (77.633)	-0.033 (0.044)	0.051 (0.039)	0.250 (0.359)	-66.464 (84.516)
wd_rose_Cat5-Q1	-0.066 (0.048)	0.016 (0.042)	0.556 (0.391)	-138.405* (81.505)	-0.020 (0.047)	0.060 (0.044)	0.168 (0.376)	-84.966 (88.103)
wd_rose_Cat6-Q1	-0.073 (0.048)	0.017 (0.044)	0.618 (0.396)	-119.186 (82.078)	-0.026 (0.049)	0.062 (0.046)	0.245 (0.389)	-66.360 (88.851)
<i>N</i>	9833	9833	9833	9833	7683	7683	7683	7683
R2a_RF	0.37	0.30	0.35	0.34	0.38	0.33	0.36	0.37

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: the estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay). The interaction of the wind direction rose with wind speed not shown to save space.

2F IV 2 STAGE: moving average 7 days-maximum

Table 2.17: IV SECOND STAGE: SO_2 (Ave. MA7) ON HEALTH

	SO2 (Ave. MA7)			
	Birth Weight (1)	Premature (2)	Low Weight (3)	Weeks Gest. (4)
Niño-Q3	0.011* (0.006)	0.014*** (0.005)	-0.086* (0.048)	-30.442*** (8.691)
Niño-Q2	-0.006 (0.007)	-0.011** (0.005)	0.041 (0.056)	12.065 (9.186)
Niño-Q1	0.008 (0.006)	0.012** (0.005)	-0.031 (0.044)	-16.271** (7.989)
Niña-Q3	0.002 (0.006)	0.006 (0.004)	0.008 (0.047)	3.104 (7.836)
Niña-Q2	-0.001 (0.006)	0.014*** (0.004)	-0.004 (0.042)	-21.704*** (7.378)
Niña-Q1	-0.014* (0.008)	-0.005 (0.006)	0.119** (0.060)	10.388 (10.102)
so2_avma7-Q3	0.000 (0.001)	-0.001 (0.001)	-0.004 (0.011)	-2.025 (2.282)
so2_avma7-Q2	-0.003** (0.002)	0.001 (0.002)	0.023* (0.012)	-0.819 (2.910)
so2_avma7-Q1	0.006*** (0.002)	0.004** (0.002)	-0.040** (0.016)	-6.757** (2.710)
Observations	11003	11003	11003	11003
NINO	0.01	0.01	-0.08	-34.65
P_NINO	0.05	0.01	0.12	0.00
NINA	-0.01	0.02	0.12	-8.21
P_NINA	0.25	0.07	0.12	0.52
POL	0.00	0.00	-0.02	-9.60
P_POL	0.24	0.04	0.25	0.00
r2_a	0.36	0.30	0.35	0.34

* p<0.05> , ** p<0.01, *** p<0.001

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the IPS, Weeks gest. is the average of weeks of gestation for all the children born in that IPS and Birth Weight is the average weight for all the children born in that IPS. For pollutants: Ave. MA7 is the moving average of the last 7 days calculated per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the three coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

Table 2.18: IV SECOND STAGE: PM_{10} - $PM_{2.5}$ (Ave. MA7) ON HEALTH

	PM10 (Ave. MA7)				PM2.5 (Ave. MA7)			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)	Premature (5)	Low Weight (6)	Weeks Gest. (7)	Birth Weight (8)
Niño-Q3	0.007 (0.006)	0.005 (0.005)	-0.054 (0.042)	-21.377*** (7.990)	-0.004 (0.011)	-0.004 (0.007)	0.032 (0.087)	-10.242 (12.922)
Niño-Q2	0.003 (0.006)	-0.003 (0.005)	-0.016 (0.047)	0.231 (7.567)	0.011 (0.009)	0.004 (0.007)	-0.065 (0.068)	-8.057 (9.587)
Niño-Q1	0.008 (0.006)	0.010* (0.006)	-0.020 (0.046)	-12.233 (9.464)	0.012 (0.010)	0.018*** (0.007)	-0.062 (0.072)	-22.852** (11.450)
Niña-Q3	0.000 (0.006)	0.001 (0.005)	0.021 (0.041)	6.481 (8.634)	-0.037** (0.017)	-0.015 (0.010)	0.305** (0.130)	18.927 (17.134)
Niña-Q2	0.002 (0.006)	0.009 (0.006)	-0.015 (0.045)	-21.097** (10.631)	0.036*** (0.012)	0.033*** (0.008)	-0.236** (0.093)	-44.788*** (10.371)
Niña-Q1	-0.002 (0.007)	0.000 (0.006)	0.037 (0.057)	-5.058 (10.350)	0.004 (0.012)	0.001 (0.010)	0.006 (0.092)	0.710 (13.547)
pm10_avma7-Q3	0.001 (0.001)	-0.001** (0.001)	-0.004 (0.005)	0.503 (0.772)				
pm10_avma7-Q2	0.000 (0.001)	0.001 (0.001)	-0.002 (0.005)	-1.448 (0.909)				
pm10_avma7-Q1	0.001 (0.001)	0.001 (0.001)	-0.003 (0.006)	-0.956 (1.256)				
pm25_avma7-Q3					0.005*** (0.002)	0.002 (0.001)	-0.037*** (0.014)	-2.315 (1.846)
pm25_avma7-Q2					-0.004** (0.002)	0.001 (0.001)	0.021 (0.014)	-1.838 (2.151)
pm25_avma7-Q1					0.000 (0.001)	-0.000 (0.001)	-0.001 (0.009)	0.137 (1.319)
Observations	11023	11023	11023	11023	4405	4405	4405	4405
NINO	0.02	0.01	-0.09	-33.38	0.02	0.02	-0.09	-41.15
P_NINO	0.02	0.06	0.09	0.01	0.24	0.05	0.43	0.01
NINA	-0.00	0.01	0.04	-19.67	0.00	0.02	0.08	-25.15
P_NINA	0.99	0.20	0.57	0.12	0.85	0.17	0.62	0.23
POL	0.00	0.00	-0.01	-1.90	0.00	0.00	-0.02	-4.02
P_POL	0.12	0.54	0.23	0.18	0.59	0.36	0.51	0.26
r2_a	0.36	0.30	0.35	0.34	0.35	0.32	0.35	0.36

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the IPS, Weeks gest. is the average of weeks of gestation for all the children born in that IPS and Birth Weight is the average weight for all the children born in that IPS. For pollutants: Ave. MA7 is the moving average of the last 7 days calculated per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the three coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

Table 2.19: IV SECOND STAGE: $CO-NO_X$ (Max.) ON HEALTH

	CO (Max.)				NOX (Max.)			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)	Premature (5)	Low Weight (6)	Weeks Gest. (7)	Birth Weight (8)
Niño-Q3	0.001 (0.006)	0.012** (0.006)	-0.011 (0.045)	-24.583*** (8.708)	0.002 (0.007)	0.009 (0.007)	-0.029 (0.052)	-20.476** (10.216)
Niño-Q2	-0.009 (0.009)	0.002 (0.006)	0.055 (0.069)	5.812 (9.745)	-0.005 (0.007)	-0.004 (0.006)	0.042 (0.049)	7.595 (10.584)
Niño-Q1	0.004 (0.007)	0.008 (0.006)	0.002 (0.053)	-11.676 (8.797)	0.010 (0.007)	0.021*** (0.007)	-0.044 (0.056)	-21.020** (9.097)
Niña-Q3	-0.003 (0.007)	0.003 (0.005)	0.051 (0.055)	2.673 (9.642)	-0.003 (0.007)	-0.003 (0.005)	0.051 (0.049)	12.851 (9.099)
Niña-Q2	-0.003 (0.006)	0.013*** (0.004)	0.017 (0.043)	-19.976** (8.549)	-0.008 (0.006)	0.015*** (0.005)	0.042 (0.044)	-19.955** (8.870)
Niña-Q1	-0.005 (0.008)	0.003 (0.005)	0.045 (0.058)	-0.254 (9.065)	-0.005 (0.008)	-0.004 (0.006)	0.061 (0.063)	4.233 (9.740)
co_max-Q3	0.003 (0.006)	0.008 (0.005)	-0.035 (0.048)	-4.968 (9.068)				
co_max-Q2	-0.008* (0.005)	0.009** (0.004)	0.048 (0.037)	-3.976 (6.459)				
co_max-Q1	0.001 (0.004)	0.003 (0.004)	-0.016 (0.035)	1.636 (8.263)				
nox_max-Q3					-0.000 (0.000)	-0.000 (0.000)	0.000 (0.002)	0.313 (0.248)
nox_max-Q2					0.000 (0.000)	0.000 (0.000)	-0.001 (0.001)	-0.338 (0.256)
nox_max-Q1					-0.000** (0.000)	-0.000 (0.000)	0.002** (0.001)	0.243 (0.205)
Observations	9963	9963	9963	9963	7705	7705	7705	7705
NINO	-0.00	0.02	0.05	-30.45	0.01	0.03	-0.03	-33.90
P_NINO	0.69	0.00	0.53	0.00	0.40	0.00	0.62	0.01
NINA	-0.01	0.02	0.11	-17.56	-0.02	0.01	0.15	-2.87
P_NINA	0.30	0.01	0.15	0.12	0.21	0.36	0.08	0.82
POL	-0.00	0.02	-0.00	-7.31	-0.00	-0.00	0.00	0.22
P_POL	0.76	0.02	0.97	0.66	0.32	0.92	0.41	0.60
r2_a	0.37	0.29	0.35	0.34	0.38	0.33	0.36	0.36

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the IPS, Weeks gest. is the average of weeks of gestation for all the children born in that IPS and Birth Weight is the average weight for all the children born in that IPS. For pollutants: Max. is the maximum per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the three coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

2G Variation in the spatial correlation over time and residential sorting

Table 2.20: Moran Index for main pollutants and weather variables (Yearly)

	PM10	Pval	PM25	Pval	O3	Pval	SO2	Pval	CO	Pval	NO	Pval
1998	0.68	0.00	-0.05	2.00	0.07	0.00	0.88	0.00	0.90	0.00	0.73	0.00
1999	0.86	0.00	-0.02	1.94	0.47	0.00	0.74	0.00	0.51	0.00	0.62	0.00
2000	0.67	0.00			0.13	0.00	0.87	0.00	0.54	0.00	0.75	0.00
2001	0.73	0.00			0.17	0.00	0.50	0.00	0.31	0.00	0.55	0.00
2002	0.71	0.00			0.24	0.00	0.68	0.00	0.63	0.00	0.63	0.00
2003	0.71	0.00			-0.00	1.15	0.71	0.00	0.10	0.00	0.01	0.21
2004	0.59	0.00			0.36	0.00	0.47	0.00	0.60	0.00	0.06	0.00
2005	0.80	0.00	-0.38	2.00	0.49	0.00	0.85	0.00	0.09	0.00		
2006	0.86	0.00	-0.18	2.00	0.36	0.00	0.49	0.00			0.14	0.00
2007	0.70	0.00	-0.00	0.98	-0.01	1.87	0.38	0.00	0.79	0.00	0.55	0.00
2008	0.51	0.00	0.44	0.00	0.00	0.62	0.51	0.00	0.46	0.00	0.35	0.00
2009	0.53	0.00	0.53	0.00	0.34	0.00	0.32	0.00	0.43	0.00	0.74	0.00
2010	0.76	0.00	0.27	0.00	0.39	0.00	0.59	0.00	0.63	0.00	0.83	0.00
2011	0.63	0.00	-0.05	2.00	0.48	0.00	0.89	0.00	0.43	0.00	0.35	0.00
2012	0.46	0.00	0.25	0.00	0.51	0.00	0.79	0.00	0.47	0.00	0.54	0.00
2013	0.55	0.00	0.38	0.00	0.54	0.00	0.79	0.00	0.27	0.00	0.70	0.00
2014	0.69	0.00	0.41	0.00	0.56	0.00	0.45	0.00	0.43	0.00	0.27	0.00
2015	0.57	0.00	0.48	0.00	0.63	0.00	0.72	0.00	0.58	0.00	0.63	0.00

	NO2	Pval	NOX	Pval	RAIN	Pval	TMP	Pval	WS	Pval
1998	0.51	0.00	0.74	0.00	0.25	0.00	0.50	0.00	0.74	0.00
1999	0.72	0.00	0.70	0.00	0.09	0.00	0.68	0.00	0.76	0.00
2000	0.49	0.00	0.82	0.00	0.22	0.00	0.53	0.00	0.83	0.00
2001	0.54	0.00	0.57	0.00	0.19	0.00	0.76	0.00	0.87	0.00
2002	0.51	0.00	0.64	0.00	0.04	0.00	0.70	0.00	0.80	0.00
2003	0.03	0.01	0.02	0.04	0.14	0.00	0.38	0.00	0.78	0.00
2004	0.09	0.00	0.15	0.00	0.12	0.00	0.49	0.00	0.68	0.00
2005					0.42	0.00	0.64	0.00	0.55	0.00
2006	0.19	0.00	0.27	0.00	0.12	0.00	0.76	0.00	0.61	0.00
2007	0.46	0.00	0.56	0.00	0.04	0.00	0.65	0.00	0.69	0.00
2008	0.07	0.00	0.31	0.00	0.17	0.00	0.29	0.00	0.80	0.00
2009	0.59	0.00	0.75	0.00	0.13	0.00	0.62	0.00	0.81	0.00
2010	0.69	0.00	0.90	0.00	0.08	0.00	0.16	0.00	0.90	0.00
2011	0.06	0.00			0.14	0.00	0.58	0.00	0.78	0.00
2012	0.28	0.00			0.05	0.00	0.78	0.00	0.79	0.00
2013	0.66	0.00			0.04	0.00	0.73	0.00	0.75	0.00
2014	0.41	0.00	0.27	0.00	0.00	0.48	0.77	0.00	0.70	0.00
2015	0.33	0.00	0.51	0.00	0.06	0.00	0.76	0.00	0.87	0.00

Note: Moran's I Statistic of spatial correlation per variable and per year, calculated for a spatial distance of 2km. It also has the respective p-value. All variables used are yearly average. The table shows how the spatial variation varies year by year for the main pollutants and the weather factors. We observe different spatial correlation, with some years with higher correlation than others for PM_{10} , and others like O_3 have higher variation, even towards negative values. Similar spatial correlation patterns are observed for weather variables too.

Table 2.21: OLS Estimations of Niño-Niña and Instrumental Variables (Wind direction and Wind Speed) on socioeconomic characteristics

	Father's age			Mother's age			Marital Status			Female's birth			Father's education			Mother's education			Type Social Security		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
nino	-0.009 (0.0967)	0.001 (0.0583)	0.001 (0.00449)	0.000 (0.00393)	-0.006 (0.00435)	0.014 (0.00790)	-0.012 (0.00951)	0.005 (0.00430)	-0.009 (0.00450)	0.004 (0.00230)	0.001 (0.00430)	0.009* (0.00406)	-0.003 (0.00525)	-0.013 (0.00461)							
niña	-0.001 (0.0992)	-0.006 (0.0623)	-0.014* (0.00472)	-0.007 (0.00380)	-0.003 (0.00366)	0.010 (0.00523)	-0.013 (0.00564)	-0.001 (0.00381)	0.008 (0.00393)	-0.008 (0.00272)	-0.000 (0.00437)	0.017*** (0.00470)	0.000 (0.00407)	-0.033** (0.00523)							
ws_ave	-0.041 (0.380)	-0.142* (0.499)	-0.062 (0.0413)	0.111 (0.0324)	0.240* (0.0317)	0.086 (0.0379)	-0.206 (0.0726)	0.221 (0.0370)	-0.010 (0.0682)	-0.155 (0.0909)	-0.038 (0.0355)	-0.018 (0.0648)	0.161*** (0.0252)	-0.121 (0.0260)							
3.wd_rose	-0.035** (0.543)	-0.013 (0.642)	-0.023 (0.0643)	0.018 (0.0362)	0.054 (0.0436)	0.009 (0.0646)	-0.034 (0.0725)	0.032 (0.0375)	0.005 (0.0770)	-0.031 (0.0956)	-0.004 (0.0580)	-0.007 (0.0839)	0.025 (0.0451)	-0.017 (0.0287)							
4.wd_rose	-0.025 (0.334)	-0.004 (0.487)	-0.027 (0.0573)	0.061 (0.0293)	0.098 (0.0334)	0.116 (0.0627)	-0.136 (0.0740)	0.065 (0.0344)	0.040 (0.0737)	-0.101 (0.0945)	-0.006 (0.0451)	-0.089 (0.0833)	0.100** (0.0315)	0.023 (0.0282)							
5.wd_rose	-0.020 (0.334)	-0.022 (0.487)	-0.025 (0.0573)	0.050 (0.0299)	0.123 (0.0330)	0.090 (0.0630)	-0.122 (0.0751)	0.088 (0.0347)	0.044 (0.0739)	-0.121 (0.0947)	-0.048 (0.0457)	-0.048 (0.0834)	0.115** (0.0310)	0.036 (0.0291)							
6.wd_rose	0.015 (0.400)	-0.013 (0.538)	-0.035 (0.0582)	0.094 (0.0327)	0.112 (0.0352)	0.057 (0.0645)	-0.101 (0.0758)	0.013 (0.0361)	0.073 (0.0748)	-0.107 (0.0949)	-0.021 (0.0480)	-0.023 (0.0844)	0.032 (0.0345)	0.047 (0.0308)							
3.wd_rose X ws_ave	0.020 (0.532)	0.030 (0.600)	0.026 (0.0464)	-0.047 (0.0366)	-0.084** (0.0384)	-0.008 (0.0605)	0.051 (0.0724)	-0.073* (0.0398)	0.006 (0.0703)	0.047 (0.0923)	0.005 (0.0442)	-0.005 (0.0712)	-0.038* (0.0367)	0.058*** (0.0256)							
4.wd_rose X ws_ave	0.028 (0.403)	0.150 (0.496)	0.073 (0.0429)	-0.177 (0.0321)	-0.273* (0.0331)	-0.126 (0.0578)	0.248 (0.0728)	-0.219 (0.0369)	-0.008 (0.0688)	0.181 (0.0918)	0.031 (0.0347)	0.005 (0.0653)	-0.167** (0.0273)	0.174* (0.0255)							
5.wd_rose X ws_ave	0.042 (0.394)	0.239* (0.492)	0.095 (0.0418)	-0.222 (0.0324)	-0.382* (0.0324)	-0.147 (0.0575)	0.318 (0.0731)	-0.339 (0.0373)	0.018 (0.0687)	0.245 (0.0918)	0.087 (0.0349)	-0.038 (0.0653)	-0.242*** (0.0268)	0.226 (0.0262)							
6.wd_rose X ws_ave	-0.010 (0.410)	0.154* (0.491)	0.083 (0.0426)	-0.207* (0.0332)	-0.281* (0.0333)	-0.093 (0.0580)	0.225 (0.0734)	-0.177 (0.0380)	-0.039 (0.0690)	0.188 (0.0916)	0.047 (0.0356)	-0.037 (0.0658)	-0.117* (0.0271)	0.130 (0.0270)							
FE YEAR-LOCALITY	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes							
FE MONTH	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes							
FE IPS HEALTH CENTER	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes							

Standardized beta coefficients; Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: the table shows the coefficient estimates of wind direction, wind speed (the instrumental variables for pollution) and Niño-Niña on some key socioeconomic characteristics of the running-quarter analysis. All variables are averaged at health center or IPS. Column 1 and 2 have the average of the father and mother's age, column 3 is the percentage of mothers with a couple (Marital Status), column 4 is the percentage of female born in the IPS, columns 5-7 have the percentage of fathers with primary, secondary and tertiary education in each IPS, columns 8-10 have the percentage of mothers with primary, secondary and tertiary education in each IPS, columns 11-14 have the percentage of mothers with different types of social security in each IPS. Wind Direction Rose: 1 "WD: N 337.5-22.5" 2 "WD: NE 22.5-67.5" 3 "WD: E 67.5-112.5" 4 "WD: SE 112.5-157.5" 5 "WD: S 157.5-202.5" 6 "WD: SW 202.5-247.5" 7 "WD: W 247.5-292.5" 8 "WD: NW 292.5-337.5". The standard errors in parentheses are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

Table 2.22: Socioeconomic characteristics from the EMB - Migration
a) Migration from locality

Locality of origin	2011		2014		Total		Diff	se
	Obs.	Per.	Obs.	Per.	Obs.	Per.		
USAQUEN	88899	7.0	90019	7.5	178919	7.2	0.003	(0.008)
CHAPINERO	36981	2.9	35667	3.0	72648	2.9	-0.002	(0.004)
SANTA FE	9920	0.8	8679	0.7	18600	0.8	0.002	(0.002)
SAN CRISTOBAL	70101	5.5	59047	4.9	129148	5.2	-0.004	(0.06)
USME	58024	4.5	57710	4.8	115734	4.7	0.002	(0.007)
TUNJUELITO	34867	2.7	24317	2.0	59184	2.4	-0.007	(0.004)
BOSA	117635	9.2	116513	9.7	234148	9.5	0.010	(0.020)
KENNEDY	159858	12.5	163085	13.6	322943	13.1	0.023	(0.030)
FONTIBON	51227	4.0	53906	4.5	105133	4.3	0.002	(0.006)
ENGATIVA	142864	11.2	117431	9.8	260295	10.5	-0.016	(0.006)**
SUBA	182169	14.3	189672	15.9	371841	15.0	-0.003	(0.007)
BARRIOS UNIDOS	23068	1.8	25638	2.1	48706	2.0	0.001	(0.006)
TEUSAQUILLO	20320	1.6	17904	1.5	38225	1.5	-0.002	(0.004)
LOS MARTIRES	14036	1.1	11044	0.9	25080	1.0	-0.001	(0.003)
ANTONIO NARINO	27111	2.1	17128	1.4	44239	1.8	-0.005	(0.004)
PUENTE ARANDA	34202	2.7	36659	3.1	70861	2.9	0.004	(0.005)
LA CANDELARIA	7884	0.6	7698	0.6	15581	0.6	0.000	(0.002)
RAFAEL URIBE URIBE	69241	5.4	50029	4.2	119270	4.8	-0.005	(0.006)
CIUDAD BOLIVAR	105299	8.3	90132	7.5	195430	7.9	-0.005	(0.007)
No Info	21813	1.7	23957	2.0	45770	1.9	0.004	(0.007)
Total	1275519	100.0	1196234	100.0	2471753	100.0		

b) Reasons to migrate

Reason to Migrate	2011	2014	2017	Total
	Per.	Per.	Per.	Per.
Economic	42.5	46.0	54.8	47.4
Family	28.2	28.4	20.9	26.0
Education	17.0	14.6	13.6	15.2
Risk for life security	9.3	7.5	4.1	7.1
Better dwelling	0.0	0.0	4.3	1.3
Health	2.2	1.9	1.2	1.8
Others	0.8	1.6	1.1	1.1
Total	100.0	100.0	100.0	100.0
Z-value	92.994	86.001	84.546	
Economic vs Health (p-val)	1	1	1	

Note: calculated using the Multipurpose Survey for Bogotá (EMB) for 2011, 2014 and 2017. Part a) uses the declared locality of origin for individuals moving inside the city, for whom the movement took place three years ago or less. Part b) shows the percentage of each declared reasons for migrating per year for the city. While part a) compares migration between localities, part b) analyses migration towards the city. Statistics are calculated using individual' weights provided by the EMB. In part b), migration is defined as migrating to the city five years ago or less; this definition allowed to homogenize the variables for the different waves. The original categories were: 1) For work or business, 2) Education opportunity, 3) Health motives, 4) Marriage or making family, 5) Risk-natural disaster, 6) Threat for life-risk, 7) Buying dwelling, 8) Better dwelling or location, 9) Problem or partner's problems, 10) Economic reason, 11) joint family member, 12) Others.

Part a) shows not large changes in the percentage of migration per each locality along the years we have in the EMB survey. Part b) shows that individuals migrate mainly for economic reasons and family reasons and once that problem is solved, they think on moving for other reasons such as health. The results again do not show large changes along the years. Individuals moving to avoid air pollution hence is less a concern in our results and migration between localities has been stable in the years we have information (no residential sorting in this sense). The column Diff of part a) does not correspond exactly to the difference between the columns of the percentages 2014 and 2011, as the test includes controls for stratus and locality fixed effects; the standard error in the tests are clustered at the locality level.

Table 2.23: Socioeconomic characteristics from the EMB

a) Descriptive Statistics

Locality	2011				2014				2017			
	Fem.	Age	Edu-year	Mig.	Fem.	Age	Edu-year	Mig.	Fem.	Age	Edu-year	Mig.
ANTONIO NARINO	0.52	32.9	10.8	0.08	0.51	33.9	10.4	0.04	0.50	35.9	11.8	0.04
BARRIOS UNIDOS	0.52	36.8	12.1	0.07	0.52	37.9	12.2	0.07	0.50	38.8	13.3	0.08
BOSA	0.51	28.4	8.6	0.09	0.51	29.5	8.7	0.09	0.50	31.1	9.6	0.08
CHAPINERO	0.54	36.8	14.6	0.19	0.53	37.9	14.6	0.18	0.51	37.1	14.9	0.13
CIUDAD BOLIVAR	0.51	27.7	8.4	0.07	0.51	28.6	8.6	0.04	0.50	30.5	9.0	0.08
ENGATIVA	0.52	33.2	11.2	0.08	0.52	34.3	11.2	0.06	0.50	35.7	11.9	0.06
FONTIBON	0.53	32.1	11.9	0.11	0.53	33.0	11.7	0.07	0.51	35.1	12.4	0.08
KENNEDY	0.51	30.8	9.9	0.09	0.51	31.7	10.2	0.07	0.51	33.2	10.9	0.08
LA CANDELARIA	0.47	35.4	11.2	0.12	0.47	36.2	11.0	0.09	0.50	38.5	12.2	0.09
LOS MARTIRES	0.50	34.1	10.6	0.09	0.50	35.2	10.2	0.07	0.50	37.2	11.6	0.09
PUENTE ARANDA	0.51	34.2	11.3	0.07	0.51	35.5	11.1	0.04	0.50	36.8	11.6	0.06
RAFAEL URIBE U	0.51	30.8	9.1	0.05	0.51	31.8	9.2	0.04	0.50	33.5	9.9	0.06
SAN CRISTOBAL	0.51	30.2	8.7	0.04	0.51	30.9	8.8	0.05	0.50	32.8	9.5	0.06
SANTA FE	0.50	32.2	10.3	0.10	0.50	33.5	9.2	0.05	0.50	35.8	10.5	0.06
SUBA	0.53	31.7	11.6	0.09	0.53	32.8	11.5	0.09	0.50	34.1	12.2	0.07
TEUSAQUILLO	0.54	38.0	14.4	0.13	0.53	39.1	14.7	0.09	0.50	38.8	15.3	0.11
TUNJUELITO	0.51	31.4	9.7	0.06	0.50	32.5	9.6	0.06	0.50	34.1	10.4	0.11
USAQUEN	0.54	34.5	13.3	0.08	0.54	36.3	13.0	0.06	0.50	36.7	13.9	0.13
USME	0.51	27.4	8.2	0.08	0.51	28.3	8.4	0.04	0.50	30.7	8.7	0.05
Total	0.52	31.5	10.5	0.08	0.52	32.5	10.5	0.07	0.50	34.0	11.2	0.08

b) T-test of difference between years

2014 vs 2011			2017 vs 2014			2017 vs 2014-2011					
Variable	Mean 2011	Mean 2014	Diff	Variable	Mean 2014	Mean 2017	Diff	Variable	Mean 2011-2014	Mean 2017	Diff
Fem.	0.518	0.517	-0.028	Fem.	0.517	0.505	-0.032	Fem.	0.517	0.505	-0.033
	(0.500)	(0.500)	(0.000)		(0.500)	(0.500)	(0.000)		(0.500)	(0.500)	(0.000)
Age	31.500	32.549	-5.846	Age	32.549	33.955	-5.760	Age	32.036	33.955	-3.969
	(20.243)	(20.622)	(0.000)		(20.622)	(20.079)	(0.000)		(20.444)	(20.079)	(0.000)
Edu-year	10.531	10.547	-4.718	Edu-year	10.547	11.243	-3.922	Edu-year	10.539	11.243	-4.251
	(4.681)	(4.665)	(0.000)		(4.665)	(4.596)	(0.000)		(4.673)	(4.596)	(0.000)
Mig.	0.084	0.067	-0.048	Mig.	0.067	0.076	0.015	Mig.	0.075	0.076	-0.006
	(0.277)	(0.250)	(0.000)		(0.250)	(0.264)	(0.000)		(0.264)	(0.264)	(0.000)

The standard error in parenthesis.

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

Note: calculated using the Multipurpose Survey for Bogotá (EMB) for 2011, 2014 and 2017. Part a) shows statistics for percentage of female (fem), average age, years of education and percentage of migrants in the locality are calculated using individual' weights. Migration is defined as migrating to the city five years ago or less (see previous part b) for reasons to migrate); this definition allowed to homogenize the variables for the different waves.

We do not see large changes in the socioeconomic variables between the different years. Migration is very low by locality and the percentage of women by locality does not change. Although there is an increasing trend in the years of education and age, this is a general aspect similar along all localities. By including yearxlocality fixed effects in our estimations, we expect that the general trend is controlled for already. In general, we do not observe drastic changes in these variables by locality that can be interpreted as evidence of a residential sorting problem. In addition, part b) of the table compares the difference in mean for the different years. The column Diff does not correspond exactly to the difference between the columns of the percentages of the years compared, as the test includes controls for stratus, yearxlocality fixed effects; the standard error in the tests are clustered at the locality level. The results of part b) show that there are no statistically significant differences in the average of the socioeconomic characteristics along the years. We do not have evidence to assert the presence of a residential sorting problem in terms of these variables, as they do not change drastically during the years of the sample.

2H Restricting sample to South-West²³ of the city

Table 2.24: IV SECOND STAGE: SO_2 (average) ON HEALTH

	SO2 (average)			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)
Niño-Q3	0.012 (0.010)	0.014** (0.006)	-0.076 (0.067)	-34.387*** (12.820)
Niño-Q2	0.004 (0.012)	0.002 (0.007)	-0.036 (0.081)	4.252 (12.864)
Niño-Q1	0.012 (0.012)	-0.004 (0.006)	-0.054 (0.086)	-8.761 (12.750)
Niña-Q3	0.019* (0.010)	0.011** (0.005)	-0.126* (0.066)	3.572 (11.265)
Niña-Q2	0.003 (0.009)	0.009 (0.006)	-0.037 (0.063)	-18.734 (12.498)
Niña-Q1	0.001 (0.009)	0.002 (0.007)	0.018 (0.059)	-11.534 (13.168)
so2_ave-Q3	-0.001 (0.001)	-0.000 (0.001)	0.007 (0.010)	-0.993 (2.079)
so2_ave-Q2	-0.003** (0.001)	-0.001 (0.001)	0.020** (0.009)	0.386 (2.276)
so2_ave-Q1	0.000 (0.002)	-0.000 (0.001)	-0.004 (0.011)	-1.648 (2.877)
Observations	3368	3368	3368	3368
NINO	0.03	0.01	-0.17	-38.90
P_NINO	0.02	0.07	0.04	0.01
NINA	0.02	0.02	-0.15	-26.70
P_NINA	0.09	0.03	0.13	0.19
POL	-0.00	-0.00	0.02	-2.26
P_POL	0.11	0.51	0.19	0.55
r2_a	0.40	0.30	0.40	0.35

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the IPS, Weeks gest. is the average of weeks of gestation for all the children born in that IPS and Birth Weight is the average weight for all the children born in that IPS. Pollution is averaged per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the three coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

²³Localities: Bosa, Ciudad Bolivar, Engativa, Fontibon, Kennedy, Puente Aranda and Tunjuelito

Table 2.25: IV SECOND STAGE: PM_{10} - $PM_{2.5}$ (average) ON HEALTH

	PM10 (average)				PM2.5 (average)			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)	Premature (5)	Low Weight (6)	Weeks Gest. (7)	Birth Weight (8)
Niño-Q3	0.011 (0.010)	0.016*** (0.006)	-0.074 (0.066)	-36.608*** (13.205)	0.019 (0.023)	0.043** (0.017)	-0.147 (0.164)	-41.666 (27.096)
Niño-Q2	0.008 (0.011)	0.004 (0.006)	-0.065 (0.074)	-9.663 (13.041)	-0.010 (0.018)	0.027** (0.012)	0.034 (0.127)	-41.416* (24.499)
Niño-Q1	0.019 (0.013)	-0.000 (0.006)	-0.102 (0.087)	-13.863 (14.661)	0.023 (0.015)	-0.002 (0.011)	-0.094 (0.101)	10.340 (19.806)
Niña-Q3	0.014 (0.010)	0.011** (0.005)	-0.101 (0.062)	-2.360 (10.073)	-0.032 (0.021)	-0.019 (0.018)	0.220 (0.147)	21.633 (25.696)
Niña-Q2	0.003 (0.010)	0.010 (0.007)	-0.039 (0.065)	-21.517 (14.001)	0.052*** (0.019)	0.041*** (0.016)	-0.327** (0.140)	-52.093* (26.752)
Niña-Q1	0.008 (0.010)	0.005 (0.007)	-0.036 (0.063)	-28.926** (12.875)	-0.002 (0.015)	-0.007 (0.013)	0.044 (0.102)	-14.910 (21.058)
pm10_ave-Q3	0.000 (0.001)	0.000 (0.001)	-0.004 (0.007)	0.075 (1.172)				
pm10_ave-Q2	0.001* (0.001)	0.001 (0.001)	-0.010** (0.005)	-2.339** (1.075)				
pm10_ave-Q1	0.001 (0.001)	0.001 (0.001)	-0.005 (0.007)	-0.917 (1.325)				
pm25_ave-Q3					0.007** (0.003)	0.004* (0.002)	-0.046** (0.018)	-6.202 (3.830)
pm25_ave-Q2					-0.004 (0.003)	0.006*** (0.002)	0.024 (0.023)	-6.507 (4.218)
pm25_ave-Q1					-0.000 (0.002)	0.003* (0.002)	0.002 (0.016)	-4.657 (3.193)
Observations	3368	3368	3368	3368	1456	1456	1456	1456
NINO	0.04	0.02	-0.24	-60.13	0.03	0.07	-0.21	-72.74
P_NINO	0.01	0.04	0.01	0.00	0.23	0.00	0.27	0.07
NINA	0.03	0.03	-0.18	-52.80	0.02	0.01	-0.06	-45.37
P_NINA	0.10	0.01	0.10	0.01	0.47	0.62	0.69	0.28
POL	0.00	0.00	-0.02	-3.18	0.00	0.01	-0.02	-17.37
P_POL	0.03	0.12	0.03	0.09	0.75	0.01	0.69	0.06
r2_a	0.39	0.29	0.39	0.34	0.41	0.29	0.42	0.32

* p<0.05>, ** p<0.01, *** p<0.001

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the IPS, Weeks gest. is the average of weeks of gestation for all the children born in that IPS and Birth Weight is the average weight for all the children born in that IPS. Pollution is averaged per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the three coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

Table 2.26: IV SECOND STAGE: $CO-NO_X$ (peak afternoon) ON HEALTH

	CO (peak afternoon)				NOX (peak afternoon)			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)	Premature (5)	Low Weight (6)	Weeks Gest. (7)	Birth Weight (8)
Niño-Q3	0.003 (0.010)	0.015** (0.007)	-0.012 (0.071)	-39.050*** (12.020)	-0.015* (0.009)	0.011 (0.008)	0.113* (0.062)	-23.697 (16.098)
Niño-Q2	-0.010 (0.013)	0.006 (0.008)	0.061 (0.090)	9.144 (14.269)	-0.004 (0.011)	0.002 (0.007)	0.011 (0.081)	6.606 (16.152)
Niño-Q1	0.016 (0.015)	0.002 (0.007)	-0.082 (0.101)	-31.199** (14.373)	0.005 (0.015)	0.006 (0.007)	-0.013 (0.105)	-18.386 (11.934)
Niña-Q3	0.011 (0.009)	0.009 (0.006)	-0.071 (0.062)	-3.618 (10.441)	0.009 (0.010)	0.005 (0.006)	-0.057 (0.072)	12.656 (11.782)
Niña-Q2	-0.006 (0.009)	0.005 (0.006)	0.030 (0.057)	-12.314 (10.747)	-0.019* (0.010)	-0.007 (0.006)	0.099 (0.070)	-17.138 (11.983)
Niña-Q1	-0.001 (0.011)	0.002 (0.007)	0.030 (0.072)	-9.442 (11.633)	0.002 (0.010)	-0.006 (0.010)	0.008 (0.068)	-11.240 (16.982)
co_peaka-Q3	0.010 (0.010)	0.004 (0.008)	-0.063 (0.067)	6.755 (14.682)				
co_peaka-Q2	-0.011 (0.009)	0.005 (0.007)	0.070 (0.063)	2.176 (12.795)				
co_peaka-Q1	-0.001 (0.006)	0.008 (0.005)	0.011 (0.042)	9.519 (8.947)				
nox_peaka-Q3					0.000 (0.001)	0.001 (0.000)	0.000 (0.004)	0.826 (0.796)
nox_peaka-Q2					-0.001** (0.000)	-0.000 (0.000)	0.005** (0.003)	0.158 (0.384)
nox_peaka-Q1					-0.001* (0.000)	0.001* (0.000)	0.004* (0.003)	-0.307 (0.517)
Observations	3056	3056	3056	3056	2256	2256	2256	2256
NINO	0.01	0.02	-0.03	-61.11	-0.01	0.02	0.11	-35.48
P_NINO	0.58	0.02	0.75	0.00	0.35	0.08	0.28	0.05
NINA	0.00	0.02	-0.01	-25.37	-0.01	-0.01	0.05	-15.72
P_NINA	0.81	0.14	0.92	0.16	0.65	0.49	0.68	0.52
POL	-0.00	0.02	0.02	18.45	-0.00	0.00	0.01	0.68
P_POL	0.92	0.32	0.90	0.54	0.13	0.16	0.11	0.61
r2_a	0.40	0.28	0.40	0.34	0.39	0.29	0.39	0.35

* p<0.05> , ** p<0.01, *** p<0.001

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the IPS, Weeks gest. is the average of weeks of gestation for all the children born in that IPS and Birth Weight is the average weight for all the children born in that IPS. Peak afternoon is the average pollution during the peak in the afternoon per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the three coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

Table 2.27: IV FIRST STAGE: IV ON SO_2 PM_{10} $PM_{2.5}$ (average)

	SO2 (average)			PM10 (average)			PM2.5 (average)		
	Q3-gest. (1)	Q2-gest. (2)	Q1-gest. (3)	Q3-gest. (4)	Q2-gest. (5)	Q1-gest. (6)	Q3-gest. (7)	Q2-gest. (8)	Q1-gest. (9)
Niño-Q3	1.152** (0.566)	1.527** (0.675)	-0.224 (0.574)	-3.337 (2.405)	0.319 (1.632)	-0.701 (1.709)	-1.074 (1.373)	-3.362*** (0.951)	-0.779 (1.387)
Niño-Q2	-1.132** (0.563)	1.564** (0.718)	2.315*** (0.594)	-0.830 (1.974)	-4.744** (2.098)	-0.493 (1.284)	2.606 (1.616)	-5.512*** (1.211)	1.094 (1.220)
Niño-Q1	0.339 (0.685)	-1.688** (0.731)	-0.495 (0.567)	-1.333 (1.735)	0.599 (2.428)	-4.485 (3.339)	-0.079 (1.209)	0.095 (1.180)	-0.382 (1.091)
Niña-Q3	1.051* (0.591)	0.103 (0.626)	0.754 (0.575)	-1.337 (2.012)	-2.473 (1.870)	-0.089 (2.234)	1.309 (1.579)	-3.595** (1.599)	7.505*** (1.632)
Niña-Q2	0.251 (0.645)	1.234* (0.683)	0.105 (0.813)	-4.576** (2.231)	1.247 (2.767)	-3.051 (2.688)	-3.987** (1.760)	-1.412 (1.623)	-5.766*** (2.094)
Niña-Q1	0.630 (0.651)	0.131 (0.695)	4.398*** (1.268)	-0.048 (2.007)	-5.659*** (1.793)	-0.498 (2.204)	0.500 (1.606)	-5.618*** (1.183)	4.543** (1.933)
ws_ave-Q3	0.511 (0.409)	-1.411*** (0.460)	-0.807* (0.414)	-2.352 (1.818)	1.163 (1.506)	0.818 (1.127)	0.297 (0.830)	-1.351* (0.789)	0.488 (0.771)
ws_ave-Q2	0.500 (0.396)	1.246** (0.524)	-0.854 (0.537)	0.959 (1.392)	-4.328** (1.842)	-0.177 (1.164)	-1.389** (0.664)	0.793 (0.760)	0.225 (0.714)
ws_ave-Q1	-0.732* (0.383)	0.094 (0.445)	1.447** (0.566)	3.576*** (0.932)	0.223 (1.215)	-4.690*** (1.495)	1.035 (0.808)	-2.787*** (0.920)	-0.592 (0.785)
wd_rose_Cat3-Q3	1.861 (3.530)	0.870 (2.460)	4.264** (2.103)	13.963** (6.668)	1.989 (6.594)	-16.041* (8.914)			
wd_rose_Cat4-Q3	0.708 (0.954)	-1.737 (1.078)	1.146 (0.968)	-2.816 (2.565)	-5.763** (2.798)	-4.918 (3.083)	-2.293* (1.197)	-1.933** (0.801)	0.941 (1.211)
wd_rose_Cat6-Q3	-3.605*** (0.773)	-1.372 (0.945)	-0.783 (1.140)	4.982** (1.987)	5.725** (2.474)	4.259 (2.929)	-0.684 (0.814)	2.784*** (0.874)	-1.301 (0.952)
wd_rose_Cat3-Q2	-13.164*** (2.908)	3.825 (3.539)	9.859*** (2.349)	0.222 (5.041)	9.831* (5.815)	-12.063 (7.524)			
wd_rose_Cat4-Q2	-1.362 (1.023)	0.382 (1.018)	-1.072 (1.108)	-2.218 (2.656)	-1.854 (2.279)	-7.067*** (2.671)	-1.337 (1.400)	-2.450*** (0.893)	-0.649 (0.924)
wd_rose_Cat6-Q2	1.670** (0.743)	-4.577*** (0.883)	-2.298** (1.142)	-2.790 (2.808)	3.583 (2.642)	5.547** (2.164)	-2.971*** (0.847)	-4.263*** (1.197)	2.354** (1.078)
wd_rose_Cat3-Q1	-12.722*** (3.234)	-9.459*** (3.569)	2.362 (2.370)	0.048 (5.643)	-0.353 (5.058)	7.526 (6.748)			
wd_rose_Cat4-Q1	0.677 (0.903)	-1.485* (0.776)	-1.695 (1.305)	-1.137 (2.774)	-2.599 (2.144)	-8.590*** (2.547)	-0.695 (1.156)	0.112 (1.211)	-3.139*** (0.843)
wd_rose_Cat6-Q1	1.780** (0.726)	0.334 (0.866)	-4.520*** (1.334)	2.284 (2.364)	-3.471 (2.874)	-0.447 (2.497)	2.513** (1.046)	-3.683*** (1.086)	-2.208** (0.876)
<i>N</i>	3368	3368	3368	3368	3368	3368	1456	1456	1456
R2_first_a	0.32	0.28	0.18	0.09	0.11	0.09	0.30	0.27	0.34
F_first_a	19.20	20.10	13.66	30.24	18.07	7.23	15.21	12.79	23.96

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Pollution is averaged per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. Each regression of the health outcomes has three equations in the first stage, one for each quarter of gestation. In this case, the three equations are the same for each set of four health outcomes presented in the second stage, for instance, the first four columns of Table 2.6 (for SO_2) share the same equations in the first stage, so we will only present once the results. The same holds for the other pollutants. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay). The interaction of the wind direction rose with wind speed not shown to save space. Wind direction baseline is category 5 here.

Table 2.28: IV FIRST STAGE: IV ON CO NO_X (peak afternoon)

	CO (peak afternoon)			NOX (peak afternoon)		
	Q3-gest. (1)	Q2-gest. (2)	Q1-gest. (3)	Q3-gest. (4)	Q2-gest. (5)	Q1-gest. (6)
Niño-Q3	0.053 (0.149)	-0.323** (0.156)	0.264* (0.151)	-4.972 (3.577)	-12.717*** (2.713)	-9.535*** (2.379)
Niño-Q2	-0.174 (0.167)	0.031 (0.134)	-0.349* (0.209)	3.227 (3.572)	-6.667*** (2.513)	-8.162*** (2.263)
Niño-Q1	0.075 (0.115)	-0.338 (0.225)	0.417** (0.164)	0.711 (3.991)	-1.785 (4.402)	-4.362 (3.013)
Niña-Q3	-0.246* (0.133)	-0.189* (0.111)	0.546*** (0.160)	0.845 (2.937)	-8.608*** (2.954)	9.646*** (3.006)
Niña-Q2	-0.014 (0.110)	-0.335* (0.173)	-0.024 (0.123)	5.463 (3.902)	1.025 (3.192)	-6.912** (3.271)
Niña-Q1	-0.481*** (0.151)	-0.178* (0.098)	-0.317** (0.152)	-5.343 (4.046)	8.855 (5.674)	8.349*** (3.165)
ws_peaka-Q3	0.009 (0.062)	0.098* (0.054)	0.003 (0.063)	0.014 (1.517)	0.515 (1.796)	-1.739 (1.444)
ws_peaka-Q2	0.119* (0.061)	-0.164*** (0.063)	0.016 (0.087)	-1.415 (2.019)	-0.022 (1.790)	3.194* (1.703)
ws_peaka-Q1	-0.148*** (0.052)	0.281*** (0.076)	0.088 (0.078)	-4.878*** (1.813)	-1.882 (2.124)	-2.616 (1.743)
wd_rose_Cat3-Q3	-1.827*** (0.564)	-2.487*** (0.660)	4.243*** (1.153)	62.925*** (22.513)	56.386** (22.235)	-21.126* (12.096)
wd_rose_Cat4-Q3	-0.137 (0.109)	0.224** (0.113)	0.086 (0.124)	-4.008 (4.093)	-5.142 (4.172)	-1.568 (3.036)
wd_rose_Cat6-Q3	0.034 (0.108)	0.104 (0.122)	0.448*** (0.124)	-0.840 (4.512)	1.768 (3.904)	5.318 (3.688)
wd_rose_Cat3-Q2	-1.864*** (0.384)	-2.097*** (0.591)	1.019** (0.517)	13.993 (14.104)	23.826 (19.294)	-3.511 (13.161)
wd_rose_Cat4-Q2	-0.186 (0.141)	-0.147 (0.145)	0.237** (0.112)	0.450 (3.821)	-7.309 (4.645)	-5.381 (3.812)
wd_rose_Cat6-Q2	0.298*** (0.081)	0.059 (0.087)	0.244** (0.112)	-1.096 (4.889)	-1.042 (4.567)	3.463 (3.477)
wd_rose_Cat3-Q1	-0.881** (0.360)	-0.555 (0.410)	-1.764** (0.706)	-32.026** (15.346)	-10.675 (19.127)	3.447 (11.540)
wd_rose_Cat4-Q1	-0.572*** (0.135)	0.124 (0.120)	0.127 (0.146)	-1.342 (5.039)	-1.608 (2.920)	-6.704** (3.209)
wd_rose_Cat6-Q1	0.205** (0.089)	0.148 (0.091)	-0.046 (0.107)	8.183* (4.486)	2.743 (3.422)	-8.318** (3.783)
<i>N</i>	3056	3056	3056	2256	2256	2256
R2_first_a	0.40	0.40	0.41	0.14	0.22	0.29
F_first_a	44.77	35.21	35.20	8.24	11.25	26.27

* p<0.05> , ** p<0.01, *** p<0.001

Note: Peak afternoon is the average pollution during the peak in the afternoon per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. Each regression of the health outcomes has three equations in the first stage, one for each quarter of gestation. In this case, the three equations are the same for each set of four health outcomes presented in the second stage, for instance, the first four columns of Table 2.6 (for SO₂) share the same equations in the first stage, so we will only present once the results. The same holds for the other pollutants. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay). The interaction of the wind direction rose with wind speed not shown to save space. Wind direction baseline is category 5 here.

2I Aggregating the sample by localities of the city

Table 2.29: IV SECOND STAGE: SO_2 (average) ON HEALTH

	SO2 (average)			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)
Niño-Q3	-0.003 (0.006)	0.001 (0.004)	0.020 (0.046)	-6.316 (7.837)
Niño-Q2	0.009 (0.007)	0.001 (0.004)	-0.065 (0.049)	-3.987 (8.292)
Niño-Q1	-0.010 (0.006)	-0.001 (0.003)	0.072* (0.043)	-5.379 (7.103)
Niña-Q3	-0.000 (0.006)	0.005 (0.004)	-0.000 (0.040)	-3.780 (8.006)
Niña-Q2	0.005 (0.006)	0.009** (0.004)	-0.036 (0.038)	-7.791 (7.553)
Niña-Q1	-0.013* (0.008)	0.003 (0.004)	0.091* (0.054)	-8.023 (8.867)
so2_ave-Q3	-0.000 (0.001)	-0.002** (0.001)	-0.001 (0.009)	3.328* (1.851)
so2_ave-Q2	-0.005*** (0.001)	-0.001 (0.001)	0.035*** (0.009)	-1.440 (1.756)
so2_ave-Q1	0.003 (0.002)	0.002* (0.001)	-0.018 (0.013)	0.721 (2.073)
Observations	3347	3347	3347	3347
NINO	-0.00	0.00	0.03	-15.68
P_NINO	0.55	0.75	0.55	0.03
NINA	-0.01	0.02	0.06	-19.59
P_NINA	0.41	0.00	0.43	0.10
POL	-0.00	-0.00	0.02	2.61
P_POL	0.28	0.37	0.30	0.32
r2_a	0.52	0.44	0.53	0.51

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the locality, Weeks gest. is the average of weeks of gestation for all the children born in that locality and Birth Weight is the average weight for all the children born in that locality. Pollution is averaged per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the three coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

Table 2.30: IV SECOND STAGE: PM_{10} - $PM_{2.5}$ (average) ON HEALTH

	PM10 (average)				PM2.5 (average)			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)	Premature (5)	Low Weight (6)	Weeks Gest. (7)	Birth Weight (8)
Niño-Q3	-0.006 (0.005)	-0.003 (0.004)	0.037 (0.037)	-8.357 (7.733)	-0.001 (0.012)	-0.004 (0.008)	0.002 (0.084)	6.652 (10.717)
Niño-Q2	0.012* (0.007)	0.007* (0.004)	-0.083* (0.048)	-15.269* (8.318)	0.023** (0.010)	0.018*** (0.007)	-0.169** (0.067)	-20.524** (8.837)
Niño-Q1	-0.007 (0.006)	-0.000 (0.004)	0.049 (0.043)	-5.244 (8.674)	-0.007 (0.009)	-0.006 (0.005)	0.042 (0.065)	5.572 (8.361)
Niña-Q3	-0.002 (0.004)	0.001 (0.003)	0.008 (0.031)	2.291 (6.683)	-0.001 (0.011)	-0.002 (0.008)	0.004 (0.077)	-0.888 (12.536)
Niña-Q2	0.003 (0.005)	0.008** (0.004)	-0.023 (0.035)	-16.482** (7.636)	0.007 (0.008)	0.017*** (0.007)	-0.045 (0.055)	-14.870 (9.480)
Niña-Q1	-0.007 (0.008)	0.006 (0.005)	0.053 (0.054)	-15.768 (11.224)	0.011 (0.009)	0.008 (0.006)	-0.096 (0.066)	-14.155 (11.886)
pm10_ave-Q3	0.001 (0.001)	0.000 (0.000)	-0.005 (0.004)	-1.043* (0.604)				
pm10_ave-Q2	0.000 (0.001)	0.000 (0.000)	-0.000 (0.004)	-1.131 (0.834)				
pm10_ave-Q1	-0.000 (0.001)	0.001 (0.001)	0.001 (0.005)	-1.318 (1.105)				
pm25_ave-Q3					0.002 (0.001)	0.001 (0.001)	-0.015 (0.010)	-2.516 (2.116)
pm25_ave-Q2					0.000 (0.002)	0.001 (0.001)	-0.007 (0.013)	1.517 (1.841)
pm25_ave-Q1					0.001 (0.001)	0.001 (0.001)	-0.009 (0.008)	0.121 (1.093)
Observations	3347	3347	3347	3347	1471	1471	1471	1471
NINO	-0.00	0.00	0.00	-28.87	0.01	0.01	-0.12	-8.30
P_NINO	0.98	0.44	0.93	0.01	0.27	0.53	0.20	0.61
NINA	-0.01	0.01	0.04	-29.96	0.02	0.02	-0.14	-29.91
P_NINA	0.59	0.02	0.60	0.03	0.31	0.03	0.24	0.11
POL	0.00	0.00	-0.00	-3.49	0.00	0.00	-0.03	-0.88
P_POL	0.39	0.41	0.45	0.02	0.33	0.25	0.21	0.85
r2_a	0.53	0.45	0.54	0.50	0.59	0.44	0.59	0.54

* p<0.05> , ** p<0.01, *** p<0.001

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the locality, Weeks gest. is the average of weeks of gestation for all the children born in that locality and Birth Weight is the average weight for all the children born in that locality. Pollution is averaged per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the three coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

Table 2.31: IV SECOND STAGE: $CO-NO_X$ (peak afternoon) ON HEALTH

	CO (peak afternoon)				NOX (peak afternoon)			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)	Premature (5)	Low Weight (6)	Weeks Gest. (7)	Birth Weight (8)
Niño-Q3	-0.010** (0.005)	-0.003 (0.004)	0.069** (0.032)	-7.410 (8.554)	-0.009 (0.006)	0.001 (0.003)	0.064 (0.041)	-14.411* (8.644)
Niño-Q2	0.016** (0.007)	0.011** (0.004)	-0.115** (0.047)	-6.657 (9.572)	0.005 (0.007)	0.003 (0.004)	-0.036 (0.047)	-0.817 (8.665)
Niño-Q1	-0.003 (0.008)	-0.002 (0.005)	0.027 (0.055)	-8.812 (7.905)	-0.014** (0.007)	-0.004 (0.004)	0.107** (0.050)	-5.646 (9.697)
Niña-Q3	0.000 (0.005)	0.005 (0.004)	-0.008 (0.033)	-7.411 (6.602)	0.000 (0.005)	-0.002 (0.003)	-0.013 (0.038)	7.873 (9.777)
Niña-Q2	-0.002 (0.005)	0.004 (0.003)	0.015 (0.034)	-3.472 (6.033)	-0.011* (0.006)	-0.005* (0.003)	0.074* (0.039)	13.714* (8.279)
Niña-Q1	0.004 (0.007)	0.013*** (0.004)	-0.023 (0.051)	-14.359* (7.336)	-0.008 (0.007)	0.004 (0.003)	0.058 (0.051)	-4.987 (8.166)
co_peaka-Q3	0.025*** (0.009)	0.010* (0.005)	-0.173*** (0.062)	-7.119 (9.476)				
co_peaka-Q2	0.005 (0.007)	0.009** (0.004)	-0.029 (0.049)	2.362 (8.368)				
co_peaka-Q1	-0.003 (0.005)	0.002 (0.003)	0.016 (0.034)	8.260 (5.964)				
nox_peaka-Q3					-0.000 (0.000)	0.000** (0.000)	0.003 (0.002)	-1.622*** (0.601)
nox_peaka-Q2					0.000 (0.000)	0.000 (0.000)	0.000 (0.002)	-0.156 (0.377)
nox_peaka-Q1					-0.001** (0.000)	0.000 (0.000)	0.006** (0.002)	-0.727* (0.428)
Observations	3027	3027	3027	3027	2348	2348	2348	2348
NINO	0.00	0.01	-0.02	-22.88	-0.02	0.00	0.13	-20.87
P_NINO	0.77	0.28	0.75	0.02	0.08	0.92	0.06	0.13
NINA	0.00	0.02	-0.02	-25.24	-0.02	-0.00	0.12	16.60
P_NINA	0.87	0.00	0.82	0.01	0.09	0.61	0.11	0.21
POL	0.03	0.02	-0.19	3.50	-0.00	0.00	0.01	-2.50
P_POL	0.08	0.05	0.09	0.87	0.08	0.14	0.06	0.02
r2_a	0.57	0.46	0.58	0.51	0.55	0.45	0.56	0.53

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the locality, Weeks gest. is the average of weeks of gestation for all the children born in that locality and Birth Weight is the average weight for all the children born in that locality. Pollution is averaged per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the three coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

Table 2.32: IV FIRST STAGE: IV ON SO_2 PM_{10} $PM_{2.5}$ (average)

	SO2 (average)			PM10 (average)			PM2.5 (average)		
	Q3-gest. (1)	Q2-gest. (2)	Q1-gest. (3)	Q3-gest. (4)	Q2-gest. (5)	Q1-gest. (6)	Q3-gest. (7)	Q2-gest. (8)	Q1-gest. (9)
Niño-Q3	0.288 (0.404)	1.351** (0.591)	0.038 (0.389)	-3.562 (2.225)	0.302 (1.553)	1.075 (1.653)	0.966 (1.985)	-3.415*** (1.050)	-1.173 (1.774)
Niño-Q2	-1.392*** (0.445)	0.442 (0.599)	1.731*** (0.481)	-2.424 (1.872)	-5.013*** (1.588)	-1.132 (1.008)	2.159 (1.599)	-4.175*** (1.097)	2.438* (1.336)
Niño-Q1	0.280 (0.497)	-1.345** (0.578)	-0.694 (0.512)	1.735 (1.669)	0.255 (2.088)	-3.838 (2.610)	-0.992 (1.343)	-0.046 (1.250)	1.242 (1.343)
Niña-Q3	0.511 (0.437)	-0.164 (0.556)	0.461 (0.404)	-1.506 (2.207)	-0.418 (1.869)	-0.797 (2.062)	0.699 (1.730)	-2.433 (2.100)	9.620*** (1.864)
Niña-Q2	0.109 (0.575)	0.583 (0.583)	-0.119 (0.618)	-4.290** (2.020)	-0.978 (3.204)	-1.375 (2.115)	-2.110 (2.251)	-0.984 (1.781)	-5.136** (2.433)
Niña-Q1	1.023** (0.435)	0.037 (0.632)	1.996*** (0.677)	-1.431 (2.051)	-5.474*** (1.849)	-2.430 (1.923)	1.158 (2.000)	-6.039*** (1.468)	4.966** (2.384)
ws_ave-Q3	0.687 (1.232)	0.756 (0.857)	-3.212*** (0.717)	-4.047 (2.567)	-4.473** (2.264)	3.330 (2.700)	3.249 (6.230)	1.177 (6.004)	-9.189 (6.744)
ws_ave-Q2	-0.249 (0.998)	-0.601 (1.043)	-0.952 (0.855)	3.938** (1.983)	0.706 (2.494)	1.091 (1.963)	-5.795 (5.498)	-5.449 (4.853)	26.457*** (9.796)
ws_ave-Q1	-0.329 (0.650)	-0.738 (0.810)	-1.852* (0.982)	5.723*** (1.850)	3.424 (2.110)	0.597 (2.479)	-3.142 (6.112)	-7.698 (6.842)	-2.071 (4.933)
wd_rose_Cat4-Q3	-2.019 (1.554)	0.949 (1.238)	-4.254*** (1.273)	-4.906 (3.447)	-6.207 (3.815)	3.923 (5.531)	2.585 (10.366)	1.090 (10.179)	-10.106 (11.332)
wd_rose_Cat5-Q3	-2.092 (1.504)	1.446 (1.098)	-4.426*** (1.318)	-0.472 (3.446)	-5.701 (3.573)	4.104 (5.223)	3.953 (10.567)	1.622 (10.051)	-12.171 (11.042)
wd_rose_Cat6-Q3	-3.399* (1.778)	1.906 (1.472)	-3.572** (1.797)	3.784 (3.871)	-0.034 (4.278)	9.796* (5.656)	3.040 (10.454)	2.438 (9.927)	-11.790 (11.350)
wd_rose_Cat4-Q2	-2.442** (1.144)	-3.318*** (0.783)	0.469 (0.934)	6.928** (2.791)	5.310** (2.200)	-1.306 (2.878)	-5.234 (8.329)	-12.173 (8.508)	44.698*** (16.414)
wd_rose_Cat5-Q2	-1.609 (1.210)	-2.964*** (0.824)	0.496 (1.017)	9.474*** (2.878)	8.507*** (2.529)	-2.280 (3.054)	-7.420 (8.559)	-10.462 (8.221)	44.219*** (16.504)
wd_rose_Cat6-Q2	-0.848 (1.500)	-4.141*** (1.304)	0.514 (1.468)	3.649 (3.903)	12.407*** (3.252)	4.708 (4.076)	-10.033 (8.703)	-12.175 (8.309)	46.316*** (17.131)
wd_rose_Cat4-Q1	0.267 (0.809)	-2.760*** (0.902)	-5.008*** (1.138)	4.639* (2.458)	3.139 (1.994)	3.524* (1.965)	-5.987 (10.243)	-7.323 (12.063)	-4.045 (7.746)
wd_rose_Cat5-Q1	0.032 (1.015)	-1.129 (0.919)	-4.853*** (0.994)	5.821** (2.633)	2.859 (2.132)	4.660** (2.131)	-5.345 (10.081)	-9.254 (11.956)	-1.591 (7.543)
wd_rose_Cat6-Q1	0.264 (1.290)	-0.133 (1.303)	-6.136*** (1.273)	6.964** (3.000)	-1.475 (2.778)	8.300** (3.230)	-2.887 (9.925)	-12.637 (11.629)	-1.567 (7.541)
<i>N</i>	3347	3347	3347	3347	3347	3347	1471	1471	1471
R2_first_a	0.17	0.18	0.10	0.15	0.13	0.08	0.22	0.24	0.38
F_first_a	29.00	15.88	14.09	17.10	10.39	7.82	6.67	20.44	33.22

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Pollution is averaged per day, then averaged by month and by quarter of gestation for each locality. The estimations include yearlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. Each regression of the health outcomes has three equations in the first stage, one for each quarter of gestation. In this case, the three equations are the same for each set of four health outcomes presented in the second stage, for instance, the first four columns of Table 2.6 (for SO_2) share the same equations in the first stage, so we will only present once the results. The same holds for the other pollutants. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay). The interaction of the wind direction rose with wind speed not shown to save space.

Table 2.33: IV FIRST STAGE: IV ON CO NO_X (peak afternoon)

	CO (peak afternoon)			NOX (peak afternoon)		
	Q3-gest. (1)	Q2-gest. (2)	Q1-gest. (3)	Q3-gest. (4)	Q2-gest. (5)	Q1-gest. (6)
Niño-Q3	0.087 (0.089)	-0.405*** (0.114)	0.163 (0.110)	-3.699 (3.191)	-5.968*** (2.301)	-3.560* (2.129)
Niño-Q2	-0.134 (0.131)	0.052 (0.116)	-0.413*** (0.154)	1.669 (3.503)	-4.340* (2.279)	-5.311*** (1.888)
Niño-Q1	0.004 (0.089)	-0.301 (0.249)	0.323** (0.135)	0.153 (3.752)	-1.405 (3.678)	-5.738** (2.775)
Niña-Q3	-0.135 (0.096)	-0.250** (0.100)	0.399*** (0.126)	0.810 (2.337)	-6.107*** (2.265)	8.619*** (2.802)
Niña-Q2	-0.080 (0.080)	-0.245* (0.143)	-0.054 (0.114)	4.515 (3.092)	-0.399 (2.303)	-2.656 (3.233)
Niña-Q1	-0.490*** (0.123)	-0.190* (0.102)	-0.175* (0.106)	-5.516* (2.989)	6.474 (3.980)	4.889* (2.635)
ws_peaka-Q3	-0.645*** (0.115)	0.790*** (0.177)	0.181* (0.105)	-2.729 (2.677)	-3.711 (3.874)	9.794*** (2.822)
ws_peaka-Q2	-0.232 (0.158)	-0.765*** (0.161)	0.406*** (0.099)	-1.195 (1.918)	0.195 (1.891)	-1.981 (1.766)
ws_peaka-Q1	0.052 (0.145)	-0.234 (0.156)	-0.068 (0.150)	-5.959*** (2.096)	-2.640 (1.872)	-3.232* (1.779)
wd_rose_Cat4-Q3	-1.136*** (0.371)	0.436 (0.483)	0.401 (0.273)	15.043** (5.894)	-0.208 (7.328)	11.170** (5.555)
wd_rose_Cat5-Q3	-1.237*** (0.370)	0.438 (0.485)	0.265 (0.289)	20.918*** (5.790)	-2.573 (7.736)	8.832 (5.918)
wd_rose_Cat6-Q3	-1.230*** (0.400)	0.361 (0.477)	0.346 (0.335)	23.612*** (6.758)	1.219 (9.083)	11.775 (7.350)
wd_rose_Cat4-Q2	-0.860*** (0.317)	-1.398*** (0.303)	0.824*** (0.127)	-2.353 (2.998)	9.139** (3.640)	-6.713*** (2.468)
wd_rose_Cat5-Q2	-0.946*** (0.342)	-1.731*** (0.325)	0.937*** (0.155)	6.092* (3.507)	14.131*** (4.270)	-10.310*** (3.247)
wd_rose_Cat6-Q2	-0.898*** (0.343)	-1.648*** (0.357)	1.068*** (0.180)	6.014 (4.559)	12.196** (5.494)	-12.500*** (4.590)
wd_rose_Cat4-Q1	0.083 (0.303)	-1.011*** (0.320)	-0.488*** (0.165)	-4.704* (2.818)	-2.061 (2.705)	2.158 (1.704)
wd_rose_Cat5-Q1	0.261 (0.296)	-1.139*** (0.347)	-0.739*** (0.190)	-2.110 (2.752)	0.658 (2.617)	3.196 (2.335)
wd_rose_Cat6-Q1	0.186 (0.297)	-1.081*** (0.382)	-0.957*** (0.231)	-1.401 (4.278)	3.223 (4.218)	-1.309 (3.874)
<i>N</i>	3027	3027	3027	2348	2348	2348
R2_first_a	0.30	0.35	0.39	0.19	0.24	0.29
F_first_a	28.30	31.67	28.29	9.45	15.95	35.87

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Peak afternoon is the average pollution during the peak in the afternoon per day, then averaged by month and by quarter of gestation for each locality. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. Each regression of the health outcomes has three equations in the first stage, one for each quarter of gestation. In this case, the three equations are the same for each set of four health outcomes presented in the second stage, for instance, the first four columns of Table 2.6 (for SO_2) share the same equations in the first stage, so we will only present once the results. The same holds for the other pollutants. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay). The interaction of the wind direction rose with wind speed not shown to save space.

2J Changing buffer to 4km

Table 2.34: IV SECOND STAGE: SO_2 (average) ON HEALTH

	SO2 (average)			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)
Niño-Q3	0.010 (0.006)	0.013*** (0.005)	-0.075 (0.050)	-29.372*** (8.261)
Niño-Q2	-0.007 (0.007)	-0.012** (0.005)	0.048 (0.055)	14.154* (8.412)
Niño-Q1	0.007 (0.006)	0.011** (0.005)	-0.019 (0.046)	-13.822 (8.715)
Niña-Q3	0.001 (0.006)	0.005 (0.004)	0.014 (0.046)	4.327 (7.439)
Niña-Q2	-0.003 (0.006)	0.011*** (0.004)	0.014 (0.045)	-17.118** (7.359)
Niña-Q1	-0.014* (0.008)	-0.007 (0.006)	0.120** (0.060)	10.680 (10.454)
so2_ave-Q3	0.000 (0.001)	-0.001 (0.001)	-0.002 (0.012)	-1.038 (2.378)
so2_ave-Q2	-0.003 (0.002)	0.001 (0.002)	0.020 (0.014)	-1.639 (3.190)
so2_ave-Q1	0.005** (0.002)	0.004** (0.002)	-0.038** (0.018)	-5.472* (3.209)
Observations	11112	11112	11112	11112
NINO	0.01	0.01	-0.05	-29.04
P_NINO	0.09	0.02	0.34	0.00
NINA	-0.02	0.01	0.15	-2.11
P_NINA	0.14	0.21	0.05	0.86
POL	0.00	0.00	-0.02	-8.15
P_POL	0.25	0.02	0.26	0.01
r2_a	0.36	0.30	0.35	0.34

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the IPS, Weeks gest. is the average of weeks of gestation for all the children born in that IPS and Birth Weight is the average weight for all the children born in that IPS. Pollution is averaged per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the three coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

Table 2.35: IV SECOND STAGE: PM_{10} - $PM_{2.5}$ (average) ON HEALTH

	PM10 (average)				PM2.5 (average)			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)	Premature (5)	Low Weight (6)	Weeks Gest. (7)	Birth Weight (8)
Niño-Q3	0.004 (0.006)	0.004 (0.005)	-0.036 (0.042)	-21.491*** (7.954)	0.011 (0.010)	0.003 (0.006)	-0.067 (0.081)	-12.562 (12.824)
Niño-Q2	0.001 (0.006)	-0.004 (0.005)	-0.007 (0.045)	2.137 (7.711)	-0.000 (0.009)	-0.001 (0.006)	0.001 (0.072)	-4.805 (10.498)
Niño-Q1	0.007 (0.006)	0.011* (0.006)	-0.018 (0.048)	-11.534 (9.358)	0.020** (0.010)	0.010 (0.007)	-0.113 (0.072)	-13.682 (11.527)
Niña-Q3	0.000 (0.006)	0.000 (0.005)	0.022 (0.043)	6.378 (8.592)	-0.030** (0.015)	-0.021** (0.009)	0.246** (0.118)	21.186 (18.497)
Niña-Q2	-0.003 (0.007)	0.007 (0.006)	0.018 (0.047)	-17.617* (10.329)	0.035*** (0.011)	0.035*** (0.007)	-0.227** (0.089)	-42.680*** (11.722)
Niña-Q1	-0.005 (0.007)	-0.002 (0.006)	0.058 (0.057)	0.081 (10.518)	-0.004 (0.012)	-0.005 (0.007)	0.037 (0.091)	5.190 (11.693)
pm10_ave-Q3	-0.000 (0.001)	-0.001*** (0.000)	0.000 (0.005)	0.826 (0.792)				
pm10_ave-Q2	0.001 (0.001)	0.001 (0.001)	-0.004 (0.005)	-1.459 (0.965)				
pm10_ave-Q1	-0.000 (0.001)	0.001 (0.001)	0.001 (0.006)	-0.496 (1.181)				
pm25_ave-Q3					0.005*** (0.002)	0.001 (0.001)	-0.040*** (0.014)	-2.491 (1.976)
pm25_ave-Q2					-0.002 (0.002)	0.002* (0.001)	0.010 (0.014)	-3.018 (2.541)
pm25_ave-Q1					0.001 (0.001)	0.001 (0.001)	-0.003 (0.008)	-0.529 (1.378)
Observations	11112	11112	11112	11112	4658	4658	4658	4658
NINO	0.01	0.01	-0.06	-30.89	0.03	0.01	-0.18	-31.05
P_NINO	0.06	0.10	0.22	0.01	0.06	0.21	0.16	0.08
NINA	-0.01	0.01	0.10	-11.16	0.00	0.01	0.06	-16.30
P_NINA	0.41	0.52	0.20	0.39	0.99	0.44	0.70	0.38
POL	0.00	0.00	-0.00	-1.13	0.00	0.00	-0.03	-6.04
P_POL	0.63	0.73	0.74	0.44	0.31	0.07	0.27	0.18
r2_a	0.37	0.30	0.35	0.34	0.35	0.30	0.34	0.33

* p<0.05>, ** p<0.01, *** p<0.001

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the IPS, Weeks gest. is the average of weeks of gestation for all the children born in that IPS and Birth Weight is the average weight for all the children born in that IPS. Pollution is averaged per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the three coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

Table 2.36: IV SECOND STAGE: $CO-NO_X$ (peak afternoon) ON HEALTH

	CO (peak afternoon)				NOX (peak afternoon)			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)	Premature (5)	Low Weight (6)	Weeks Gest. (7)	Birth Weight (8)
Niño-Q3	0.002 (0.006)	0.012** (0.005)	-0.019 (0.043)	-24.052*** (7.971)	0.000 (0.006)	0.007 (0.006)	-0.014 (0.047)	-24.999*** (9.080)
Niño-Q2	-0.005 (0.008)	-0.003 (0.005)	0.030 (0.064)	10.653 (8.895)	-0.002 (0.007)	-0.009* (0.005)	0.023 (0.051)	10.103 (8.986)
Niño-Q1	0.006 (0.007)	0.013** (0.005)	-0.015 (0.050)	-15.238* (9.216)	0.007 (0.006)	0.015*** (0.005)	-0.011 (0.048)	-19.045** (8.565)
Niña-Q3	0.001 (0.006)	0.007* (0.004)	0.019 (0.042)	0.496 (8.046)	-0.004 (0.007)	0.001 (0.005)	0.060 (0.048)	12.673 (9.276)
Niña-Q2	-0.003 (0.005)	0.012*** (0.004)	0.017 (0.038)	-18.348** (7.876)	-0.008 (0.007)	0.010** (0.005)	0.045 (0.054)	-16.426* (9.393)
Niña-Q1	-0.004 (0.007)	0.000 (0.005)	0.046 (0.056)	6.581 (9.692)	-0.009 (0.007)	-0.008 (0.007)	0.088 (0.055)	10.784 (9.760)
co_peaka-Q3	0.010 (0.007)	0.010 (0.006)	-0.074 (0.054)	3.570 (12.833)				
co_peaka-Q2	-0.007 (0.006)	0.005 (0.005)	0.044 (0.044)	12.911 (10.361)				
co_peaka-Q1	-0.003 (0.005)	0.001 (0.004)	0.013 (0.037)	9.787 (8.309)				
nox_peaka-Q3					0.000 (0.000)	0.000 (0.000)	0.000 (0.003)	-0.432 (0.506)
nox_peaka-Q2					0.000 (0.000)	-0.000 (0.000)	-0.000 (0.002)	-0.239 (0.472)
nox_peaka-Q1					-0.000 (0.000)	0.000 (0.000)	0.003 (0.002)	-0.504 (0.421)
Observations	9965	9965	9965	9965	7879	7879	7879	7879
NINO	0.00	0.02	-0.00	-28.64	0.00	0.01	-0.00	-33.94
P_NINO	0.69	0.00	0.95	0.01	0.54	0.08	0.98	0.01
NINA	-0.01	0.02	0.08	-11.27	-0.02	0.00	0.19	7.03
P_NINA	0.49	0.00	0.25	0.36	0.06	0.80	0.02	0.61
POL	-0.00	0.02	-0.02	26.27	-0.00	0.00	0.00	-1.18
P_POL	0.98	0.15	0.87	0.26	0.72	0.69	0.69	0.33
r2_a	0.37	0.30	0.35	0.34	0.37	0.32	0.36	0.37

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the IPS, Weeks gest. is the average of weeks of gestation for all the children born in that IPS and Birth Weight is the average weight for all the children born in that IPS. Peak afternoon is the average pollution during the peak in the afternoon per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the three coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

2K OLS estimations of pollution on Health - NO instruments

Table 2.37: OLS: SO_2 - PM_{10} (average) CO - NOX (peak afternoon) ON HEALTH

	SO2 (average)				PM10 (average)				CO (peak afternoon)				NOX (peak afternoon)			
	Premature	Low Weight	Weeks Gest.	Birth Weight	Premature	Low Weight	Weeks Gest.	Birth Weight	Premature	Low Weight	Weeks Gest.	Birth Weight	Premature	Low Weight	Weeks Gest.	Birth Weight
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Niño-Q3	0.005 (0.005)	0.010** (0.005)	-0.045 (0.039)	-26.713*** (7.896)	0.065 (0.065)	0.069* (0.005)	-0.037 (0.039)	-24.881*** (7.728)	0.003 (0.006)	0.009 (0.003)	-0.025 (0.045)	-22.953*** (8.205)	0.002 (0.006)	0.008 (0.005)	-0.029 (0.043)	-23.299** (9.161)
Niño-Q2	-0.003 (0.006)	-0.007 (0.005)	0.023 (0.046)	10.869 (7.653)	-0.003 (0.006)	-0.007 (0.005)	0.026 (0.045)	9.259 (7.885)	-0.009 (0.008)	-0.003 (0.005)	0.063 (0.064)	9.312 (8.989)	-0.004 (0.007)	-0.009* (0.005)	0.031 (0.050)	11.339 (9.130)
Niño-Q1	0.007 (0.006)	0.008 (0.005)	-0.017 (0.044)	-10.628 (8.324)	0.007 (0.006)	0.008 (0.005)	-0.014 (0.043)	-9.542 (8.385)	0.007 (0.007)	0.013*** (0.005)	-0.017 (0.050)	-15.221* (8.313)	0.009 (0.006)	0.018*** (0.005)	-0.027 (0.048)	-19.452** (8.584)
Niña-Q3	0.002 (0.006)	0.005 (0.004)	0.005 (0.042)	2.086 (7.868)	0.002 (0.006)	0.003 (0.004)	0.012 (0.042)	3.484 (8.107)	0.004 (0.006)	0.005 (0.004)	0.001 (0.043)	-0.225 (8.531)	-0.004 (0.007)	0.000 (0.005)	0.061 (0.050)	12.916 (9.774)
Niña-Q2	-0.003 (0.006)	0.013*** (0.004)	0.012 (0.043)	-20.921** (8.412)	-0.003 (0.006)	0.013*** (0.005)	0.018 (0.039)	-21.455** (8.462)	-0.006 (0.005)	0.012*** (0.004)	0.032 (0.038)	-19.912** (8.251)	-0.010 (0.006)	0.011** (0.005)	0.057 (0.046)	-17.238* (9.866)
Niña-Q1	-0.007 (0.007)	-0.003 (0.006)	0.068 (0.053)	4.600 (9.034)	-0.009 (0.007)	-0.005 (0.006)	0.081 (0.059)	4.901 (9.302)	-0.006 (0.008)	-0.001 (0.005)	0.068 (0.062)	2.763 (9.454)	-0.009 (0.007)	-0.007 (0.007)	0.088 (0.058)	7.521 (11.157)
so2_ave-Q3	-0.001 (0.001)	-0.001** (0.000)	0.007 (0.004)	0.901 (0.737)												
so2_ave-Q2	-0.000 (0.001)	0.001 (0.000)	0.001 (0.004)	-0.692 (0.788)												
so2_ave-Q1	0.000 (0.001)	0.000 (0.000)	-0.005 (0.004)	-1.334 (0.925)												
pm10_ave-Q3						-0.000 (0.000)	0.000 (0.001)	-0.103 (0.274)								
pm10_ave-Q2					-0.000* (0.000)	-0.000 (0.000)	0.003* (0.001)	0.242 (0.238)								
pm10_ave-Q1					-0.000 (0.000)	-0.000 (0.000)	0.001 (0.002)	-0.169 (0.258)								
co_ave-Q3									0.005 (0.004)	0.003 (0.003)	-0.022 (0.033)	-6.384 (5.016)	0.000 (0.000)	0.000 (0.000)	-0.001 (0.001)	-0.328 (0.203)
co_ave-Q2									-0.003 (0.003)	-0.003 (0.002)	0.020 (0.020)	8.671** (4.345)	-0.000 (0.000)	-0.000 (0.000)	0.000 (0.001)	0.036 (0.183)
co_ave-Q1									-0.006* (0.004)	-0.001 (0.002)	0.035 (0.027)	6.855 (4.222)	0.000 (0.000)	0.000 (0.000)	-0.001 (0.001)	-0.515*** (0.200)
nox_ave-Q3																
nox_ave-Q2																
nox_ave-Q1																
Observations	11027	11027	11027	11027	11027	11027	11027	11027	9833	9833	9833	9833	7683	7683	7683	7683
Niño	0.01	0.01	-0.04	-26.47	0.01	0.01	-0.03	-25.16	0.00	0.02	0.02	-25.86	0.01	0.02	-0.02	-31.41
P_Niño	0.14	0.03	0.41	0.00	0.21	0.05	0.39	0.01	0.95	0.00	0.77	0.01	0.36	0.00	0.65	0.00
Niña	-0.01	0.01	0.09	-14.24	-0.01	0.01	0.11	-13.07	-0.01	0.02	-0.10	-17.37	-0.02	0.00	0.21	3.20
P_Niña	0.44	0.06	0.23	0.15	0.33	0.17	0.17	0.26	0.43	0.02	0.23	0.10	0.04	0.75	0.01	0.81
POL	-0.00	-0.00	0.00	-1.13	-0.00	-0.00	0.00	-0.03	-0.00	-0.00	0.03	9.14	0.00	0.00	-0.00	-0.81
P_POL	0.49	0.95	0.66	0.45	0.26	0.36	0.18	0.94	0.57	0.81	0.55	0.34	0.33	0.13	0.16	0.03
r2_a	0.37	0.31	0.35	0.34	0.37	0.31	0.35	0.34	0.37	0.30	0.35	0.34	0.38	0.33	0.36	0.37

* p<0.05, ** p<0.01, *** p<0.001

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the locality, Weeks gest. is the average of weeks of gestation for all the children born in that locality and Birth Weight is the average weight for each IPS. The estimations include yearlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *Niño*, *Niña* and *POL* correspond to the sum of the three coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_Niño*, *P_Niña* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

2L Heterogeneity: mothers with tertiary education

Table 2.38: IV SECOND STAGE: SO_2 (average) ON HEALTH

	SO2 (average)			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)
Niño-Q3	0.008 (0.007)	0.015*** (0.004)	-0.057 (0.051)	-30.405*** (8.749)
Niño-Q2	-0.004 (0.008)	-0.011** (0.005)	0.043 (0.060)	22.256** (9.628)
Niño-Q1	0.004 (0.006)	0.008 (0.005)	-0.005 (0.050)	-11.267 (9.598)
Niño-EDU-Q3	0.011 (0.024)	-0.007 (0.019)	-0.107 (0.184)	-0.840 (33.094)
Niño-EDU-Q2	-0.012 (0.026)	-0.009 (0.020)	0.053 (0.188)	-8.107 (36.652)
Niño-EDU-Q1	-0.021 (0.018)	-0.018 (0.015)	0.114 (0.128)	41.831 (34.449)
Niña-Q3	0.003 (0.006)	0.009* (0.005)	-0.018 (0.043)	-3.896 (8.397)
Niña-Q2	0.005 (0.007)	0.013** (0.005)	-0.043 (0.051)	-27.676*** (8.027)
Niña-Q1	-0.016* (0.008)	0.004 (0.005)	0.100 (0.064)	0.761 (8.500)
Niña-EDU-Q3	-0.014 (0.019)	-0.014 (0.016)	0.045 (0.149)	-6.374 (29.367)
Niña-EDU-Q2	0.009 (0.022)	-0.005 (0.016)	-0.072 (0.165)	41.345 (29.095)
Niña-EDU-Q1	-0.045** (0.018)	-0.044*** (0.015)	0.349** (0.141)	71.075*** (25.640)
so2_ave-Q3	-0.001 (0.001)	-0.001 (0.001)	0.002 (0.009)	0.322 (1.608)
so2_ave-Q2	-0.003** (0.001)	0.001 (0.001)	0.022** (0.010)	-0.990 (1.766)
so2_ave-Q1	0.006*** (0.002)	0.003** (0.001)	-0.049*** (0.015)	-7.127*** (2.147)
Observations	9537	9537	9537	9537
NINO	0.01	0.01	-0.02	-19.42
P_NINO	0.25	0.01	0.74	0.03
NINA	-0.01	0.03	0.04	-30.81
P_NINA	0.53	0.00	0.61	0.01
NINO_EDU	-0.02	-0.03	0.06	32.88
P_NINO_EDU	0.28	0.03	0.71	0.29
NINA_EDU	-0.05	-0.06	0.32	106.05
P_NINA_EDU	0.02	0.00	0.05	0.00
POL	0.00	0.00	-0.02	-7.80
P_POL	0.11	0.10	0.08	0.00
r2_a	0.45	0.43	0.44	0.44

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the IPS, Weeks gest. is the average of weeks of gestation for all the children born in that IPS and Birth Weight is the average weight for all the children born in that IPS. Pollution is averaged per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while P_NINO , P_NINA and P_POL are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

Table 2.39: IV SECOND STAGE: PM_{10} $PM_{2.5}$ (average) ON HEALTH

	PM10 (average)				PM2.5 (average)			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)	Premature (5)	Low Weight (6)	Weeks Gest. (7)	Birth Weight (8)
Niño-Q3	0.000 (0.006)	0.010** (0.004)	0.003 (0.042)	-20.201** (8.023)	0.010 (0.013)	0.006 (0.007)	-0.053 (0.098)	-2.612 (13.168)
Niño-Q2	0.008 (0.007)	-0.005 (0.005)	-0.045 (0.048)	6.093 (8.778)	0.019 (0.013)	0.008 (0.007)	-0.137 (0.095)	-6.554 (9.735)
Niño-Q1	0.003 (0.006)	0.006 (0.005)	-0.002 (0.043)	-8.283 (9.069)	0.019* (0.010)	0.019** (0.008)	-0.130* (0.072)	-22.687* (12.633)
Niño-EDU-Q3	0.016 (0.024)	-0.005 (0.019)	-0.142 (0.181)	2.829 (33.066)	0.014 (0.035)	-0.005 (0.019)	-0.128 (0.269)	-27.853 (36.999)
Niño-EDU-Q2	-0.018 (0.025)	-0.015 (0.021)	0.100 (0.184)	9.009 (36.884)	-0.020 (0.037)	-0.014 (0.023)	0.114 (0.271)	20.988 (38.460)
Niño-EDU-Q1	-0.018 (0.018)	-0.014 (0.015)	0.093 (0.127)	31.849 (33.748)	-0.020 (0.025)	-0.006 (0.020)	0.144 (0.176)	3.382 (26.490)
Niña-Q3	0.004 (0.006)	0.007 (0.005)	-0.023 (0.041)	-0.364 (9.350)	0.006 (0.014)	-0.019* (0.010)	-0.000 (0.107)	22.271 (16.266)
Niña-Q2	0.004 (0.007)	0.011** (0.005)	-0.028 (0.049)	-25.653*** (8.615)	0.028** (0.014)	0.023*** (0.009)	-0.191* (0.102)	-48.046*** (12.434)
Niña-Q1	-0.006 (0.006)	0.006 (0.005)	0.027 (0.049)	-11.467 (8.490)	-0.001 (0.016)	-0.002 (0.009)	-0.011 (0.123)	11.382 (11.984)
Niña-EDU-Q3	-0.024 (0.019)	-0.015 (0.017)	0.117 (0.148)	1.757 (29.574)	-0.041 (0.042)	-0.038 (0.042)	0.136 (0.326)	-9.336 (67.038)
Niña-EDU-Q2	0.012 (0.023)	-0.002 (0.016)	-0.091 (0.175)	39.517 (29.819)	0.004 (0.043)	-0.020 (0.025)	-0.016 (0.307)	110.891*** (29.780)
Niña-EDU-Q1	-0.044** (0.019)	-0.043*** (0.015)	0.341** (0.145)	73.908*** (26.394)	-0.012 (0.030)	-0.012 (0.021)	0.113 (0.243)	-4.377 (34.296)
pm10_ave-Q3	0.000 (0.001)	-0.000 (0.000)	0.000 (0.004)	0.900 (0.551)				
pm10_ave-Q2	0.001* (0.001)	0.000 (0.000)	-0.006 (0.004)	-1.369** (0.681)				
pm10_ave-Q1	-0.000 (0.001)	-0.000 (0.000)	0.003 (0.005)	-0.470 (1.045)				
pm25_ave-Q3					0.004** (0.001)	0.001 (0.001)	-0.026** (0.011)	-0.746 (1.104)
pm25_ave-Q2					0.000 (0.002)	0.002** (0.001)	-0.007 (0.011)	-2.119 (1.565)
pm25_ave-Q1					0.001 (0.001)	0.002*** (0.000)	-0.009 (0.006)	-1.822* (1.080)
Observations	9537	9537	9537	9537	3838	3838	3838	3838
NINO	0.01	0.01	-0.04	-22.39	0.05	0.03	-0.32	-31.85
P_NINO	0.10	0.04	0.44	0.04	0.01	0.00	0.01	0.08
NINA	0.00	0.02	-0.02	-37.48	0.03	0.00	-0.20	-14.39
P_NINA	0.85	0.00	0.72	0.00	0.12	0.94	0.18	0.47
NINO_EDU	-0.02	-0.03	0.05	43.69	-0.03	-0.03	0.13	-3.48
P_NINO_EDU	0.32	0.02	0.75	0.16	0.26	0.11	0.45	0.91
NINA_EDU	-0.06	-0.06	0.37	115.18	-0.05	-0.07	0.23	97.18
P_NINA_EDU	0.01	0.00	0.03	0.00	0.20	0.06	0.42	0.11
POL	0.00	-0.00	-0.00	-0.94	0.01	0.01	-0.04	-4.69
P_POL	0.54	0.62	0.58	0.42	0.05	0.00	0.03	0.11
r2_a	0.46	0.43	0.45	0.44	0.46	0.45	0.46	0.43

* p<0.05> , ** p<0.01, *** p<0.001

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the IPS, Weeks gest. is the average of weeks of gestation for all the children born in that IPS and Birth Weight is the average weight for all the children born in that IPS. Pollution is averaged per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

Table 2.40: IV SECOND STAGE: $CO\ NO_X$ (peak afternoon) ON HEALTH

	CO (peak afternoon)				NOX (peak afternoon)			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)	Premature (5)	Low Weight (6)	Weeks Gest. (7)	Birth Weight (8)
Niño-Q3	-0.005 (0.006)	0.012** (0.005)	0.027 (0.045)	-24.237** (9.493)	0.001 (0.007)	0.014** (0.005)	-0.013 (0.051)	-37.692*** (11.636)
Niño-Q2	0.006 (0.008)	-0.001 (0.005)	-0.030 (0.062)	15.712 (10.025)	0.004 (0.009)	-0.005 (0.006)	-0.022 (0.064)	13.560 (11.987)
Niño-Q1	0.005 (0.007)	0.012** (0.006)	-0.026 (0.049)	-8.353 (8.505)	0.009 (0.007)	0.012** (0.005)	-0.054 (0.050)	-17.472 (10.870)
Niño-EDU-Q3	0.018 (0.025)	-0.009 (0.017)	-0.157 (0.192)	12.203 (30.938)	-0.034 (0.037)	-0.034 (0.025)	0.256 (0.264)	85.154* (48.574)
Niño-EDU-Q2	-0.022 (0.024)	-0.011 (0.017)	0.141 (0.180)	1.620 (32.078)	0.006 (0.041)	0.005 (0.024)	-0.066 (0.294)	-21.908 (53.445)
Niño-EDU-Q1	-0.016 (0.018)	-0.020 (0.015)	0.090 (0.126)	19.597 (22.722)	-0.041 (0.025)	-0.024 (0.018)	0.301* (0.178)	53.734 (44.595)
Niña-Q3	0.006 (0.006)	0.011** (0.005)	-0.028 (0.044)	-6.579 (8.567)	0.001 (0.007)	-0.001 (0.006)	-0.021 (0.048)	9.027 (10.250)
Niña-Q2	0.001 (0.006)	0.012*** (0.005)	-0.013 (0.047)	-26.490*** (7.880)	-0.004 (0.007)	0.005 (0.005)	0.027 (0.052)	-21.978** (9.654)
Niña-Q1	-0.002 (0.007)	0.011** (0.005)	0.002 (0.055)	-6.121 (8.073)	-0.014* (0.008)	0.001 (0.007)	0.085 (0.057)	0.763 (9.350)
Niña-EDU-Q3	-0.018 (0.020)	-0.011 (0.015)	0.065 (0.150)	-3.486 (28.121)	0.003 (0.035)	0.027 (0.026)	-0.033 (0.246)	-40.589 (47.962)
Niña-EDU-Q2	0.013 (0.021)	-0.005 (0.015)	-0.089 (0.164)	34.279 (28.722)	0.031 (0.032)	0.013 (0.018)	-0.236 (0.239)	34.376 (42.053)
Niña-EDU-Q1	-0.044** (0.018)	-0.040*** (0.014)	0.348** (0.144)	60.356** (26.576)	-0.023 (0.029)	-0.019 (0.020)	0.162 (0.210)	51.348 (35.499)
co_peaka-Q3	0.022*** (0.007)	0.012** (0.005)	-0.140*** (0.048)	-8.605 (8.709)				
co_peaka-Q2	0.001 (0.005)	0.006 (0.004)	-0.014 (0.039)	3.426 (6.971)				
co_peaka-Q1	-0.002 (0.005)	-0.001 (0.004)	-0.005 (0.035)	4.781 (7.090)				
nox_peaka-Q3					-0.000 (0.000)	0.000 (0.000)	0.001 (0.002)	-0.136 (0.338)
nox_peaka-Q2					-0.000 (0.000)	0.000 (0.000)	0.001 (0.002)	-0.353 (0.309)
nox_peaka-Q1					-0.000 (0.000)	0.000 (0.000)	0.001 (0.002)	-0.554* (0.336)
Observations	8439	8439	8439	8439	6457	6457	6457	6457
NINO	0.01	0.02	-0.03	-16.88	0.01	0.02	-0.09	-41.60
P_NINO	0.48	0.00	0.68	0.14	0.08	0.00	0.10	0.00
NINA	0.01	0.03	-0.04	-39.19	-0.02	0.00	0.09	-12.19
P_NINA	0.59	0.00	0.60	0.00	0.12	0.67	0.25	0.39
NINO_EDU	-0.02	-0.04	0.07	33.42	-0.07	-0.05	0.49	116.98
P_NINO_EDU	0.28	0.01	0.62	0.25	0.12	0.02	0.11	0.03
NINA_EDU	-0.05	-0.06	0.32	91.15	0.01	0.02	-0.11	45.13
P_NINA_EDU	0.02	0.00	0.05	0.00	0.84	0.61	0.80	0.54
POL	0.02	0.02	-0.16	-0.40	-0.00	0.00	0.00	-1.04
P_POL	0.09	0.06	0.09	0.98	0.61	0.21	0.49	0.19
r2_a	0.47	0.42	0.46	0.43	0.47	0.45	0.48	0.48

* p<0.05> , ** p<0.01, *** p<0.001

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the IPS, Weeks gest. is the average of weeks of gestation for all the children born in that IPS and Birth Weight is the average weight for all the children born in that IPS. Pollution is averaged per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

2M Effects on additional outcomes

Table 2.41: IV SECOND STAGE: SO_2 - PM_{10} (average) ON HEALTH

	SO2 (average)			PM10 (average)		
	Apgar1 (1)	Pre-Low Weight (2)	Sex ratio (3)	Apgar1 (4)	Pre-Low Weight (5)	Sex ratio (6)
Niño-Q3	-0.018 (0.014)	0.013*** (0.004)	0.014 (0.026)	-0.006 (0.015)	0.005 (0.004)	0.028 (0.024)
Niño-Q2	0.015 (0.017)	-0.010*** (0.004)	-0.048* (0.025)	-0.009 (0.016)	-0.001 (0.004)	-0.072** (0.028)
Niño-Q1	0.012 (0.015)	0.007* (0.004)	0.047* (0.026)	0.005 (0.017)	0.007 (0.005)	-0.042 (0.027)
Niña-Q3	-0.006 (0.013)	0.004 (0.003)	0.025 (0.025)	0.007 (0.015)	-0.002 (0.004)	0.021 (0.023)
Niña-Q2	0.017 (0.014)	0.004 (0.003)	-0.036 (0.024)	0.019 (0.017)	0.001 (0.004)	-0.041 (0.025)
Niña-Q1	0.009 (0.016)	-0.006 (0.005)	0.013 (0.032)	-0.009 (0.019)	-0.001 (0.005)	-0.033 (0.036)
so2_ave-Q3	0.005* (0.003)	-0.002* (0.001)	-0.009* (0.005)			
so2_ave-Q2	0.006* (0.003)	0.001 (0.001)	0.010* (0.005)			
so2_ave-Q1	-0.005 (0.004)	0.003*** (0.001)	-0.017** (0.008)			
pm10_ave-Q3				0.001 (0.001)	-0.001*** (0.000)	-0.001 (0.002)
pm10_ave-Q2				-0.003 (0.002)	0.001** (0.000)	-0.002 (0.003)
pm10_ave-Q1				-0.001 (0.002)	0.001 (0.001)	-0.002 (0.003)
Observations	11024	11027	10612	11024	11027	10612
NINO	0.01	0.01	0.01	-0.01	0.01	-0.00
P_NINO	0.57	0.01	0.66	0.59	0.05	0.94
NINA	0.02	0.00	0.00	0.02	-0.00	-0.05
P_NINA	0.30	0.82	0.96	0.44	0.74	0.23
POL	0.01	0.00	-0.02	-0.00	0.00	-0.01
P_POL	0.16	0.06	0.04	0.19	0.39	0.14
r2_a	0.98	0.35	0.08	0.98	0.35	0.09

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Apgar1 is the average of the APGAR score at 1 minute of birth (1-10) of the children born in the IPS, Pre-Low Weight is the percentage of premature and low birth weight children born in the IPS, and sex-ratio is the ratio of the percentage of boys over the percentage of girls for the children born in that IPS. Pollution is averaged per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

Table 2.42: IV SECOND STAGE: $CO-NO_X$ (peak afternoon) ON HEALTH

	CO (peak afternoon)			NOX (peak afternoon)		
	Apgar1 (1)	Pre-Low Weight (2)	Sex ratio (3)	Apgar1 (4)	Pre-Low Weight (5)	Sex ratio (6)
Niño-Q3	-0.016 (0.015)	0.010** (0.004)	0.026 (0.026)	0.005 (0.015)	0.008* (0.005)	0.066** (0.027)
Niño-Q2	0.015 (0.019)	-0.002 (0.005)	-0.057* (0.034)	0.010 (0.014)	-0.006 (0.004)	-0.051** (0.023)
Niño-Q1	0.022 (0.018)	0.008* (0.004)	0.003 (0.025)	0.007 (0.014)	0.013*** (0.004)	0.057* (0.030)
Niña-Q3	-0.005 (0.013)	0.005 (0.003)	0.007 (0.028)	0.002 (0.012)	0.000 (0.004)	0.026 (0.021)
Niña-Q2	0.016 (0.015)	0.004 (0.003)	-0.020 (0.024)	0.016 (0.013)	0.003 (0.004)	-0.021 (0.024)
Niña-Q1	-0.002 (0.017)	-0.001 (0.005)	-0.027 (0.035)	0.004 (0.016)	-0.009 (0.005)	0.032 (0.034)
co_peaka-Q3	-0.030 (0.020)	0.013*** (0.005)	-0.008 (0.046)			
co_peaka-Q2	-0.023 (0.016)	0.006 (0.004)	0.018 (0.031)			
co_peaka-Q1	-0.022** (0.011)	0.004 (0.003)	0.025 (0.022)			
nox_peaka-Q3				0.001 (0.001)	0.000 (0.000)	-0.000 (0.001)
nox_peaka-Q2				0.002** (0.001)	0.000 (0.000)	-0.000 (0.001)
nox_peaka-Q1				0.001 (0.001)	0.000 (0.000)	0.002 (0.001)
Observations	9830	9833	9460	7680	7683	7408
NINO	0.02	0.02	-0.03	0.02	0.01	0.07
P_NINO	0.30	0.01	0.44	0.30	0.02	0.07
NINA	0.01	0.01	-0.04	0.02	-0.01	0.04
P_NINA	0.66	0.14	0.42	0.29	0.44	0.28
POL	-0.08	0.02	0.03	0.00	0.00	0.00
P_POL	0.05	0.01	0.63	0.04	0.30	0.73
r2_a	0.98	0.34	0.09	0.98	0.36	0.08

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Apgar1 is the average of the APGAR score at 1 minute of birth (1-10) of the children born in the IPS, Pre-Low Weight is the percentage of premature and low birth weight children born in the IPS, and sex-ratio is the ratio of the percentage of boys over the percentage of girls for the children born in that IPS. Pollution is averaged per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay).

2N Index of pollution - Principal Component Analysis

Table 2.43: IV SECOND STAGE: PCA-A and PCA-B ON HEALTH

	PCA version A				PCA version B			
	Premature (1)	Low Weight (2)	Weeks Gest. (3)	Birth Weight (4)	Premature (5)	Low Weight (6)	Weeks Gest. (7)	Birth Weight (8)
Niño-Q3	-0.024 (0.025)	-0.009 (0.012)	0.133 (0.188)	2.073 (17.419)	-0.004 (0.007)	0.016** (0.007)	0.007 (0.051)	-25.281** (10.529)
Niño-Q2	0.023 (0.015)	0.009 (0.011)	-0.117 (0.113)	-20.322 (15.799)	-0.004 (0.009)	-0.002 (0.007)	0.036 (0.065)	4.078 (10.102)
Niño-Q1	0.004 (0.009)	0.013* (0.008)	-0.008 (0.071)	-27.088* (14.144)	0.005 (0.006)	0.018*** (0.006)	-0.004 (0.049)	-18.875** (9.125)
Niña-Q3	-0.009 (0.015)	-0.023** (0.011)	0.106 (0.123)	-0.468 (14.985)	-0.008 (0.006)	-0.002 (0.005)	0.084* (0.044)	10.247 (8.930)
Niña-Q2	0.007 (0.022)	0.019* (0.011)	-0.050 (0.174)	-22.326 (16.223)	-0.002 (0.006)	0.018*** (0.005)	-0.002 (0.048)	-24.796*** (9.190)
Niña-Q1	0.026 (0.019)	-0.003 (0.016)	-0.122 (0.154)	-14.733 (19.764)	-0.005 (0.008)	-0.002 (0.007)	0.061 (0.065)	-2.782 (10.575)
pcaA1-Q3	-0.014 (0.193)	0.030 (0.122)	0.162 (1.406)	263.387 (170.478)				
pcaA1-Q2	-0.253 (0.163)	0.166 (0.142)	1.695 (1.137)	-5.202 (174.836)				
pcaA1-Q1	-0.276** (0.127)	0.019 (0.095)	1.991** (0.962)	151.374 (132.264)				
pcaB1-Q3					0.115* (0.062)	0.059 (0.059)	-0.902* (0.479)	-128.453 (94.796)
pcaB1-Q2					-0.056 (0.038)	0.078* (0.043)	0.316 (0.304)	-41.247 (58.204)
pcaB1-Q1					-0.021 (0.034)	0.050** (0.025)	0.151 (0.230)	25.838 (55.389)
Observations	2765	2765	2765	2765	6732	6732	6732	6732
NINO	0.00	0.01	0.01	-45.34	-0.00	0.03	0.04	-40.08
P_NINO	0.94	0.48	0.97	0.06	0.76	0.00	0.64	0.00
NINA	0.02	-0.01	-0.07	-37.53	-0.01	0.01	0.14	-17.33
P_NINA	0.59	0.82	0.85	0.28	0.20	0.13	0.11	0.21
POL	-0.54	0.22	3.85	409.56	0.04	0.19	-0.43	-143.86
P_POL	0.20	0.49	0.21	0.31	0.71	0.07	0.58	0.35
r2_a	0.36	0.30	0.34	0.39	0.37	0.31	0.35	0.36

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: Premature is the percentage of premature births in the IPS, Low Birth is the percentage of low birth weight in the IPS, Weeks gest. is the average of weeks of gestation for all the children born in that IPS and Birth Weight is the average weight for all the children born in that IPS. Pollution is averaged per day, then averaged by month and by quarter of gestation for each IPS. The estimations include yearxlocality fixed effects, month fixed effects, health center fixed effects, controls for household characteristics and use the instruments mentioned in section 2.4.1. *NINO*, *NINA* and *POL* correspond to the sum of the coefficients (the cumulative effect) of the variables for El Niño, for La Niña and for the pollutant, respectively, while *P_NINO*, *P_NINA* and *P_POL* are their respective p-values. The standard errors in parenthesis are robust to autocorrelated cross-panel disturbances (Driscoll-Kraay). The principal components (PCA) are constructed using the standardized version of the pollutants of the main specifications (*SO₂*, *PM₁₀*, *PM_{2.5}* (average), *CO*, *NO_X* (peak afternoon)). The estimations use the first principal component, which has an eigenvalue larger than one. As *PM_{2.5}* has many missing values, PCA-A version includes *PM_{2.5}* (explained variation of 48.9%), while PCA-B does not include *PM_{2.5}* (explained variation of 44.1%).

Chapter 3

Droughts and Agricultural Adaptation to Climate Change

Abstract

This article analyses the effects of droughts and climate variability on short-term and medium-term adaptation of Colombian rural households. I measure drought in a Differences-in-Differences (DID) framework, as an alternative to the standard approaches decomposing the effects from climate and yearly weather deviations on agricultural productivity and those using the growing degree days and harmful degree days. In the short-term and medium-term, rural households adapt to the drought of 2010 by increasing the total area planted in crops and livestock, (increasing also the total gross agricultural productivity in value terms) and by working more on the farm. The droughts also increased the use of external sources of water in the farm and made rural households postpone non-housing investments in the farm. I find heterogeneous effects according to the long run mean of temperature in the municipality. Higher temperature affects positively gross agricultural productivity in low-temperature municipalities but negatively high-temperature municipalities. Cereals and coffee seem to benefit from higher temperatures, while vegetables and fruits are more affected.

Keywords: climate change, weather, agriculture, gross productivity, adaptation, rural impacts.
JEL: Q54, O13, Q12, Q15, R20.

3.1 Introduction

Climate change is affecting many aspects of human life, not only through distortions of weather patterns, or increases of the intensity of phenomena such as the El Niño Southern Oscillation (ENSO events or El Niño-La Niña events), but primarily through the impact it has on health, migration and agriculture, among many other aspects. The threat it can pose on food safety and food production of rural households in many countries is undeniable, and one of the most important impacts is on agriculture. In their review of the literature, [D'Agostino and Schlenker \(2016\)](#) note that climate change could reverse the gains in average yields obtained with the Green Revolution, by increasing the volatility of food production. According to [Ortiz-Bobea et al. \(2021\)](#), anthropogenic climate change has reduced global agricultural total factor productivity (TFP) by about 21% since 1961, equivalent to losing the last 7 years of productivity growth. The damaging effect of climate change on TFP has been more severe (a reduction of around 26% to 34%) in warmer regions such as Africa and Latin America and the Caribbean. In spite of the magnitude of the effect, [Auffhammer and Schlenker \(2014\)](#) point out that the main studies of crop responses to climate change have been focused on important food crops and major producers rather than on low income and small producer countries. In this sense, the literature for developing countries and small rural households still remains scarce and particularly focused on U.S, as pointed out by [Kolstad and Moore \(2020\)](#).

Only recently, [Aragón et al. \(2021\)](#) analysed how subsistence rural households respond to extreme heat in Peru. In this case, high temperatures reduce gross agricultural productivity (in value terms), and rural households attenuate the effect on output by increasing the area planted, using more crop mixing and selling livestock. In this article, I study small rural households in Colombia and their adaptation over the medium-term to contribute to this scarce literature on responses to climate change in low-income countries. The article aims to analyse first, how droughts affect agricultural decisions and food production of rural households in Colombia, in the short-term and medium term and for consecutive droughts. To what extent, and by which means, do rural households adapt to these shocks? And what are the differences in rural households' behaviour in the short-term and the medium-term versus the adaptation to consecutive shocks? The analysis also focuses on an important aspect of heterogeneity: the differences between rural households located in

municipalities with high average temperature and those located in municipalities with low average temperature in the past.

In order to study rural households' adaptation to these shocks, I use the Colombian panel survey (ELCA) which has very rich information on agricultural production and investments made in the land of rural households. It also has information about the problems faced by the household such as losing crops, weather shocks among others, but I mainly rely on weather data from satellite images in order to avoid the measurement errors that could arise from using subjective self-reported shocks. Since the panel follows the same household over three waves, I can study adaptation such as using more some crops than others, change in the use of inputs or changes towards other types of agricultural activities. This is a main advantage compared to studies such as [Aragón et al. \(2021\)](#), which use repeated cross-section data. Also, the use of panel data to study weather impacts on agriculture allows for a better causal identification, as weather deviations around the mean are random and exogenous ([Blanc and Schlenker \(2017\)](#)).

In terms of agricultural adaptation, [Costinot et al. \(2016\)](#) and [Burke and Emerick \(2016\)](#) show that crop switching could be a possible response to climate change. Rural households might adapt by making investments in the unit of agricultural production, trying to get technical assistance or modifying the use of fertilizers. The information on these adaptive margins are also available in the panel survey ELCA. Similarly, [Rosenzweig and Wolpin \(1993\)](#) also highlight that households could respond to droughts on the intensive margin by increasing off-farm work (see [Jayachandran \(2006\)](#)), selling cattle (see [Fafchamps et al. \(1998\)](#)), or on the extensive margin, through migration (see [Cattaneo et al. \(2019\)](#)), among several others. The roughness, altitude and the different climatic zones are characteristic of the Colombian territory, which have posed some challenges for agricultural production, affecting also transportation. This heterogeneity affects not only weather but also the type of crops that can be produced, and together with the non-linear effects for the agricultural production (see [Deschênes and Greenstone \(2007\)](#) and [Schlenker and Roberts \(2009\)](#)) are additional aspects to consider.

The analysis of the article can be divided in three parts. First, I separate the effect of climate

from yearly weather deviations for the Colombian rural households as has been done in the literature (Kelly et al., 2005; Deschênes and Kolstad, 2011; Burke and Emerick, 2016; Bento et al., 2020). Second, I explore different adaptive margins in response to extreme temperature using two measures: the Growing Degree Days (GDD) and the Harmful Degree Days (HDD) (Schlenker and Roberts, 2009; Deschênes and Greenstone, 2007; Aragón et al., 2021). And third, I suggest that an alternative measure of drought can be used in a Differences-in-Differences (DID) framework to analyse how rural households adapt to climate and weather. The DID analyses the short-term and the medium-term adaptation and then, the adaptation with respect to consecutive droughts.

The main findings of the article can be broken down in five. First, with respect to the analysis of climate and yearly weather deviations of section 3.4.2, gross agricultural productivity (in value terms) is positively affected by the long-run mean temperature, while there is no effect from temperature shocks (deviations from the long-run mean); with respect to the marginal effects, higher temperature affects positively gross agricultural productivity (in value terms) in low-temperature municipalities but negatively in high-temperature municipalities. With the exception of Aragón et al. (2021), the heterogeneous effects of temperature have been explored for the case of U.S. (Deschênes and Greenstone, 2007; Schlenker and Roberts, 2009) but less for developing countries. Second, across several of the analyses of the short-term and medium-term (climate-yearly weather deviations, GDD-HDD and the DID), the gross agricultural productivity (in value terms) of cereals and coffee benefit from droughts or when facing higher temperatures. The gross agricultural productivity increases in low-temperature municipalities for cereals and coffee, while those in high-temperature areas are negatively effected. Vegetables seem to be more affected following droughts (DID analysis). Across the different analyses, the gross agricultural productivity (in value terms) of vegetables and fruits present reductions, particularly in low-temperature municipalities (increases in high-temperature municipalities). Third, as regards short-term and medium-term, the drought of 2010 made the rural households re-allocate land by reducing the type of land left fallow and assigning it to production of crops and for livestock. This captures some of the trade-offs proposed in the theoretical model of section 3.3. It also goes in line with the increase of farm inputs (land and labor) as in Aragón et al. (2021) and with increases in the total gross agricultural productivity (in value terms). The drought of 2010 also affected labor market outcomes, making household heads

and his/her partner to work more on the farm. In addition, rural households reduce the investments after droughts, which could be interpreted as postponing investment decisions and a way to smooth consumption. They also increase the use of external sources of water as their own water sources dries up during droughts. Fourth, for the consecutive drought of 2010 and 2013, rural households continue implementing some strategies as in the short-term and medium-term, but the adaptation becomes more difficult once the droughts start to be more frequent. On the one hand, some strategies observed before like working more in agriculture, using external sources of water and smoothing consumption by postponing investments, are less likely to be implemented. On the other hand, the rural households keep using more land available for crops and there is an increase in the land area with investments (excluding housing). In this sense, droughts might make rural households aware of the potential future benefits of making investments, which is compatible with the findings of [Burke and Emerick \(2016\)](#). Fifth, the analysis of the climate-yearly weather deviations of section [3.4.2](#) and the DID of section [3.5.2](#) gave qualitatively similar findings, with a positive effect of temperature shocks and droughts (short-term and medium-term in the DID) on gross agricultural productivity (in value terms), with a negative marginal effect in high-temperature municipalities, while a positive marginal effect in low-temperature ones. This article thus proposes an alternative manner to examine the short-term and medium-term adaptation decisions of the rural households.

The article contributes in four ways: first, it adds to the scarce literature in low and middle income countries, and proposes an alternative measure to assess climate impact and adaptation in agriculture in the short-term and in medium-term in Colombia. Second, it explores the heterogeneity between households living in high versus low-temperature areas, an aspect not considered yet in the literature. Third, it shows that in the short-term and medium-term, rural households in Colombia adapt by using more available land, which could lead to higher gross agricultural productivity (in value terms). Another novelty is to analyse consecutive droughts. Fourth, it explores in more detail how rural households adapt to droughts by analysing a broader set of crops, vegetables, fruits, cereals and coffee. As pointed out by [Hertel and de Lima \(2020\)](#), the FAO identifies 175 distinct crops but the majority of climate impact studies have focused on changes in yields for four staple crops, maize, wheat, rice and soybeans. These four staple crops account for only one-quarter of the total value of agricultural output. The article gives evidence of climate impacts outside of

the staple crop domain and explores other farm inputs such as labor, the use of investments and access to water. According to [Hertel and de Lima \(2020\)](#), the literature should move beyond the yield impact where we have better data and models, and move towards other food products, farm inputs and nutritional impacts.

The article proceeds in the following manner: section [3.2](#) describes the ELCA panel survey and the descriptive statistics on gross agricultural productivity; section [3.3](#) proposes a simple theoretical framework that helps to explain the mechanisms captured by the empirical findings; section [3.4.2](#) analyses the impact of climate and weather deviations on gross agricultural productivity, distinguishing between weather deviations and long-term climate averages; section [3.4.4](#) analyses the effect of gradual changes in temperature on gross agricultural productivity (in value terms) to see how rural households adapt in Colombia; section [3.5](#) proposes a DID strategy to analyse if there are differences in short-term and medium term adaptation compared to the adaptation with respect to consecutive droughts; and section [3.6](#) concludes.

3.2 Data and Descriptive Statistics

3.2.1 ELCA 2010-2013-2016

I use the ELCA survey (The Colombian panel survey), a nationally representative panel following the same households for three waves, in 2010, in 2013 and 2016. It is conducted by the University of Los Andes in rural and urban areas, surveying also units of agricultural production, which vary in terms of land size, the type of crops and type of production (crops, livestock, etc.). In the rural areas, it is representative for the small rural households (peasants) of the four micro-regions sampled (Atlantica Media-Cundi Boyacense-Eje Cafetero-Centro Oriental) and the attrition is very low, only 3.6% (see [Fuertes et al. \(2017\)](#)). It covers 17 municipalities and 224 veredas (small rural areas inside a municipality).

The ELCA panel follows the same household for different rounds, with information on the type of crops used by each agricultural unit. It also helps to identify long-term adaptive strategies such as migration, changes in agricultural technology and investment, or dynamic effects of short term

responses (changes in the type of crops or effects on the rural labor market).

The panel survey is composed of 10,800 households among which 4,800 are rural households and 6,000 are urban households (see Table 3.13 in Appendix 3C for the sample per municipality and year). It has a section of questions by household, a section of different land parcels belonging to the household (land section) and another section by crops and livestock that the household produces (production section). Thus, the land section and the production section are aggregated by rural household and merged with the household section. I will focus on the rural sample (44.4% of the total), restricted to households with at least one parcel used in agricultural activities, excluding land given to someone else or sold (the final sample comprises 95% of reported land in 2010, 92% in 2013 and 93.2% in 2016). I also restrict the main sample to households with complete information on land use and investment and I exclude households moving to municipalities outside of the initial sample of 17 municipalities (647 of the total of 12,804 household observations for the three waves in Table 3.13 of Appendix 3C). I explore in more detail some aspects of migration for the larger sample in the section of results.

Table 3.1 and Table 3.2 summarize the descriptive statistics of the outcome variables, the control and weather variables used.¹ For the outcomes, I construct the variable "Per. crops (YEARLY)" as the percentage of annual crops among the total number of parcels that the rural household has. Yearly crops constitutes 61% of the sample, so inter-annual cropping is less frequent in this sample. The ELCA provides detailed information on agricultural costs, sales, total agricultural production (see the description of the construction in the next paragraph), as well as important data on investment. For the latter variables, the survey records a dummy if there was an investment on the parcel of land and I adjust multiplying by the land size of that plot; for example, if a household has two plots, one with two hectares and another one with three hectares, and made an investment in the first one, the variable aggregated by household takes a value of two hectares.² The aggre-

¹I added a value of 1 to all the variables using logarithm in order to avoid missing information for cases when the household reported a production or values of zero, which could make the logarithm of zero to be undefined. The use of logarithm is justified to reduce the variability, avoid extreme values and have outcome variables that resemble better a normal distribution. An alternative was to use the inverse hyperbolic sine (IHS) transformation. Although logarithm and IHS gave similar distributions, I prefer to use the logarithm in order to have a more straightforward interpretation of the results.

²The majority of the times, there was only one investment made in each plot. For the cases where the household

gated variable per rural household "Land (ha) ANY INVEST" captures the hectares of parcels in which the household did any investment. Table 3.1 shows that on average the rural households made investments for housing on 0.32 hectares and investments for structural investments on 0.27 hectares, compared to the average land size of the household of 2.89 hectares. Investments excluding housing were made on 0.69 hectares, which represents 23.9% of the average total land size of the rural households. Other relevant outcomes are "Land size (hectares)" of the household which is separated by the different type of uses (in permanent crops, mixed crops, livestock, etc.). Finally, the ELCA provides data of livestock, which were homogenized in Tropical Livestock Units according to the guidelines of FAO units for international comparison (see FAO (2016)).³ In terms of labor, the ELCA has information on whether or not the household head (or partner) were employed, or looked for a job, if she or he has an agricultural job, etc. Finally, "Land (ha) (ANY WATER)" is the hectares of land of the household that were declared to have access to water. The descriptive statistics of all outcome variables are shown in Table 3.1. The ELCA panel survey records the variables of land size and investment questions in a separate section asking the rural household questions by plots, which is another section of the rural agricultural production. Both sections, the agricultural production and the one of parcels are linked only by the rural household number.

For the control variables, Table 3.2 includes data of the household characteristics such as household size, age of the household head, "N. persons <14" as the number of persons below or equal to 14 years old in the household, and a dependency ratio variable as the ratio between the household members below or equal to 14 years old plus household members above or equal to 64 years old, divided by the household members between 15 to 64 years old. "Percentage crops (DROUGHT problem)" as the percentage of crop parcels for which drought was reported as a problem, "Dummy community prob. (lackwater)" is a question if the community where the household lives faced a problem of lack of water, "PCA 1: HH WEALTH"⁴ as the first principal component for the analysis of wealth measures of the household, "Dummy HH has credit" as one or zero if the household has

made more than one investment in the same plot, I divide the number of hectares among the total investments. For a plot of three hectares making two investments, I will assign 1.5 hectares to each investment.

³Cattle corresponds to 0.7 Tropical Livestock Units (TLU) in South America, sheep and goats to 0.1 TLU, pigs to 0.25 TLU, asses to 0.5 TLU, horses to 0.65 TLU, mules to 0.6 TLU and chickens to 0.01 TLU.

⁴For the estimations, I create indicator variables for five quintiles in the wealth distribution in 2010 (initial period), in which the household belongs to.

access to credit in that year, "Dummy access cabecera is ok (community)"⁵ as a dummy if the community has good access to the town center of the municipality, "Minutes to reach cabecera (community)" for the minutes from the community to reach the town center of the municipality, and finally, "Land (ha) (OWNED)" a variable which the hectares of land that were reported as owned. The descriptive statistics of all control variables are shown in Table 3.2.

In the estimation section, I only use the dependency ratio (0-14 years old and +65 years old with respect to the household size) variable instead of "HH size", "N. persons <=14", "N. persons 15-64" and "N. persons >=65". I also exclude some potentially endogenous variables in the estimations: "Dummy HH FAMILIAS EN ACCION", "Percentage crops (DROUGHT problem)", "Percentage crops (RAIN problem)", "Dummy community prob. (prices)" and "Dummy community prob. (lackwater)". "Edu. HH head" is excluded for the many missing. However, I include the variables here for illustration.

⁵1 for "Carretera pavimentada en buen estado (paved road in good shape)", "Carretera pavimentada en mal estado (paved road in bad shape)", "Carretera sin pavimentar en buen estado (no paved road in good shape)". 0 for "Carretera sin pavimentar en mal estado (no paved road in bad shape)", "Trocha o carreteable (bad shape road)", "Camino de herradura (bridle path)", "Mar, río o caño (sea or river)."

Table 3.1: Outcome variables rural households - Descriptive Statistics

	Observations	Mean	SD	Min	Max
Percentage crops (YEARLY)	9056	0.61	0.40	0	1
Percentage crops (SEMIANNUAL)	9056	0.18	0.31	0	1
Percentage crops (QUARTERLY)	9056	0.04	0.14	0	1
Percentage crops (MONTHLY)	9056	0.09	0.21	0	1
Percentage crops (BIMONTHLY-OTHER)	9056	0.08	0.20	0	1
Total Agro costs (Millions Col. Pesos/Ha)	10454	4.70	35.32	0	1580
Agro sale (Millions Col. Pesos/Ha)	10454	2.07	21.90	0	1662
vr prodtot (Millions Col. Pesos/Ha of crops)	3403	21.12	263.55	0	12770
vr prodtot-vegetable (Millions Col. Pesos/Ha of crops)	3403	7.37	93.14	0	2505
vr prodtot-fruit (Millions Col. Pesos/Ha of crops)	3403	4.81	183.98	0	9670
vr prodtot-cereal (Millions Col. Pesos/Ha of crops)	3403	7.49	163.97	0	12770
vr prodtot-coffee (Millions Col. Pesos/Ha of crops)	3403	1.45	11.18	0	425
Land (ha) ANY INVEST	11633	1.01	3.38	0	103
Land (ha) IRRIGATION INVEST	11633	0.05	0.78	0	53
Land (ha) STRUCTURES INVEST	11633	0.27	1.74	0	100
Land (ha) CONSERVATION INVEST	11633	0.12	1.03	0	50
Land (ha) FRUITS INVEST	11633	0.10	0.85	0	35
Land (ha) WOOD INVEST	11633	0.03	0.43	0	30
Land (ha) COMMERC. INVEST	11633	0.03	0.37	0	12
Land (ha) HOUSING INVEST	11633	0.32	1.55	0	40
Land (ha) OTHER INVEST	11633	0.09	0.89	0	40
Land (ha) ANY NO HOUSING INVEST	11633	0.69	2.77	0	103
Land size HH (Ha)	11633	2.89	6.35	0	118
Land PERMANENT crops (Ha)	11633	0.36	1.02	0	21
Land TRANSITIONAL crops (Ha)	11633	0.27	0.92	0	40
Land MIXED crops (Ha)	11633	0.17	0.89	0	30
Land LIVESTOCK (Ha)	11633	1.14	4.27	0	118
Land PASTURE (Ha)	11633	0.15	1.05	0	64
Land FOREST (Ha)	11633	0.13	1.22	0	90
Land OTHER USES (Ha)	11633	0.07	0.44	0	32
Land NO USED (Ha)	11633	0.35	1.73	0	62
Total Area planted (Ha) (perman.+trans.+mixed)	12157	0.78	1.59	0	40
Dummy after moving to another comm inside mpio 2010	12157	0.11	0.31	0	1
Tropical Livestock Units (FAO reference)	11644	6.42	37.66	0	2520
HH head-partner employed	12157	0.67	0.47	0	1
HH head-partner look for job	12157	0.15	0.36	0	1
HH head-partner agro work	12157	0.50	0.50	0	1
HH head-partner no agro work	12157	0.23	0.42	0	1
HH head-partner Ave. wage (Millions Col. Pesos)	4473	0.35	0.32	0	6
HH head-partner Ave. hours worked month	8227	35.76	18.56	1	126
Land (ha) (ANY WATER)	11633	2.14	5.66	0	118
Land (ha) (OWN WATER)	11633	1.27	4.28	0	100
Land (ha) (EXTERNAL WATER)	11633	0.87	2.93	0	118

Note: based on the ELCA panel survey 2010-2013-2016, using expansion factors for 2010 as recommended by the ELCA team.

Table 3.2: Controls and weather variables rural households - Descriptive Statistics

	Obs.	Mean	SD	Min	Max
Dependency ratio ($\leq 14 + \geq 65$)/(15-64) Per. 2010	12157	83.62	73.69	0	700
HH size	12157	4.59	2.14	1	19
N. persons ≤ 14	12157	1.44	1.41	0	10
N. persons 15-64	12157	2.74	1.42	0	12
N. persons ≥ 65	12157	0.41	0.66	0	4
Dummy HH FAMILIAS EN ACCION	12157	1.51	0.50	1	2
Age HH head	12157	49.76	12.81	16	100
Dummy women HH head	12157	0.21	0.40	0	1
Edu. HH head	9231	4.23	3.31	0	21
Percentage crops (DROUGHT problem)	10411	0.30	0.38	0	1
Percentage crops (PEST problem)	10411	0.26	0.32	0	1
Percentage crops (BRUSH problem)	9518	0.03	0.14	0	1
Percentage crops (RAIN problem)	10411	0.03	0.12	0	1
Percentage crops (SEEDS problem)	9518	0.01	0.07	0	1
Percentage crops (VANDALISM problem)	10411	0.01	0.06	0	1
Percentage crops (OTHER problem)	10411	0.04	0.15	0	1
Percentage crops (NONE problem)	10411	0.47	0.39	0	1
Dummy community prob. (land quality)	11111	0.36	0.48	0	1
Dummy community prob. (transport)	11111	0.70	0.46	0	1
Dummy community prob. (internal abuse)	11111	0.72	0.45	0	1
Dummy community prob. (costs)	11111	0.95	0.22	0	1
Dummy community prob. (armed groups)	11111	0.03	0.17	0	1
Dummy community prob. (lack credit)	11111	0.39	0.49	0	1
Dummy community prob. (others)	11111	0.12	0.32	0	1
Dummy community prob. (selling goods)	11111	4.84	1.47	0	8
Dummy community prob. (lackwater)	11111	0.63	0.48	0	1
Dummy community prob. (prices)	11111	0.95	0.23	0	1
Land (ha) (OWNED)	11633	2.15	5.27	0	118
Land (ha) (OWNED WITH TITLE)	11633	1.66	4.90	0	118
PCA 1: HH WEALTH	12157	0.46	0.24	0	1
PCA 1: HH (LIVESTOCK ASSETS)	12157	0.06	0.84	-1	4
PCA 1: HH (AGR. ASSETS)	12157	-0.01	0.54	-0	6
Dummy HH has credit	12157	0.50	0.50	0	1

continue next page ...

... continuation

	Obs.	Mean	SD	Min	Max
Dummy access cabecera is ok (community)	12157	0.28	0.45	0	1
Minutes to reach cabecera (community)	11111	42.47	36.22	1	240
Annual Ave daily Rainfall 1981-2010-CHIRPS (mm) agro-season	12157	4.07	1.10	2	7
Annual Ave daily Temp. 2001-2010 (Celsius MODIS) agro-season	12157	29.16	5.42	19	38
Yearly Ave days with rainfall 1981-2010-CHIRPS (mm)	12157	136.36	30.40	65	176
Yearly Ave days with temp. > 34C 2001-2010 (MODIS)	12157	22.94	15.67	0	52
Rainfall Trend 30 years-MEAN (mm) agro-season	12157	4.32	0.76	3	6
Rainfall deviation from trend-MEAN (mm) agro-season	12157	-0.25	0.81	-2	1
Temperature Trend 15 years-MEAN (Celsius) agro-season	12157	28.85	5.02	20	35
Temperature deviation from trend-MEAN (Celsius) agro-season	12157	0.31	1.06	-2	3
Temperature Trend 15 years-MAX (Celsius) agro-season	12157	31.11	4.30	22	37
Temperature deviation from trend-MAX (Celsius) agro-season	12157	0.25	1.27	-2	3
Temperature GDD-MEAN (Celsius) agro-season	12157	12.78	4.51	4	18
Temperature HDD-MEAN (Celsius) agro-season	12157	1.00	1.13	0	5
Temperature GDD-MAX (Celsius) agro-season	12157	15.02	3.07	8	18
Temperature HDD-MAX (Celsius) agro-season	12157	2.74	2.17	0	8
SPI Index 12 months CHIRPS	12157	0.22	1.14	-2	2
SPI Index 12 months CHIRPS(opposite)	12157	-0.22	1.14	-2	2
SPI Index 12 months CHIRPS(opposite)-agro season	12157	-0.55	1.24	-2	2
SPI Index 12 months CHIRPS(opposite)-positive values	12157	0.47	0.72	0	2
SPI Index 12 months CHIRPS(opposite)-negative values	12157	-0.69	0.49	-2	0
Per. months droughts(<=-1)-yearly for SPI Index 12	12157	0.21	0.37	0	1
Per. months droughts(<=-1)-yearly for SPI Index 12-agro season	12157	0.19	0.36	0	1
Per. months droughts(<=-2)-yearly for SPI Index 12	12157	0.08	0.18	0	1
Per. months droughts(<=-2)-yearly for SPI Index 12-agro season	12157	0.03	0.10	0	1
Dummy 1 or more months drought(<=-1)(SPI12)	12157	0.32	0.47	0	1
Dummy 1 or more months drought(<=-1)(SPI12)-agro season	12157	0.23	0.42	0	1
Dummy 1 or more months drought(<=-2)(SPI12)	12157	0.21	0.41	0	1
Dummy 1 or more months drought(<=-2)(SPI12)-agro season	12157	0.15	0.35	0	1

Note: based on the ELCA panel survey 2010-2013-2016, using expansion factors for 2010 as recommended by the ELCA team. "PCA 1: HH WEALTH" is the first principal component and it is already provided in the ELCA. "PCA 1: (LIVESTOCK ASSETS)" and "PCA 1: (AGR. ASSETS)" are also the first principal component but constructed here.

Table 3.3: Values of gross agricultural productivity by municipality

Municipalities	vr prodtot (ha) Vegetables	vr prodtot (ha) Fruits	vr prodtot (ha) Cereal	vr prodtot (ha) Coffee	vr prodtot (ha) Average
Saboya	1.33	0.84	5.96	0.00	8.13
Cerete	13.63	2.00	8.08	0.00	23.71
Chinu	6.20	0.32	41.94	0.00	48.46
Ciénaga de Oro	24.73	2.12	3.74	0.01	30.60
Sahagun	6.93	0.01	14.70	0.00	21.63
Simijaca	0.24	0.39	12.90	0.00	13.53
Susa	25.97	1.55	148.25	0.00	175.77
Tocaima	0.16	1.63	2.14	0.00	3.94
Circasia	27.49	1.46	6.07	4.28	39.30
Cordoba	3.40	0.15	5.49	4.21	13.25
Filandia	5.57	0.24	8.27	3.83	17.92
Belen de Umbria	7.47	2.08	1.73	4.94	16.21
Puente Nacional	0.91	0.59	1.01	1.12	3.63
Sampues	1.59	0.11	4.22	0.00	5.92
Natagaima	1.50	0.76	4.97	0.00	7.23
Ortega	5.54	25.08	4.25	3.33	38.21
Purificacion	4.41	0.69	5.27	0.00	10.37
Average	6.90	4.13	10.73	1.88	23.65

Note: ELCA Panel. The data section describes how the variables were constructed. Each column has the value in million of Colombian Pesos/hectares of household's land in crops. Sample restricted to only crop producers.

Groups of crops:

Vegetables: eggplant, broccoli, onion, green peas, chickpeas, any vegetable, sweet beans, tomato, carrots, cabbage, potatoes, etc.;

Fruits: coconut, avocado, anon, araza, banana, cacao, chirimoya, plum, curuba, guanabana, guava, lemon, etc.; aromatic herbs, palm, wood, others, were included here as the total of them had a very low percentage;

Cereal: beans, soybeans, rice, corn, sorghum, cotton, wheat, etc.;

Coffee: all types of coffee.

The variables of agricultural yields per hectares are calculated in the following manner (see Table 3.3). The survey in 2010 disaggregates the total production in kilograms for each crop, how many kilograms for consumption and how many kilograms for sale. It also has the sales in Colombian Pesos of the kilograms sold. This allows to infer an implicit price for each crop of the kilograms sold ($\text{Price} = \text{Value} / \text{Quantity}$), which can be averaged to have a price of different crops inside each municipality. Prices are aggregated by municipality in order to avoid endogeneity when using the implicit price of each crop of the rural household. These prices can then be used to value the total production of each crop by each household

in Colombian pesos and sum up all the values to have a measure of yields in pesos per household. As the subsequent surveys in 2013 and 2016 did not disaggregate the total production of kilograms by consumption and sales, I calculate the average implicit prices for 2010 at the level of municipality (or for the whole country) and apply those prices to estimate a measure of yields in Colombian Pesos for the three surveys. I then have a variable in Colombian pesos per household. A drawback of the construction of the gross agricultural production in value (Colombian Pesos) is that it is not able to distinguish exactly price effects and physical yield effects. In the particular case of Colombia and the ELCA panel survey, rural households tend to cultivate several crops, livestock, etc., which can make very hard to sum products that have different units. For this reason, I choose to calculate agricultural yields valued in Colombian Pesos (gross agricultural productivity-value terms), rather than physical units of production. Using total production in physical units can be a better measure of yields in contexts with farmers producing mainly one crop or specializing in a few ones. It is also important to mention that no prices were reported for livestock, and it is not possible to construct a measure of value of livestock production in Colombian Pesos. Therefore, the variables of total gross agricultural productivity (in value terms) do not include livestock.⁶ For this reason, the variables of gross agricultural productivity are restricted to only crop producers and capture productivity coming from crops.

Table 3.3 shows the average value in million of Colombian Pesos/hectares of household's land used in crops for the municipalities of the sample (the gross agricultural productivity). Coffee is mainly produced in Circasia, Córdoba, Filandia, Belén de Umbría and Ortega while cereals are important in Susa and Chinú and Cereté. Vegetables and fruits are produced in the majority of the municipalities. The table also has a description of the crops of each category.

⁶The literature of agricultural economics usually uses measures of "agricultural productivity", which corresponds to either yield by hectare (in physical terms) or some "net profitability" measure after deducting the costs of the inputs used in production. The measure used here corresponds more to gross agricultural productivity in value terms or a gross revenue per hectare.

3.2.2 Weather variables - The Standard Precipitation Index (SPI)

The article uses external data to construct measures of droughts and capture them with more precision. Two measures are extensively used in the literature to capture more adequately droughts, the Standardised Precipitation Index (SPI) proposed by [McKee et al. \(1993\)](#) and the Standardised Precipitation-Evapotranspiration Index (SPEI) of [Vicente-Serrano et al. \(2010\)](#). The SPI is defined as the number of standard deviations by which an observed anomaly deviates from the long-term mean. It can be calculated for any monthly scale (12 or 6 months here), which is the number of months over which water deficits accumulate. It considers the long-term time series of precipitation accumulated over the desired time scale to estimate an appropriate probability density function. Thus, a long period of data is necessary for the calculation (longer than 30 years is desirable). As the rainfall frequency distribution is positive skewed (like Gamma, Pearson III, etc.), the SPI entails a transformation to represent it according to a normal (Gaussian) distribution. Compared to the SPI, the SPEI adjusts by temperature, considering the climatic water balance (the difference between precipitation and potential evapotranspiration (PET) for each month in each location). In both cases, the SPEI and SPI takes values around -3 to 3, and lower values are associated with more intense droughts and higher values with excess water. The SPI and SPEI have started to be used more often in the literature of climatology as in [Spinoni et al. \(2014\)](#) or in agricultural economics as in [Branco and Féres \(2020\)](#). In this sense, [Branco and Féres \(2020\)](#) prefer to use the "SPI rather than annual fluctuations of precipitation because its probabilistic nature gives it historical context". In addition, [Spinoni et al. \(2014\)](#) prefer to use the SPI rather than the SPEI or other measures like the Palmer Drought Severity Index (PDSI) (see [Palmer \(1965\)](#)), as the SPEI can mistake a heat wave for a meteorological drought or because the PDSI needs too many input variables.

To construct the measures, I use satellite images of the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) which provide monthly rainfall from 1981 to 2016

at a higher resolution of 0.05 X 0.05 degree - around 5km x 5km at the Equator (Funk et al., 2015). As the average temperature stayed relatively stable during the period of analysis and for the municipalities of the sample (see figure in appendix 3A), I rely mainly on the SPI, without adjusting by temperature.⁷

The construction of the SPI made use of the software ArcGIS and Google Earth Engine to manipulate the raster images, calculate the averages of precipitation by municipality and by month and generate some additional maps with the help of the Python platform inside the software. The drought indices are calculated using the R-software and the package provided by Vicente-Serrano et al. (2010) and Beguería et al. (2014), which allows me to generate the SPI by months and for different locations (municipalities); then, the monthly SPI is collapsed by year. Table 3.2 gives the descriptive statistics for different types of drought measures. Here, the variable "SPI index 12 months CHIRPS" is the raw measure while the opposite of the SPI is the measure used in the estimations, so that positive values of SPI are associated with more intense droughts. As the interviews of the ELCA were conducted around February-July of 2010, 2013 and 2016, I define the calendar year 2010 as the months of February 2009 to January 2010. In fact, the rural outcome questions ask the rural household about the production during the twelve months before the interview, and defining the calendar year for this period captures much better the agricultural year.

Additionally, I also construct a monthly SPI and collapsed only for the months of the agricultural growing season in Colombia, for example, the average from February to August 2009 is the value for the year 2010. The measure is defined based on the potential growing period for many crops in Colombia.⁸ I also include some extra variables ("Per. months

⁷An alternative index could be constructed by using CHIRPS data for precipitation and temperature from MODIS-TERRA dataset from NASA (see Wan et al. (2015)). The construction of the potential evapotranspiration could add noise to the estimation, if it does not include information of wind speed and solar radiation. This gives additional arguments for why I prefer to use the SPI as the main measure to capture droughts.

⁸See FAO link <http://www.fao.org/giews/countrybrief/country.jsp?code=COL>

droughts (≤ -1) yearly for SPI index 12") as the percentage of months during the year that the municipality faced a drought or SPI less than -1. Based on [McKee et al. \(1993\)](#), a moderate drought is defined when the SPI goes below -1 while a severe drought is defined when the SPI goes below -2. Finally, "Dummy 1 month or more with drought (≤ -1)-yearly" corresponds to a dummy when there is one month or more of drought (index ≤ -1) in the municipality and zero otherwise. Appendix [3B](#) shows the SPI index using the CHIRPS data for the monthly scale 12 and 6 months.⁹ I also consider an extra variable for consecutive droughts in 2010 and 2013, as a way to measure consecutive impacts.

To sum up, all weather variables are defined according to the crop season, as this has been shown to be important in [Auffhammer and Schlenker \(2014\)](#) and [Burke and Emerick \(2016\)](#), among others.

Table [3.2](#) also includes descriptive statistics for the average rainfall and the average temperature per municipality and before 2010. The temperature is calculated using satellite images from MODIS-TERRA dataset from NASA (see [Wan et al. \(2015\)](#)) for the period 2001-2016, using the satellite MOD11A1.006 Terra Land Surface for Temperature and Emissivity Daily Global with a resolution of 0.01 X 0.01 degree - around 1km x 1km at the equator. Additionally, I calculate the yearly average of number of days with rainfall and yearly average of number of days with temperature above 34 degree Celsius. Sections [3.4.2](#) and [3.4.4](#) use the same data from satellite images described here for rainfall and temperature. I make a detailed description there on how the measures in these sections are built.

⁹Surveys were collected in the first semester of each year.

3.3 A theoretical framework for analysing adaptation of rural households in Colombia

Several empirical studies in the literature of adaptation to climate change have relied on reduced-form estimations to assess the effects of climate and weather on total gross agricultural productivity. However, in many cases adaptation is not explicitly described, making it difficult to identify the mechanisms through which weather impacts total agricultural productivity (with exceptions such as [Kaminski et al. \(2013\)](#), [Burke and Emerick \(2016\)](#)¹⁰ and [Sesmero et al. \(2018\)](#)). Several mechanisms could be at work. Weather can affect crop growth, productivity of existing inputs or prices of the outputs and inputs at the same time. It could also change the use of inputs such as land, fertilizers and pesticides and generate a re-allocation of labor. I develop a simple model based on [Cui \(2020\)](#) and [Ortiz-Bobea and Just \(2013\)](#) for a representative rural household who assigns hectares of land to different crops and can use adaptive inputs in the production. The empirical part will capture allocation of land between crop and livestock and the use of adaptive inputs such as investment in irrigation, other non-housing investments and the use of external sources of water for the farm.

The rural household in the municipality allocates land to two crops (A_1 , A_2) or prefers not to use land, which is denoted as A_3 . Crop production also depends on θ , a variable that gathers the effects of the weather conditions for agriculture (weather and climate inputs). This climate variable is one-dimensional and aims to capture "temperature" mainly. [Ortiz-Bobea and Just \(2013\)](#) explore deeper how climate might affect input and production prices. As this is out of the scope of my study, I prefer to assume a price-taking rural household, which helps me to focus more on how the rural household allocates land. The total amount of land is normalized to one and crop prices are p_1 and p_2 . The return per unit of the land

¹⁰[Burke and Emerick \(2016\)](#) define adaptation as switching from one crop variety to another (from a variety that performs better in cooler climates to another one that performs much better in warmer climates). In this sense, the farmer maximizes the expected output by choosing the type of crop variety based on climate and weather realisations.

that is not used is r and the constant marginal cost of each unit of land is s . Hence, the rural household maximizes profit as:

$$\begin{aligned} \max_{A_1, A_2, A_3, x(\theta)} \quad & p_1 q_1(A_1, \theta) + p_2 q_2(A_2, \theta, x(\theta)) - wx(\theta) + rA_3 - s(A_1 + A_2) \\ \text{s.t.} \quad & A_1 + A_2 + A_3 = 1 \end{aligned} \quad (3.1)$$

Here, crop production for crop 1 and crop 2 are defined as $q_1(A_1, \theta)$ and $q_2(A_2, \theta, x(\theta))$ and are increasing in land use A_1 and A_2 , with decreasing marginal returns with respect to land ($\frac{\partial^2 q_1(A_1, \theta)}{\partial A_1^2} < 0$ and $\frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2^2} < 0$). The adaptive input $x(\theta)$ corresponds to the inputs that the rural household use to mitigate the impact of climate on the crop production, with w as its price. This could include investments in irrigation, investments in the land other than housing (from now on called non-housing investments) such as investments in fruit trees or infrastructure, or the use of additional or external sources of water for the farm. I assume that the adaptive input only affects the production of crop 2 and that irregardless of the climate threshold that is introduced later on, it increases the marginal productivity of land for crop 2, $\frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2 \partial x(\theta)} > 0$.

With respect to climate, I assume that $\frac{\partial^2 q_1(A_1, \theta)}{\partial A_1 \partial \theta} > 0$ if $\theta < \theta_1^*$ and $\frac{\partial^2 q_1(A_1, \theta)}{\partial A_1 \partial \theta} < 0$ for $\theta > \theta_1^*$ and θ_1^* reflects the optimal climate to produce crop 1. The rural household starts producing more crop 1 for low levels of climate until reaching the optimal climate θ_1^* . Similarly, it can be assumed for crop 2 that $\frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2 \partial \theta} > 0$ if $\theta < \theta_2^*$ and $\frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2 \partial \theta} < 0$ for $\theta > \theta_2^*$, with θ_2^* the optimal climate to produce crop 2. This reflects that the impact of climate on the marginal productivity of land of crop 2 is positive before the threshold and the rural household has incentives to start to produce crop 2 as the climate variable increases (higher temperature) and up to the threshold. For the adaptive input, I assume that $\frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial x(\theta) \partial \theta} > 0$ if $\theta < \theta_x^*$ and $\frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial x(\theta) \partial \theta} < 0$ for $\theta > \theta_x^*$, with θ_x^* the threshold from which the rural household starts to use the adaptive input. The conditions for optimal land use when the rural household can modify the land left fallow (A_3) are:

$$p_1 \frac{\partial q_1(A_1^*, \theta)}{\partial A_1} = p_2 \frac{\partial q_2(A_2^*, \theta, x(\theta))}{\partial A_2} = s + r \quad (3.2)$$

$$p_2 \frac{\partial q_2(A_2^*, \theta, x(\theta))}{\partial x(\theta)} = w \quad (3.3)$$

Equation 3.2 is obtained by equalising the marginal values of land (MVL) with the cost of land $s + r$ for the two crops, while equation 3.3 results from equating the marginal value of adaptive input to its price. I calculate the total differentials for both crops from equation 3.2 and 3.3 to analyse how climate (θ) affects marginally the optimal choices of land. From equation 3.2, when MVL of crop 1 equals $s + r$:

$$\frac{dA_1^*}{d\theta} = - \frac{\frac{\partial^2 q_1(A_1, \theta)}{\partial A_1 \partial \theta}}{\frac{\partial^2 q_1(A_1, \theta)}{\partial A_1^2}} \quad (3.4)$$

As it was assumed that $\frac{\partial^2 q_1(A_1, \theta)}{\partial A_1^2} < 0$, how climate affects the hectares of land allocated to crop 1 depends on the climate impact on its marginal productivity of land (MPL). As long as climate benefits the marginal productivity of crop 1, the rural household starts to use more land available for that crop (allocate A_3 to be used in A_1). This corresponds to the result of Cui (2020). However, the effect on crop 2 depends not only on the marginal productivity of crop 2 with respect to climate, but also, on the use of the adaptive input $x(\theta)$. The total differential from equation 3.2 when the MVL of crop 2 equals $s + r$ gives:

$$\frac{dA_2^*}{d\theta} = - \frac{\frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2 \partial \theta} + \frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2 \partial x(\theta)} \frac{dx(\theta)^*}{d\theta}}{\frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2^2}} \quad (3.5)$$

As for crop 1, $\frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2^2} < 0$. In this case, how climate affects the hectares of land used for crop 2 (A_2) depends on the sign of the numerator. The numerator will be positive if first, climate (θ) benefits the MPL of crop 2 (the first term is positive), and second, if climate increases the use of the adaptive input ($\frac{dx(\theta)^*}{d\theta} > 0$ in the second term); or if the difference between both terms is positive. Notice also that the latter equation depends on

$\frac{dx(\theta)^*}{d\theta}$. I can take the total differential of equation 3.3 to know how climate affects the use of the adaptive input $x(\theta)$ in crop 2:

$$\frac{dx(\theta)^*}{d\theta} = - \frac{\frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial x(\theta) \partial \theta} + \frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial x(\theta) \partial A_2} \frac{dA_2^*}{d\theta}}{\frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial x(\theta)^2}} \quad (3.6)$$

While equation 3.4 can be treated separately, equations 3.5 and 3.6 form a system that depends on $\frac{dx(\theta)^*}{d\theta}$ and $\frac{dA_2^*}{d\theta}$. I use Cramer's rule to solve the system and give the effect of climate (θ) on the use of land for crop 2 and the adaptive input:

$$\frac{dA_2^*}{d\theta} = \frac{\frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2 \partial \theta} \frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial x(\theta)^2} - \frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial x(\theta) \partial \theta} \frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2 \partial x(\theta)}}{\left(\frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2 \partial x(\theta)} \right)^2 - \frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2^2} \frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial x(\theta)^2}} \quad (3.7)$$

and,

$$\frac{dx(\theta)^*}{d\theta} = \frac{\frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial x(\theta) \partial \theta} \frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2^2} - \frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2 \partial \theta} \frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2 \partial x(\theta)}}{\left(\frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2 \partial x(\theta)} \right)^2 - \frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2^2} \frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial x(\theta)^2}} \quad (3.8)$$

Given the previous assumptions of the production functions for crops, in equations 3.7 and 3.8, $\frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial x(\theta)^2} < 0$, $\frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2^2} < 0$, and $\frac{\partial^2 q_2(A_2, \theta, x(\theta))}{\partial A_2 \partial x(\theta)} > 0$. I use the thresholds over climate defined previously to determine the effect of climate on the use of land for crop 2 and the adaptive input. Those thresholds determine the point until which the land use of the crops is optimal and also the optimal use of the adaptive input. As the denominator in equations 3.7 and 3.8 is assumed negative to have a solution and given the previous assumptions, two cases¹¹ are particularly interesting to explore:

Case A: $\theta_1^* < \theta_2^* < \theta_x^*$

¹¹Other cases such as $\theta_2^* < \theta_1^* < \theta_x^*$ or $\theta_x^* < \theta_1^* < \theta_2^*$ are not relevant as I expect that crop 2 is more resilient to extreme climate than crop 1 and the use of adaptive input becomes an option in the production of crop 2 with higher values of the variable of climate, affecting crop 2 and not crop 1.

1. $\theta \leq \theta_1^*$ which gives as result $\frac{dA_1^*}{d\theta} > 0$, $\frac{dA_2^*}{d\theta} > 0$ and $\frac{dx(\theta)^*}{d\theta} > 0$
2. $\theta_1^* < \theta \leq \theta_2^*$ which gives as result $\frac{dA_1^*}{d\theta} < 0$, $\frac{dA_2^*}{d\theta} > 0$ and $\frac{dx(\theta)^*}{d\theta} > 0$
3. $\theta_2^* < \theta \leq \theta_x^*$ which gives as result $\frac{dA_1^*}{d\theta} < 0$, $\frac{dA_2^*}{d\theta} \begin{matrix} \geq \\ \leq \end{matrix} 0$ and $\frac{dx(\theta)^*}{d\theta} \begin{matrix} \geq \\ \leq \end{matrix} 0$
4. $\theta_x^* < \theta$ which gives as result $\frac{dA_1^*}{d\theta} < 0$, $\frac{dA_2^*}{d\theta} < 0$ and $\frac{dx(\theta)^*}{d\theta} < 0$

Case B: $\theta_1^ < \theta_x^* < \theta_2^*$*

1. $\theta \leq \theta_1^*$ which gives as result $\frac{dA_1^*}{d\theta} > 0$, $\frac{dA_2^*}{d\theta} > 0$ and $\frac{dx(\theta)^*}{d\theta} > 0$
2. $\theta_1^* < \theta \leq \theta_x^*$ which gives as result $\frac{dA_1^*}{d\theta} < 0$, $\frac{dA_2^*}{d\theta} > 0$ and $\frac{dx(\theta)^*}{d\theta} > 0$
3. $\theta_x^* < \theta \leq \theta_2^*$ which gives as result $\frac{dA_1^*}{d\theta} < 0$, $\frac{dA_2^*}{d\theta} \begin{matrix} \geq \\ \leq \end{matrix} 0$ and $\frac{dx(\theta)^*}{d\theta} \begin{matrix} \geq \\ \leq \end{matrix} 0$
4. $\theta_2^* < \theta$ which gives as result $\frac{dA_1^*}{d\theta} < 0$, $\frac{dA_2^*}{d\theta} < 0$ and $\frac{dx(\theta)^*}{d\theta} < 0$

Both cases give the same results, independently of whether the optimal climate for the adaptive input is smaller or larger than the optimal climate for the land used for crop 2. The results show that when $\theta \leq \theta_1^*$, there are incentives for the rural household to increase the land used for crop 1, the land for crop 2 and the use of the adaptive input. Notice that the situation is feasible as the rural household has land not used in the production (A_3). Once the optimal climate for crop 1 is crossed and $\theta_1^* < \theta \leq \theta_2^*$ in case A (or $\theta_1^* < \theta \leq \theta_x^*$ in case B), the rural household reduces the land used for crop 1, but keep using more land for crop 2 and using the adaptive input. As climate increases and reaches the optimal threshold to use the adaptive input in case A $\theta_x^* < \theta$ (or the threshold for crop 2 $\theta_2^* < \theta$, in case B), the results are undetermined. As climate exceeds the thresholds of land to use for crop 2 and for the adaptive input $\theta_2^* < \theta_x^* < \theta$ for case A (or $\theta_x^* < \theta_2^* < \theta$ for case B), the rural household reduces the use of land for crops 1 and 2, as well as the use of the adaptive input to produce crop 2. Notice also that $\theta_x^* = \theta_2^*$ is a particular sub-case and gives the same results as in case A and case B (subcategories 1, 2 and 4). Although this does not change the interpretation of the findings, it could be more realistic to assume that the threshold

of climate to produce crop 2 is the same as the optimal value to start using the adaptive input.

For case A, when $\theta < \theta_1^*$ and climate increases, the rural household can adapt by using more land available in the production of the farm. If climate increases up to the threshold for crop 2 (θ_2^*), the rural household has incentives to incorporate the adaptive input and focuses mainly on the production of crop 2, with the optimal level of climate for its production determined by θ_2^* for case A and θ_x^* for case B. Once that thresholds are exceeded in each case, climate is far from the optimal and becomes too extreme, affecting the production of crops and the use of the adaptive input. To some extent, some of the strategies used before (increase the land used or using more adaptive inputs) might be harder to implement as they are less profitable for the rural household. This could be the case when the rural household faces extreme droughts or consecutive droughts, which is something I explore in the empirical part.

3.4 Effects of climate and weather deviations on agricultural productivity

I study the effects of climate and weather on agriculture using three different frameworks. First, I start by separating the effect of climate from yearly weather deviations as in previous studies (Kelly et al., 2005; Deschênes and Kolstad, 2011; Burke and Emerick, 2016; Bento et al., 2020). The results of this can be found in section 3.4.2. Second, I explore in more detail how rural households adapt on different adaptation margins in section 3.4.4, using different temperature thresholds. While section 3.4.2 analyses medium-term and short-term variation in climate and weather, section 3.4.4 relies more on weather variation. Also, section 3.4.2 and section 3.4.4 permit me to have a comparison with the established literature on the effects of climate change on agriculture. Third, in section 3.5, I propose an alternative approach to study the effects of climate variability on agriculture. I explore if those affected by previous

drought events adapt differently to a drought than those who had not experienced drought in the past (the DID approach). This section uses the SPI indices which aim to capture anomalies with respect to the long-term mean.

3.4.1 Identification Strategy for climate and yearly weather deviations

Burke and Emerick (2016) and Dell et al. (2014), among others, estimate adaptation as the difference between the coefficients estimated using panel data (weather shocks) and the coefficient of a cross-section estimation (climate trends). As mentioned by Blanc and Schlenker (2017), panel data allow for a better causal identification on weather impacts on agriculture, as weather deviations around the mean are random and exogenous. Although, panel models can solve identification problems of cross-sectional approaches, this comes with the cost of poorly approximating the impact of climate change (see Burke and Emerick (2016)). If adaptation is defined as in Burke and Emerick (2016) (switching from a less-tolerant to a more heat-tolerant crop) and a temperature shock occurs during the years of the panel, the estimated coefficient of the impact will be weighted depending on the time that the adaptation happened. If adaptation occurs at the end of the sample, the estimation will weight more the less-tolerant and conventional crop, magnifying equilibrium losses. Bento et al. (2020) and Kolstad and Moore (2020) also point out that the cross-section coefficient can suffer from omitted variable bias; and even in the case that the omitted variable bias does not exist, the coefficients (the one from the cross-section and the one from the panel) come from two different estimating equations. I follow previous studies in the literature in equation 3.9 that capture climate trends and weather shocks (deviations from those long-term patterns) in the same estimations (Kelly et al., 2005; Deschênes and Kolstad, 2011), in particular Bento et al. (2020)¹². I use this framework to test the effects of climate and weather deviations on gross agricultural productivity (in value terms) in Colombia, which corresponds to a long-term and short-term analysis.

¹²See also Kolstad and Moore (2020) for a detailed description of other "partitioning approaches" that capture climate and weather together in the same estimations.

$$\begin{aligned}
\mathbf{Y}_{h,m,r,t} = & \mu + \beta_{\mathbf{C}} \times \mathbf{TEMP}^{\mathbf{C}}_{m,t} + \beta_{\mathbf{W}} \times \mathbf{TEMP}^{\mathbf{W}}_{m,t} + \Gamma_{\mathbf{C}} \times \mathbf{RAIN}^{\mathbf{C}}_{m,t} + \Gamma_{\mathbf{W}} \times \mathbf{RAIN}^{\mathbf{W}}_{m,t} \\
& + \theta_x \times X_{h,m,r,t} + \alpha_r + \alpha_t + \alpha_m + \alpha_h + \xi_{h,m,r,t}
\end{aligned} \tag{3.9}$$

$\mathbf{Y}_{h,m,r,t}$ agriculture outcomes for rural household h (or agricultural unit), of municipality m , region r in year t ;

$\mathbf{TEMP}^{\mathbf{C}}_{m,t}$ represents the lagged 30-year annual moving average (MA) of past temperatures;

$\mathbf{TEMP}^{\mathbf{W}}_{m,t}$ represents weather shocks and is defined as the deviation of the annual average of the daily temperature in year t from the lagged 30-year annual MA of past temperatures;

$\mathbf{RAIN}^{\mathbf{C}}_{m,t}$ represents the lagged 30-year annual MA of past rainfall;

$\mathbf{RAIN}^{\mathbf{W}}_{m,t}$ represents rainfall shocks and is defined as the deviation of the annual average of the daily rainfall in year t from the lagged 30-year annual MA of past rainfall;

$X_{h,m,r,t}$ vector of controls for household and community characteristics¹³;

$\alpha_r, \alpha_t, \alpha_m, \alpha_h$ dummies by *region*, by *year*, by *municipality* and by *household*.

The standard errors are clustered at the municipality level m to account for the correlation within municipality and the estimations use the inverse weights of 2010. I include household fixed effects to account for potential differences in skills of the rural household, or the fact that some of them can be of different types (subsistence versus commercial rural households). The municipality and region fixed effects aim to control for differences in the regional and municipal characteristics, such as the altitude or the fact that the ELCA was designed to collect the rural information representative for some specific regions. The coefficients of

¹³Controls for the "Dependency ratio (<=14 + >=65)/(15-64) Per. 2010" of the household in 2010, "Age HH head" for the age of the household head, a "Dummy women HH head" for whether or not the household head is a woman, a "Dummy HH has credit" for whether or not the household has a credit in that year, a "Dummy access cabecera is ok (comm)" at the community level for whether or not the access to the cabecera (center of the town) is reachable and in good shape, the "Minutes to reach cabecera (comm)" for the community, and some dummies for reported problems faced by the household in that year

interest for temperature are $\beta_{\mathbf{C}}$, $\beta_{\mathbf{W}}$, and for rainfall $\Gamma_{\mathbf{C}}$ and $\Gamma_{\mathbf{W}}$. Equation 3.9 captures both slow-moving cross-section climate effect, and weather deviations. As such, equation 3.9 relies on estimating a slow-moving cross sectional climate effect (the coefficient $\beta_{\mathbf{C}}$) similar to what has been done in studies such as Deschênes and Kolstad (2011).

3.4.2 Results of climate and yearly weather deviations

Table 3.4 summarizes the results of the model estimated using the logarithm of the value of the total production (Millions Colombian Pesos/Ha)¹⁴ explained by the meteorological variables. I use the daily rainfall and temperature data from CHIRPS and MODIS-TERRA satellite images, and I average by month and by year, and then, construct the lagged moving average for 30 years. Rainfall monthly data come from CHIRPS satellite images that are available from 1981 to 2016. However, temperature daily data come from MODIS-TERRA dataset for the period 2001 to 2016 and the measures of long-term temperature constructed here thus capture a shorter period of time. Although some data exist for temperature to capture longer periods of time in the past, I prefer to use the MODIS-TERRA dataset because it has a higher level of resolution (0.01 X 0.01 degree -around 1km x 1km at the equator) that allows to capture much better temperature variability. Using a higher resolution comes with the cost of reducing the time variation, but given the size of the municipalities in Colombia, it should be more pertinent to weight more a better spatial variability and resolution rather than having a longer period of time. Otherwise, I could risk assigning the same value of temperature to many municipalities if data with more time variation and less spatial variation are used. For Kolstad and Moore (2020), some panel approaches use the fact that climate varies over time in a location to estimate short-run and long-run effects in the same panel by including location and period fixed effects. However, the interpretation of the climate term can change, capturing more a medium-run effect. According to Kolstad and Moore

¹⁴As the rural households of the ELCA sample reported to cultivate several crops on their farms, it would be harder to construct a variable to capture agricultural production in terms of units of production. For this reason, I use the variables of total gross agricultural productivity in values (see section 3.2), as they allow me to sum up all the different crops in values and have a comparable measure for all the rural households.

(2020), in practice this appears not to matter substantially when the studies use long spans of time of 30 years for example. As in this study the temperature variables were constructed using a shorter period of time, the climate trend should be interpreted as a medium-term impact. Also note that all the estimations for the gross agricultural productivity (in value terms) are restricted to the sample of crop producers only; the variables then capture the gross agricultural productivity of crops. Livestock producers and those producing crop and livestock at the same time are excluded, as it was not possible to recover livestock prices to calculate the income coming from livestock production.

The agricultural year (AGRO YEAR) 2010 is defined from February to August 2009 to capture better the agricultural growing season. I calculate an extra variable for the maximum during the month when I move from the daily to the yearly data. In addition, the histogram of the average temperature shows a bi-modal distribution, with municipalities of high-temperature (above 27 degrees Celsius) and low-temperature (below 27 degrees Celsius) (see Figure 3.4 in Appendix 3A).¹⁵ Therefore, I interact the variables of temperature with a dummy for municipalities with high and low-temperature in order to see if there are differences between both groups.¹⁶

Table 3.4 shows that rainfall trend (AGRO YEAR) and rainfall shocks are beneficial for productivity.¹⁷ Regarding temperature, the long-run mean increases the gross agricultural productivity in columns 1) and 3), while temperature shocks have no effect on productivity

¹⁵In Appendix 3D, I check if temperature variables could be modelled in squared terms. The squared terms seem to capture some of the relation between temperature and agricultural productivity but the results of column 1) and 4) capture much better the bi-modal distribution of Figure 3.4 in Appendix 3A. I also include both temperature squared and the interaction of temperature variables with the dummy of high-temperature in columns 3) and 6), to see which one dominates the other. However, the coefficients for the low-temperature municipalities are washed out of the estimations. In this regard, I prefer to use the dummy of high-temperature municipalities as it resembles better the bi-modal distribution of temperature.

¹⁶High-temperature group: Cereté, Chinú, Ciénaga de Oro, Sahagún, Tocaima, Sampedra, Natagaima, Ortega and Purificación. Low-temperature group: Saboyá, Simijaca, Susa, Circasia, Córdoba, Filandia, Belén de Umbía and Puente Nacional.

¹⁷The large values estimated could be due to the large variations and should be interpreted cautiously as the maximum possible percentage increase is 100%. Log-Lin Models are recommended when the variations are smaller, less than 10% for example.

in the same estimations. When moving to the estimations with the interaction (columns 2 and 4), in terms of the marginal effects, the increase in the long-run temperature has a positive effect for rural households in municipalities with low-temperature, but no effect for rural households in high-temperature municipalities. However, the marginal effects of the temperature shocks decrease the gross agricultural productivity (in value terms) in high-temperature municipalities, with the net effect (the sum of the marginal effects for low and high-temperature) still negative. This divergence is a new interesting point in terms of the winners or losers from climate and weather shocks. It also has some similarities with [Aragón et al. \(2021\)](#) who find that the coast areas suffer losses for the higher temperatures, while the highlands, with a cooler and wetter climate, would benefit from warmer temperatures. [Bento et al. \(2020\)](#) consider as evidence of adaptation if the difference between the temperature shock and climate is positive.¹⁸ Using this, column 2 and 4 show evidence of adaptation by the rural households facing high temperature, although not statistically significant at conventional levels. However, caution has to be taken in the interpretation, as the variable used is agricultural production in values and it could mix price and volume effects. As mentioned previously, it is harder to construct measures of total production for the rural households of the sample, who cultivate multiple crops.

In order to go further, [Table 3.5](#) explores climate and weather deviations by groups of crops, using the maximum of temperature during the agricultural growing season (AGRO YEAR). Long-run rainfall and the deviations with respect to the long-run seem to be more beneficial for the gross agricultural productivity (in value terms) of vegetables, cereals and coffee. In terms of temperature, cereals and coffee are the groups that benefit from higher temperature in the long-run and the deviations (columns 1 to 4). The results using the interaction in columns 5 to 8 show that the marginal effect of the long-run temperature

¹⁸Other studies such as [Burke and Emerick \(2016\)](#) define adaptation for the case of U.S. agriculture as the difference between a long difference estimate and the estimate from an annual panel model. A positive difference should be interpreted as adaptation, as farmers are adjusting to long-run changes in climate compared to shorter run changes in weather.

affects positively fruits of high-temperature areas, while it damages those in low-temperature areas. The marginal effects of the long run temperature and its deviations increase the gross agricultural productivity of cereals and coffee in low-temperature municipalities, but reduces productivity in high-temperature areas. Interestingly, the net effect (the sum of the low and high-temperature coefficients) of the long run temperature and its deviations are close to zero for fruits and coffee in high-temperature municipalities; only for cereals, the net effect of the long-run mean of temperature is still positive in high-temperature municipalities, and negative with respect to the temperature shocks. There are no statistically significant marginal effects of temperature on the gross agricultural productivity of vegetables. A comparison by type of crops, perennial (fruits and coffee), and annual (vegetables and cereals) yields some interesting results. On the one hand, I should expect that perennial crops in general generate more stable yields than annual crops, and correspond to agricultural production that would be less affected by weather variation. In Table 3.5, I observe that the marginal effect on perennial crops differs based on the high versus low-temperature municipalities and the type of crop. In low-temperature municipalities, it benefits coffee while reducing the gross agricultural productivity of fruits; in high-temperature municipalities the marginal effect is the opposite and benefits fruits while reducing the gross agricultural productivity of coffee. On the other hand, it should be expected that annual crops are associated with less stable yields, and the agricultural productivity would be more affected by weather variation. However, the results of the marginal effects show in fact that some annual crops such as cereals benefit from higher temperature (long run and shocks). Similar to the case of coffee, they benefit in low-temperature municipalities while the gross agricultural productivity decreases in municipalities with high temperature. It appears that cereals and coffee could take advantage of the higher temperatures in areas at high altitude, which are in general the municipalities with lower temperatures. On the contrary, fruits can take advantage of the higher temperatures in municipalities with high temperature.

Table 3.4: Climate and Yearly weather deviations

	Value prod. total (Millions Col. Pesos/Ha) ln			
	MEAN TEMPERATURE		MAXIMUM TEMPERATURE	
	NO INTERACTION (1)	INTERACTION (2)	NO INTERACTION (3)	INTERACTION (4)
Rainfall Trend 30 years (MEAN) (AGRO YEAR)	4.362** (1.621)	4.054** (1.798)	4.242*** (1.371)	4.358*** (1.446)
Rainfall Shock (MEAN) (AGRO YEAR)	0.305** (0.119)	0.250** (0.094)	0.331*** (0.092)	0.285*** (0.086)
Temperature Trend 15 years (MEAN) (AGRO YEAR)	1.287** (0.556)	2.208*** (0.645)		
Temperature Shock (MEAN) (AGRO YEAR)	0.032 (0.085)	0.091 (0.083)		
High-Temp=1 × Temperature Trend 15 years (MEAN) (AGRO YEAR)		-1.202 (0.811)		
High-Temp=1 × Temperature Shock (MEAN) (AGRO YEAR)		-0.204*** (0.061)		
Temperature Trend 15 years (MAX) (AGRO YEAR)			1.108*** (0.337)	1.652*** (0.520)
Temperature Shock (MAX) (AGRO YEAR)			0.060 (0.043)	0.104** (0.036)
High-Temp=1 × Temperature Trend 15 years (MAX) (AGRO YEAR)				-0.695 (0.619)
High-Temp=1 × Temperature Shock (MAX) (AGRO YEAR)				-0.169*** (0.057)
Observations	1770	1770	1770	1770
r2	0.62	0.63	0.62	0.63
r2_a	0.30	0.30	0.30	0.30

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: based on the ELCA panel survey 2010-2013-2016. The standard errors (in parentheses) are clustered at the municipality level m to account for the correlation in treatment within municipality and the estimations use the inverse weights of 2010. Includes *region*, *year*, *municipality* and *household* fixed effects as well as agricultural problems of the household (except drought or other weather-related problems). The estimations also include controls for the dependency ratio in 2010, the age of the household head, a dummy for whether or not the household head is a woman, a dummy for whether or not the household has a credit in that year, a dummy at the community level for whether or not the access to the cabecera (center of the town) is reachable and in good shape, and the minutes to reach the cabecera for the community. High-temp is a dummy equal to one for the municipalities with average temperature higher than 27 Celsius degrees, and zero for the other municipalities. Sample restricted to rural households producing uniquely crops (excludes livestock producers or those producing livestock and crops).

Table 3.5: Climate and Yearly weather deviations-
Maximum monthly Temperature- Agricultural Growing Season

	Value prod. total (Millions Col. Pesos/Ha) ln							
	NO INTERACTION				INTERACTION			
	Vegetables (1)	Fruits (2)	Cereals (3)	Coffee (4)	Vegetables (5)	Fruits (6)	Cereals (7)	Coffee (8)
Rainfall Trend 30 years (MEAN) (AGRO YEAR)	3.690*	0.660	1.023	1.553	3.968**	-0.152	2.338*	1.233
	(1.764)	(1.788)	(1.283)	(0.963)	(1.648)	(1.312)	(1.325)	(0.967)
Rainfall Shock (MEAN) (AGRO YEAR)	0.042	-0.022	0.190*	0.188**	0.079	-0.007	0.183**	0.134**
	(0.102)	(0.098)	(0.103)	(0.068)	(0.081)	(0.077)	(0.072)	(0.056)
Temperature Trend 15 years (MAX) (AGRO YEAR)	-0.419	-0.342	1.546**	0.362**	-0.445	-1.391***	3.089***	0.494
	(0.568)	(0.259)	(0.672)	(0.150)	(0.501)	(0.317)	(0.728)	(0.308)
Temperature Shock (MAX) (AGRO YEAR)	-0.069	-0.162**	0.137**	0.092**	-0.094	-0.199***	0.181***	0.132***
	(0.056)	(0.064)	(0.050)	(0.036)	(0.063)	(0.057)	(0.038)	(0.040)
High-Temp=1 × Temperature Trend 15 years (MAX) (AGRO YEAR)					0.012	1.385***	-2.047***	-0.141
					(0.849)	(0.312)	(0.685)	(0.344)
High-Temp=1 × Temperature Shock (MAX) (AGRO YEAR)					0.076	0.186***	-0.245***	-0.124***
					(0.062)	(0.041)	(0.057)	(0.026)
Observations	1770	1770	1770	1770	1770	1770	1770	1770
r2	0.63	0.51	0.72	0.73	0.63	0.53	0.73	0.73
r2_a	0.31	0.09	0.47	0.49	0.31	0.12	0.49	0.50

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: based on the ELCA panel survey 2010-2013-2016. The standard errors (in parentheses) are clustered at the municipality level m to account for the correlation in treatment within municipality and the estimations use the inverse weights of 2010. Includes *region*, *year*, *municipality* and *household* fixed effects as well as agricultural problems of the household (except drought or other weather-related problems). It also includes controls for the dependency ratio in 2010, the age of the household head, a dummy for whether or not the household head is a woman, a dummy for whether or not the household has a credit in that year, a dummy at the community level for whether or not the access to the cabecera (center of the town) is reachable and in good shape, and the minutes to reach the cabecera for the community. High-temp is a dummy equal to one for the municipalities with average temperature higher than 27 Celsius degrees, and zero for the other municipalities. Sample restricted to rural households producing uniquely crops (excludes livestock producers or those producing livestock and crops).

3.4.3 Identification strategy for nonlinear effects of temperature

I follow [Aragón et al. \(2021\)](#) to estimate the effect of weather on gross agricultural productivity as a function of cumulative exposure to temperature and rainfall. In particular, the relationship is modeled by using two measures: the Growing Degree Days (GDD) and the Harmful Degree Days (HDD) (see [Deschênes and Greenstone \(2007\)](#) and [Schlenker and Roberts \(2009\)](#)). GDD measures the cumulative exposure to temperatures between a lower bound θ_{low} and an upper threshold θ_{high} ,¹⁹ while HDD captures non-linear exposure to extreme temperature (above θ_{high}). With respect to how to model temperature, [Deschênes and Greenstone \(2007\)](#) note that the standard agronomic approach is to convert daily temperatures into degree-days, which represent heating units. In this sense, there is a nonlinear

¹⁹Studies like [Schlenker and Roberts \(2009\)](#) use a lower bound of 8 degree Celsius for corn and soybeans.

cumulative effect of heat accumulation where temperature must be above a threshold for plants to absorb heat and below a ceiling as plants cannot absorb extra heat when temperature is too high. While the analysis made in 3.4.2 decomposes the effect of climate variability of temperature in terms of climate trends and weather shocks, the GDD and HDD might capture better the nonlinear effects of temperature on agricultural production and productivity. It is important to notice that the GDD and HDD measures exploit more the yearly weather variation rather a long-term variation, in comparison to section 3.4.2 that explores climate and weather variation (long-term and short-term variation).

GDD is defined as: $GDD = \frac{1}{n} \sum_d g^{GDD}(h_d)$, with,

$$g^{GDD}(h) = \begin{cases} 0, & \text{if } h \leq \theta_{low} \\ h - \theta_{low}, & \text{if } \theta_{low} < h \leq \theta_{high} \\ \theta_{high} - \theta_{low}, & \text{if } \theta_{high} < h \end{cases}$$

With h_d the daytime temperature in day d and n the number of days during the calendar year or agricultural growing year. The calendar year and agricultural year are defined as in section 3.4.2. Similarly, HDD is defined as: $HDD = \frac{1}{n} \sum_d g^{HDD}(h_d)$, with,

$$g^{HDD}(h) = \begin{cases} 0, & \text{if } h \leq \theta_{high} \\ h - \theta_{high}, & \text{if } \theta_{high} < h \end{cases}$$

The variable of rainfall corresponds to the average daily precipitation and then, it is averaged by year or agricultural growing year. I include rainfall and rainfall squared in the estimations, to capture non-linear effects and also, to take into account the correlation between temperature and rain. The function relating weather to production corresponds to the variables of rainfall and temperature in the next equation:

$$\begin{aligned}
\mathbf{Y}_{h,m,r,t} = & \mu + \lambda_1 \times \mathbf{GDD}_{m,t} + \lambda_2 \times \mathbf{HDD}_{m,t} + \lambda_3 \times \mathbf{RAIN}_{m,t} + \lambda_4 \times \mathbf{RAIN}_{m,t}^2 \\
& + \theta_x \times X_{h,m,r,t} + \alpha_r + \alpha_t + \alpha_m + \alpha_h + \epsilon_{h,m,r,t}
\end{aligned}
\tag{3.10}$$

This corresponds to a non-linear version of equation 3.9, using the same control variables and fixed effects. However, the weather factors are measured with the variables GDD and HDD for temperature, while rainfall is the daily average during the year (or AGRO YEAR) and its square. In this sense, equation 3.10 captures weather variation rather than climate variation. This is a difference compared to equation 3.9 which analyses in the same estimation, weather and climate variation.

3.4.4 Results for nonlinear effects of temperature

An important aspect of the specification in equation 3.10 is to define the upper threshold θ_{high} , above which temperature starts to have a negative impact on production. In studies like Deschênes and Greenstone (2007) or Schlenker and Roberts (2009) for the U.S. context, this threshold is defined as 29-32 degree Celsius and it is crop and location-dependent.

I follow here Aragón et al. (2021) who estimates equation 3.10, replacing GDD and HDD by a set of temperature bins that measure the proportion of days in the calendar year (or AGRO YEAR) on which the temperature fell in a given temperature bracket. Using the distribution of temperature for my sample (see Figure 3.4 in Appendix 3A), I define 9 temperature bins: <15C, 15-18C, 18-21C, 21-24C, 24-27C, 27-30C, 30-33C, 33-36C and >36C, with bin 24-27C the baseline category in the estimations.²⁰ Figure 3.1 presents the estimation of the coefficients with those bins with a confidence interval at the 95% level. It uses the average temperature during the AGRO YEAR (agricultural growing season). The results show that temperature values above bin 8 (33-36C) start to affect negatively the values of

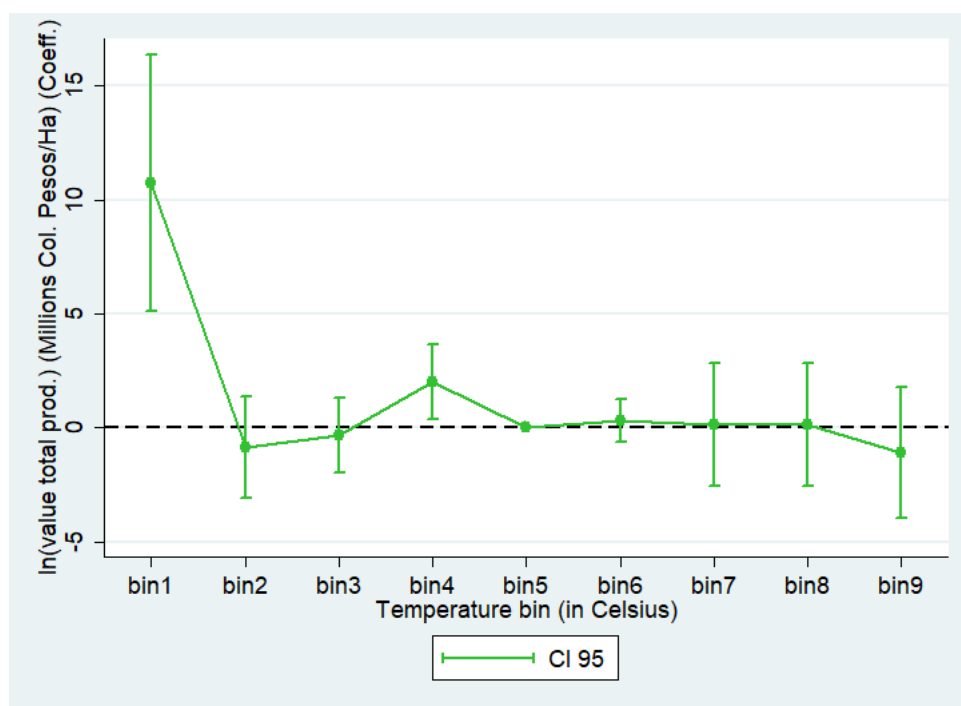
²⁰Rainfall monthly data come from CHIRPS satellite images that expand from 1981 to 2016. Temperature daily data comes from MODIS-TERRA dataset for the period 2001 to 2016.

the total production. With this in mind, I determine the upper threshold for my sample as a temperature of 36 degree Celsius and I use it to calculate the measures of GDD and HDD.²¹

Table 3.6 shows the results of the estimation of equation 3.10, using the GDD and HDD for the logarithm of the value of the total gross agricultural productivity. Columns 1) and 3) show the results using the mean and the maximum of the temperature without the interaction with the dummy for the high-temperature municipalities. Across all the specifications in the table, there is no statistically significant effect of rainfall or its square or of the GDD-HDD, which could be due to the small sample size (or because the upper threshold of 36 degree Celsius is too high for some municipalities). When analysing the estimations with the interaction of column 2), one extra HDD decreases the gross agricultural productivity in low-temperature municipalities by 15.5%. Regarding the marginal effects, Column 4) shows no negative effect of HDD or GDD on the low or high-temperature municipalities. In column 4), the marginal effects shows that the HDD decrease the total productivity in the high-temperature municipalities but increase it in the low-temperature municipalities with the results no statistically significant at conventional levels. However, these results go in line with the findings in section 3.4.2 where the high-temperature municipalities were more affected by higher temperatures.

²¹It is important to notice that the daily measures of temperature for the calculation of the GDD and HDD have several days in which there was no information recorded.

Figure 3.1: Coefficients estimated for the temperature bins - Average agricultural growing season



Note: estimations based on the ELCA panel survey 2010-2013-2016 and the weather data. The temperature bins are (in degree celsius): <15C, 15-18C, 18-21C, 21-24C, 24-27C, 27-30C, 30-33C, 33-36C and >36C. The baseline category in the estimation is bin 5 (24-27C). Confidence intervals at the 95% level. Sample restricted to rural households producing uniquely crops (excludes livestock producers or those producing livestock and crops).

In Table 3.7, I explore further the effect of the GDD and HDD on the value of the total production by group of crop and using the maximum of temperature during the agricultural growing season. The increase in rainfall tends to have a positive effect on the value of the production of vegetables (columns 1 and 5 first line), with results statistically significant at the 1% level in column 1). There is a statistically significant effect of rainfall on coffee in column 4). Interestingly, rainfall tends to increase the gross agricultural productivity for cereals and coffee but rainfall squared decreases it. This indicates that these crops might reach a maximum up to which the level of rainfall can benefit the gross agricultural productivity but after some level of rainfall, the extra rainfall can damage the productivity. Columns 1) to 4) show the results of the effect of GDD and HDD without the interaction. There are no statistically significant effects on the different groups of crops.

When separating the effects of GDD and HDD by high and low-temperature municipalities in Table 3.7, the marginal effects of columns (5) to (8) show that HDD can damage the productivity of vegetables and fruits in low municipality areas while it benefits them in high-temperature municipalities. On the contrary, the marginal effects show that HDD can damage the productivity of cereals and coffee in high-temperature municipalities while it benefits them in low-temperature municipalities. Interestingly, the net effect of the HDD in the high-temperature municipalities is close to zero across the different crops. The analysis of the marginal effects are similar to the findings in section 3.4.2 where the agricultural productivity of fruits in high-temperature municipalities benefit from higher temperature, while cereals and coffee are the winners in low-temperature areas. An additional GDD seems to benefit more cereals and coffee in high-temperature municipalities while damaging the same crops in low-temperature municipalities. In some sense, cereals and coffee in low-temperature municipalities correspond to the group of crop 2 in the theoretical framework of section 3.3. The findings of the model show that crop 2 benefits from higher values of climate (higher temperature), compared to crop 1.

Table 3.6: GDD and HDD for temperature

	Value prod. total (Millions Col. Pesos/Ha) ln			
	MEAN TEMPERATURE		MAXIMUM TEMPERATURE	
	NO INTERACTION (1)	INTERACTION (2)	NO INTERACTION (3)	INTERACTION (4)
Rainfall (MEAN) (AGRO YEAR)	0.337 (0.288)	0.204 (0.440)	0.075 (0.330)	0.254 (0.310)
Rainfall (MEAN) (AGRO YEAR) × Rainfall (MEAN) (AGRO YEAR)	-0.029 (0.023)	-0.018 (0.034)	-0.008 (0.025)	-0.023 (0.024)
Temp GDD (average AGRO YEAR)	-0.034 (0.042)	0.009 (0.113)		
Temp HDD (average AGRO YEAR)	-0.203 (0.210)	-0.169 (0.225)		
High-Temp=1 × Temp GDD (average AGRO YEAR)		-0.080 (0.198)		
High-Temp=1 × Temp HDD (average AGRO YEAR)		0.000 (.)		
Temp GDD (MAX) (AGRO YEAR)			0.020 (0.066)	-0.099 (0.110)
Temp HDD (MAX) (AGRO YEAR)			-0.244 (0.156)	16.608 (11.366)
High-Temp=1 × Temp GDD (MAX) (AGRO YEAR)				0.338 (0.213)
High-Temp=1 × Temp HDD (MAX) (AGRO YEAR)				-17.069 (11.411)
Observations	1770	1770	1770	1770
r2	0.62	0.62	0.62	0.63
r2_a	0.30	0.30	0.30	0.30

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: based on the ELCA panel survey 2010-2013-2016. The standard errors (in parentheses) are clustered at the municipality level m to account for the correlation in treatment within municipality and the estimations use the inverse weights of 2010. Includes *region*, *year*, *municipality* and *household* fixed effects as well as agricultural problems of the household (except drought or other weather-related problems). It also includes controls for the dependency ratio in 2010, the age of the household head, a dummy for whether or not the household head is a woman, a dummy for whether or not the household has a credit in that year, a dummy at the community level for whether or not the access to the cabecera (center of the town) is reachable and in good shape, and the minutes to reach the cabecera for the community. High-temp is a dummy equal to one for the municipalities with average temperature higher than 27 Celsius degrees, and zero for the other municipalities. Sample restricted to rural households producing uniquely crops (excludes livestock producers or those producing livestock and crops).

Table 3.7: GDD and HDD for temperature-
Maximum monthly Temperature- Agricultural Growing Season

	Value prod. total (Millions Col. Pesos/Ha) ln							
	NO INTERACTION				INTERACTION			
	Vegetables (1)	Fruits (2)	Cereals (3)	Coffee (4)	Vegetables (5)	Fruits (6)	Cereals (7)	Coffee (8)
Rainfall (MEAN) (AGRO YEAR)	-0.385 (0.220)	0.009 (0.344)	0.126 (0.285)	0.236* (0.129)	-0.359 (0.247)	-0.169 (0.514)	0.428*** (0.141)	0.275* (0.133)
Rainfall (MEAN) (AGRO YEAR) × Rainfall (MEAN) (AGRO YEAR)	0.049*** (0.016)	0.013 (0.029)	-0.029 (0.022)	-0.023* (0.013)	0.047*** (0.019)	0.028 (0.043)	-0.054*** (0.013)	-0.027* (0.013)
Temp GDD (MAX) (AGRO YEAR)	0.077 (0.109)	-0.008 (0.104)	-0.047 (0.069)	0.037 (0.050)	0.084 (0.111)	0.106 (0.197)	-0.247*** (0.072)	-0.001 (0.068)
Temp HDD (MAX) (AGRO YEAR)	-0.190* (0.105)	0.029 (0.158)	-0.018 (0.147)	-0.107 (0.068)	-68.856*** (19.097)	-2.378 (21.720)	25.238** (9.184)	38.303*** (10.809)
High-Temp=1 × Temp GDD (MAX) (AGRO YEAR)					0.178 (0.195)	-0.364 (0.320)	0.578*** (0.184)	0.011 (0.100)
High-Temp=1 × Temp HDD (MAX) (AGRO YEAR)					68.487*** (19.103)	2.654 (21.743)	-25.630** (9.255)	-38.386*** (10.834)
Observations	1770	1770	1770	1770	1770	1770	1770	1770
r2	0.62	0.49	0.70	0.73	0.63	0.49	0.71	0.73
r2_a	0.30	0.04	0.45	0.49	0.30	0.05	0.46	0.49

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: based on the ELCA panel survey 2010-2013-2016. The standard errors (in parentheses) are clustered at the municipality level m to account for the correlation in treatment within municipality and the estimations use the inverse weights of 2010. Includes *region*, *year*, *municipality* and *household* fixed effects as well as agricultural problems of the household (except drought or other weather-related problems). It also includes controls for the dependency ratio in 2010, the age of the household head, a dummy for whether or not the household head is a woman, a dummy for whether or not the household has a credit in that year, a dummy at the community level for whether or not the access to the cabecera (center of the town) is reachable and in good shape, and the minutes to reach the cabecera for the community. High-temp is a dummy equal to one for the municipalities with average temperature higher than 27 Celsius degrees, and zero for the other municipalities. Sample restricted to rural households producing uniquely crops (excludes livestock producers or those producing livestock and crops).

3.5 Adaptation responses to drought impacts

I start by using the three waves of the ELCA to verify if drought measures constructed with the SPI affect the agricultural outcomes. I argue that the SPI drought measures can be used as an alternative way to capture the impact of climate variability, in addition to the ones considered already in the literature (see section 3.4.2 and section 3.4.4). The SPI drought measures also allow me to determine better extreme climate events and facilitates the interpretation. First, I study the reaction by households after 2013 to droughts in 2010 (*short-term* and *medium-term adaptation*); and second, the reaction of the households after 2016 to *consecutive droughts* in 2010 and 2013. Categorizing the effects in this way should allow to disentangle how the households react in the short-term and in the medium-term to shocks happening in 2010. For the medium-term effects, as [Burke and Emerick \(2016\)](#)

mention, with panel data the effects found in the short-term could be underestimated, as rural households could adapt more to weather shocks (assuming that they can). To some extent, the results of *short-term* and *medium-term adaptation* and *adaptation following consecutive droughts* could also be compared to the results of weather and climate of section 3.4.2, respectively.

Figure 3.5 of Appendix 3C shows the timeline of the three waves of the ELCA 2010-2013-2016 and the ENSO events. As observed, there was neither El Niño nor La Niña around 2013, while El Niño phenomena occurred between July 2009 and March 2010 and before and during 2016 (November 2014 to May 2016). The event in 2010 could have affected the agricultural outcomes in 2010 (the baseline); in an attempt to rule out this possibility, I show later on that the socioeconomic control variables and the outcome variables were balanced with respect to the drought in 2010. In addition, Appendix 3A and 3B give evidence of a moderate drought in 2010 and an extreme drought in 2016, using the SPI measure. Also, Table 3.12 of Appendix 3A shows the different measures of SPI and drought by year of the survey.

3.5.1 Identification strategy using the DID approach and the SPI measure of droughts

The analysis aims to capture rural households' changes of their agricultural practice and in their labor market participation in the short-term and medium-term. I proceed by 1) using the three years 2010, 2013 and 2016, to see if the agricultural outcomes in 2013-2016 were affected by the drought in 2010 (*short-term* and *medium-term adaptation*); and 2) comparing years 2010, 2013 and 2016, to see if the agricultural outcomes in 2016 were affected differently for those who suffered a drought in 2010 and also in 2013 (*consecutive droughts adaptation*). Appendix 3E studies the effects of SPI on gross agricultural productivity (in value terms) and the additional outcomes of this section and it shows the impact from the SPI. As

the SPI takes continuous values, I prefer to construct dummy drought measures using the indices. This allows to determine better when a drought event happened and to capture easily the effects of droughts on agricultural outcomes in a DID framework. Using a dummy of drought also allows to make an easier interpretation of the impact of SPI. In this sense, the drought in 2010 determines the treatment and control groups (drought-no drought) in all the specifications. Table 3.12 in Appendix 3A describes the measures of droughts by year of the survey, which were constructed using the SPI. In the municipalities, for 18% of the households of the sample there was a drought in 2010 if I use the SPI at 12 months ("Drought 2010 (≤ -1) (SPI12)"). Using the SPI at 6 months indicates a drought for 62% of the households of the sample ("Drought 2010 (≤ -1) (SPI6)"). The equation to estimate for the rural household h is then:

$$\begin{aligned} \mathbf{Y}_{h,m,r,t} = & \nu + \beta_1 \times \mathbf{D}_{2010_m} \times A_t + \beta_2 \times D_{2010_m} + \beta_3 \times A_t \\ & + \theta \times X_{h,m,r,t} + \delta_r + \delta_t + \delta_m + \delta_h + \varepsilon_{h,m,r,t} \end{aligned} \quad (3.11)$$

$\mathbf{Y}_{h,m,r,t}$ agriculture outcomes for rural household h (or agricultural unit), of municipality m , region r in year t (2010-2013-2016);

$\mathbf{D}_{2010_{m,h}}$ treatment dummy for drought in municipality m of household h in year 2010; and 0 otherwise;

A_t is a 'post-treatment' dummy;

$X_{h,m,r,t}$ vector of controls for household characteristics;

$\delta_r, \delta_t, \delta_m, \delta_h$ dummies by *region*, by *year*, by *municipality* and by *household*.

Following Abadie et al. (2017), the standard errors are clustered at the municipality level m to account for the correlation in treatment within municipality. The estimations use the inverse weights of 2010, as recommended by the ELCA team. The coefficient of interest is β_1 . The identification assumption relies on the fact that rural households in treatment and control groups (affected versus not affected by drought in 2010) were similar

in 2010 - the parallel trends assumption. This assumption is tested in the results section. Additionally, I test if the control variables are balanced in the same treatment and control groups. Furthermore, I include *year* and *municipalities* fixed effects to avoid any omitted variable bias in this aspect.

3.5.2 Effects of drought on agricultural productivity in the short and medium term

The DID uses the drought shock in 2010 to determine treatment and control groups (drought in 2010 versus no drought in 2010). "Dummy 1 or more months drought (≤ -1) (SPI12)" is the variable used to define a moderate drought in 2010 in the main specification (see Table 3.2 for the drought definition). Table 3.16 in Appendix 3F shows the socioeconomic control variables in 2010 for the groups of drought 2010 versus no drought in 2010, aggregated by municipality. Although I find that all the control variables are balanced, they are included in the estimations. Interestingly, the variables of the percentage of crops facing problems in the parcels of the household are balanced as well as the household structure between the two groups of drought versus no drought in 2010. Also, there is no statistical difference in the variables of wealth and the access to the town center.

Table 3.17 in Appendix 3F presents the balance of the outcome variables for the group of drought versus no drought in 2010, aggregated by municipality. There is no statistical difference between the groups affected by drought in 2010 compared to the one without a drought. Also, it is compelling to find that the variables of land use, of gross agricultural productivity (in value terms) and investment, among others, are similar on average in both groups in 2010, giving evidence of the validity of the parallel trends assumption.

Table 3.8 summarizes the main findings of the DID, comparing outcomes in survey year

2010, 2013 and 2016.²² I find evidence that the drought measures can affect different channels. In panel a), having a drought in 2010²³ increases the land size planted in crops in 2013 and the area dedicated to livestock along the different measures, with the results statistically significant for crops but not for livestock. The variable captures only the effect on the total area used for livestock, but not gross agricultural productivity (in value terms) or total production of livestock. For crops, the drought in 2010 increases the land used by 11.5%-14.4% and the coefficients are statistically significant at the 5% and 1% level. This goes in line with the findings of [Aragón et al. \(2021\)](#) and [Costinot et al. \(2016\)](#) where rural households respond to droughts by planting and using more land, as well as the findings of the theoretical framework proposed in section 3.3. The more planted area is also in line with a higher productivity in terms of the value of the total production, particularly for drought measures using the SPI at 6 months. This result is similar to the findings in section 3.4.2, also by crop type. The gross agricultural productivity (in value terms) of vegetables and fruits are not affected by droughts. Cereals, and to a lesser extent, coffee seem to benefit from droughts (except for coffee production during extreme droughts). However, the results depend on the high versus low-temperature municipalities as discussed in section 3.4.2. Also here, the interpretation of these results should be taken with caution as the variables of gross agricultural productivity use a restricted sample of only crop producers.

²²Although the measures of adaptation could have started before 2010, the ELCA panel survey starts in 2010 and it makes it difficult to check exactly if rural households implemented similar strategies before. However, Figure 3.3 shows no evidence of extreme droughts during the three or four years previous to 2010, with very few municipalities affected by moderate droughts in the three years before 2010. As such, the drought measure is constructed based on the SPI index, which is the anomaly with respect to this long-term mean.

²³The dummies for drought in 2010 during the agricultural growing season perform worse than those defined in the calendar year and were all the time zero. Also, the extreme drought in 2010 using the SPI 12 months give only zeros, which makes it impossible to use them in the estimations of the tables.

Table 3.8: DID for 2010-2013-2016 using Drought in 2010 on main agricultural outcomes
a) Production

	Land (Ha) ln		Value prod. total (Millions Col. Pesos/Ha) ln				
	Crops (1)	Livestock (2)	Total (3)	Vegetables (4)	Fruits (5)	Cereals (6)	Coffee (7)
Drought 2010 (≤ -1) (SPI12)xPost(2013)	0.057 (0.037)	0.056 (0.095)	-0.190 (0.192)	-0.263 (0.209)	0.215 (0.168)	-0.385 (0.359)	0.138 (0.127)
Drought 2010 (≤ -1) (SPI6)xPost(2013)	0.109** (0.044)	0.068 (0.083)	0.386*** (0.112)	-0.182 (0.315)	0.248 (0.198)	0.181 (0.305)	0.166** (0.072)
Drought 2010 (≤ -2) (SPI6)xPost(2013)	0.135*** (0.031)	-0.064 (0.044)	-0.122 (0.108)	0.164 (0.103)	-0.084 (0.083)	0.340*** (0.105)	-0.341*** (0.093)
N	7902	7902	1770	1770	1770	1770	1770
r2	0.73	0.78	0.62	0.62	0.48	0.70	0.72
r2_a	0.56	0.64	0.29	0.29	0.04	0.45	0.48

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

b) Employment and Investment

	HH Head or partner work			Land Water (ha) ln			Land INV(ha) ln
	Employed (1)	Agro. (2)	Off-farm (3)	Any (4)	Own (5)	Ext. (6)	ANY no house (7)
Drought 2010 (≤ -1) (SPI12)xPost(2013)	0.110 (0.066)	0.132 (0.089)	-0.011 (0.077)	0.182* (0.102)	-0.031 (0.046)	0.223** (0.091)	-0.124 (0.134)
Drought 2010 (≤ -1) (SPI6)xPost(2013)	0.171*** (0.049)	0.301*** (0.032)	0.031 (0.073)	0.068 (0.089)	-0.046 (0.054)	0.125 (0.080)	-0.186** (0.075)
Drought 2010 (≤ -2) (SPI6)xPost(2013)	0.288*** (0.036)	0.333*** (0.059)	-0.033 (0.031)	0.106* (0.057)	-0.227*** (0.038)	0.362*** (0.055)	-0.142** (0.056)
N	7902	7902	7902	7902	7902	7902	7902
r2	0.66	0.63	0.49	0.76	0.71	0.61	0.55
r2_a	0.44	0.40	0.17	0.61	0.52	0.36	0.26

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: based on the ELCA panel survey 2010-2013-2016. The standard errors (in parentheses) are clustered at the municipality level m to account for the correlation in treatment within municipality and the estimations use the inverse weights of 2010. Includes *region*, *year*, *municipality* and *household* fixed effects as well as agricultural problems of the household (except drought or other weather-related problems). It also includes controls for the dependency ratio in 2010, the age of the household head, a dummy for whether or not the household head is a woman, a dummy for whether or not the household has a credit in that year, a dummy at the community level for whether or not the access to the cabecera (center of the town) is reachable and in good shape, and the minutes to reach the cabecera for the community. Each row corresponds to the interaction of post-treatment and Drought in 2010 using the SPI on the outcome variable in the column. Each row is part of different estimations. The N, r2 and r2_a in each column corresponds to the estimations of the drought measure in the first row but the explanatory power of the other estimations was almost the same when using the other ones. In columns 3)-7) of Panel a), sample restricted to rural households producing uniquely crops (excludes livestock producers or those producing livestock and crops).

3.5.3 Effects of consecutive droughts on agricultural productivity

In Table 3.9, I compare the waves 2010, 2013 and 2016 to see how rural households react to a drought in 2010 and again in 2013. I change the variable of drought used in the previous DID to represent cumulative droughts, a dummy for having experienced a drought in 2010 and also in 2013.²⁴ In panel a) of Table 3.9, consecutive drought in 2010 and 2013 increases the total land dedicated to livestock by 15.3%, and to crops by 10.1%. There is an increase in the total gross agricultural productivity by 67.7%, statistically significant at the 1% level. The value of the total production of fruits and cereals increases by 104.2% and 29.7% respectively (columns 5 and 6 in panel a), while gross agricultural productivity of vegetables decrease by 46.7%. With respect to perennial crops, there is an increase in the productivity of fruits after droughts, but no change in the value of coffee production. For annual crops, vegetables show a reduction in the gross agricultural productivity after droughts, while cereals tend to benefit more. Similar to section 3.4.2 for annual crops, vegetables are more affected while cereals tend to benefit from droughts and weather variation.

²⁴The drought variables using the agricultural growing season and the extreme droughts were always zero, which prevents me from using them in the estimations.

Table 3.9: DID for 2010-2013-2016 using the consecutive Drought 2010 and 2013 on main agricultural outcomes

a) Production

	Land (Ha) ln		Value prod. total (Millions Col. Pesos/Ha) ln				
	Crops (1)	Livestock (2)	Total (3)	Vegetables (4)	Fruits (5)	Cereals (6)	Coffee (7)
Drought 2010&2013 (≤ -1) (SPI12)xPost(2016)	0.015 (0.042)	0.142*** (0.037)	0.517*** (0.100)	-0.383*** (0.086)	0.714*** (0.058)	0.260* (0.130)	-0.076 (0.065)
Drought 2010&2013 (≤ -1) (SPI6)xPost(2016)	0.096* (0.052)	0.027 (0.059)	0.120 (0.141)	0.157 (0.170)	-0.035 (0.112)	0.124 (0.140)	-0.146 (0.155)
N	7902	7902	1770	1770	1770	1770	1770
r2	0.73	0.78	0.62	0.62	0.49	0.70	0.72
r2_a	0.56	0.64	0.29	0.29	0.05	0.44	0.48

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

b) Employment and Investment

	HH Head or partner work			Land Water (ha) ln			Land INV(ha) ln
	Employed (1)	Agro. (2)	Off-farm (3)	Any (4)	Own (5)	Ext. (6)	ANY no house (7)
Drought 2010&2013 (≤ -1) (SPI12)xPost(2016)	0.070 (0.041)	0.018 (0.067)	0.007 (0.033)	-0.075* (0.039)	-0.204*** (0.046)	-0.018 (0.029)	0.171*** (0.046)
Drought 2010&2013 (≤ -1) (SPI6)xPost(2016)	0.099** (0.035)	0.174** (0.065)	0.020 (0.041)	0.027 (0.065)	0.034 (0.080)	0.022 (0.030)	0.006 (0.102)
N	7902	7902	7902	7902	7902	7902	7902
r2	0.66	0.63	0.49	0.76	0.71	0.61	0.55
r2_a	0.44	0.40	0.17	0.61	0.52	0.36	0.26

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: based on the ELCA panel survey 2010-2013-2016. The standard errors (in parentheses) are clustered at the municipality level m to account for the correlation in treatment within municipality and the estimations use the inverse weights of 2010. Includes *region*, *year*, *municipality* and *household* fixed effects as well as agricultural problems of the household (except drought or other weather-related problems). It also includes controls for the dependency ratio in 2010, the age of the household head, a dummy for whether or not the household head is a woman, a dummy for whether or not the household has a credit in that year, a dummy at the community level for whether or not the access to the cabecera (center of the town) is reachable and in good shape, and the minutes to reach the cabecera for the community. Each row corresponds to the interaction of post-treatment and Drought in 2010-13 using the SPI on the outcome variable in the column. Each row is part of different estimations. The N, r2 and r2_a in each column corresponds to the estimations of the drought measure in the first row but the explanatory power of the other estimations were almost the same when using the other ones. In columns 3)-7) of Panel a), sample restricted to rural households producing uniquely crops (excludes livestock producers or those producing livestock and crops).

3.5.4 How do households adapt?

Labor Market

In terms of the labor market, panel b) of Table 3.8 shows that the drought in 2010 increases the probability of the household head or partner being employed by 17.1 to 28.8 percentage points and the probability of doing agricultural work by 30.1 to 33.3 percentage points, but there is no effect in terms of off-farm work. This could be associated to the use of more hectares of land used by the rural household. As the rural household expands the use of land for crops and for livestock, household members might need to work more on the farm.

Panel b) of Table 3.9 also shows some results indicating adaptation in the labor market after consecutive droughts. There is an increase in the probability of being employed and doing agricultural work by 9.9 and by 14.4 percentage points respectively (column 1 and column 2), statistically significant at the 5% level. This could indicate that after two consecutive droughts, rural households might still follow similar strategies to the ones found previously in the short-term and medium-term in section 3.5.2. They increase the total land area dedicated to crops, keep working more on the farm, and using some perennial crops such as fruits. The increase in land devoted to fruits could be explained by the fact that perennial crops tend to be a more stable source of income rather than annual crops. Following consecutive droughts, there are smaller increases in land and on-farm labor, which indicates a different response, and may also be linked to the fact that the drought in 2013 was not as severe as the one in 2010 (see Figure 3.3).

Water Use

In terms of access to water (short-term and medium-term), the drought in 2010 increased the land with any water in 2013 by 10.6%-20%, particularly from external sources by 25%-43.6% (panel b) of Table 3.8). In the short-term and medium-term, rural households might

try to look for additional sources of water when facing a drought shock and they depend less on own sources, which is consistent with the decrease observed in own water in column 5).

In addition, Panel b) of Table 3.9 shows the results for consecutive droughts on water access. There is a reduction of land with access to any water (own and external sources), statistically significant at the 10% level (columns 4 of panel b)), driven by the reduction in the land with access to own sources. This is similar to the result found in the *short-term* and *medium-term adaptation*, so droughts affect the chances to use own sources of water by the rural household. External water sources may also be more difficult to come by following consecutive droughts, as the sources of water of the rural households start to be depleted. This is an interesting difference between adaptation after one drought only and consecutive droughts.

Investment

Regarding investment (short-term and medium-term) in panel b) of Table 3.8, the drought in 2010 reduces the investments in the land plot (excluding housing), which could be seen as a way to smooth consumption for the household and postpone investments in the short-term and medium-term. For the consecutive droughts, I observe an increase in the hectares of land with investments (excluding housing) for rural households facing a consecutive drought (in 2010 and 2013). This corresponds to an increase by 18.6% of investments of this type in 2016. A potential explanation could be that as droughts become more frequent, rural households might want to make investments in the farm to reduce the impact of potential future droughts. I explore on which particular items the rural households focus more by analysing the type of investments (irrigation, structures, etc.).

While there was a reduction of investment excluding housing as *short-term* and *medium-term adaptation* in Table 3.8, an increase was observed following *consecutive droughts* in

Table 3.9. Panel a) of Table 3.10 shows that the reduction of 20.4% (second line of column 2) of any investment excluding housing is explained by the reductions in investments, mainly on structures and then on other items. The reduction by 15.3% in the third line of column 2) is explained by the investments in fruits, wood and commercial items, but the item of structures shows an increase. After consecutive droughts, there is an increase in the investment excluding housing, that appears to be driven mainly by the investments in fruits, a perennial crop. Interestingly, there is a reduction on the investments in housing after consecutive droughts. There is also a slight increase in the investment in irrigation of around 3% following droughts in the short-term and medium-term but a reduction of the same size after consecutive droughts. To summarize, the initial reduction in investment is observed mainly through reductions in investment in fruit crops, commercial items and structures for the farm; this could be interpreted as a consumption smoothing strategy for the household to reduce the impact of droughts. The increase after consecutive droughts is explained by more investments to produce fruits and a reduction on housing investments. Relating to the theoretical model of section 3.3, as temperature increases (higher values of climate) or for consecutive droughts, the production of some crops can benefit by increasing the use of adaptive input, up to an optimal point after which it is no longer profitable.

Table 3.10: DID for Drought in 2010 by type on investment

a) Investment 1

	Land INV(ha) ln				
	Any (1)	Any no house (2)	Irrigation (3)	Structures (4)	Conservation (5)
Drought 2010 (≤ -1) (SPI12)xPost(2013)	-0.120 (0.145)	-0.124 (0.134)	-0.008 (0.014)	-0.082 (0.053)	0.042 (0.030)
Drought 2010 (≤ -1) (SPI6)xPost(2013)	-0.187** (0.076)	-0.186** (0.075)	0.028** (0.013)	-0.152*** (0.043)	-0.023 (0.031)
Drought 2010 (≤ -2) (SPI6)xPost(2013)	-0.128** (0.057)	-0.142** (0.056)	0.021* (0.012)	0.126** (0.046)	0.038** (0.018)
Drought 2010&2013 (≤ -1) (SPI12)xPost(2016)	0.053 (0.061)	0.171*** (0.046)	-0.030** (0.011)	-0.004 (0.036)	0.021 (0.018)
Drought 2010&2013 (≤ -1) (SPI6)xPost(2016)	-0.061 (0.121)	0.006 (0.102)	0.001 (0.021)	-0.020 (0.082)	0.009 -0.025
N	7902	7902	7902	7902	7902
r2	0.58	0.55	0.42	0.49	0.47
r2_a	0.31	0.26	0.04	0.17	0.13

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

b) Investment 2

	Land INV(ha) ln				
	Fruits (1)	Wood (2)	Commer. (3)	House (4)	Others (5)
Drought 2010 (≤ -1) (SPI12)xPost(2013)	0.065 (0.108)	-0.026 (0.016)	-0.115 (0.071)	-0.007 (0.045)	-0.023 (0.028)
Drought 2010 (≤ -1) (SPI6)xPost(2013)	0.005 (0.033)	-0.006 (0.009)	0.005 (0.031)	-0.029 (0.044)	-0.052** (0.021)
Drought 2010 (≤ -2) (SPI6)xPost(2013)	-0.134*** (0.019)	-0.034*** (0.006)	-0.286*** (0.025)	0.033 (0.028)	-0.011 (0.016)
Drought 2010&2013 (≤ -1) (SPI12)xPost(2016)	0.158*** (0.012)	0.019*** (0.006)	0.010 (0.009)	-0.141*** (0.034)	0.022** (0.008)
Drought 2010&2013 (≤ -1) (SPI6)xPost(2016)	0.035** (0.015)	0.008 (0.011)	0.005 (0.019)	-0.097** (0.045)	-0.013 (0.019)
N	7902	7902	7902	7902	7902
r2	0.51	0.42	0.42	0.50	0.45
r2_a	0.20	0.05	0.05	0.17	0.10

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: ELCA panel survey 2010-2013-2016. Standard errors (parentheses) clustered at the municipality level m . Using inverse weights of 2010. Includes *region*, *year*, *municipality* and *household* fixed effects as well as the same controls used in previous estimations (see section 3.4.2). Each row corresponds to the interaction of post-treatment and Drought in 2010 using the SPI on the outcome variable in the column. Each row is part of different estimations. N, r2 and r2_a in each column from the first drought measure estimations.

Land use

In order to explore deeper the type of land uses of the rural households, Table 3.11 shows the estimations by the different categories of land use following droughts and consecutive droughts used in the DID. Panel a) shows that the increase observed previously of the land used for crops is mainly driven by permanent and to a lesser extent by transitional crops, while mixed crops present a reduction in the land used following droughts. Interestingly, the total area of land increases after a consecutive drought in column 1) of panel a). A potential explanation of this is that rural households might have bought additional plots of land or perhaps by deforesting nearby areas. However, this is an aspect that should be explored in more detail in future research.

Panel b) of Table 3.11 shows that after droughts, there is a reduction of land used for pastures and land left fallow. When facing droughts, the rural households re-allocate land by reducing the type of land less exploited and assigning it to production of crops and for livestock. This trade-off appears in the theoretical framework of section 3.3, where the rural household can transfer the land left fallow into the production of the different crops that she produces. In fact, the results of the theoretical framework show that the rural households adapt to extreme climate by incorporating the land left fallow and allocating it to the production of the different crops that they produce.

Table 3.11: DID for Drought in 2010 by type of land use

a) Land use 1					
	Land (Ha) ln				
	Total (1)	Crops (2)	Perm. (3)	Transit. (4)	Mixed (5)
Drought 2010 (≤ -1) (SPI12)xPost(2013)	0.065 (0.051)	0.057 (0.037)	0.119 (0.092)	0.047 (0.033)	-0.111 (0.096)
Drought 2010 (≤ -1) (SPI6)xPost(2013)	0.053 (0.049)	0.109** (0.044)	0.041 (0.062)	0.072** (0.029)	0.008 (0.045)
Drought 2010 (≤ -2) (SPI6)xPost(2013)	-0.038 (0.027)	0.135*** (0.031)	0.252*** (0.038)	-0.053** (0.023)	-0.078** (0.028)
Drought 2010&2013 (≤ -1) (SPI12)xPost(2016)	0.162*** (0.034)	0.015 (0.042)	0.097 (0.058)	-0.035 (0.029)	-0.042 (0.032)
Drought 2010&2013 (≤ -1) (SPI6)xPost(2016)	0.093* (0.049)	0.096* (0.052)	0.131 (0.078)	0.007 (0.037)	-0.030 (0.061)
N	7902	7902	7902	7902	7902
r2	0.86	0.73	0.63	0.61	0.50
r2_a	0.77	0.56	0.39	0.35	0.17
Standard errors in parentheses					
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$					
b) Land use 2					
	Land (Ha) ln				
	Livestock (1)	Pasture (2)	Forest (3)	Others (4)	No used (5)
Drought 2010 (≤ -1) (SPI12)xPost(2013)	0.056 (0.095)	-0.073 (0.068)	0.032 (0.035)	0.017 (0.016)	0.005 (0.071)
Drought 2010 (≤ -1) (SPI6)xPost(2013)	0.068 (0.083)	-0.013 (0.086)	0.026 (0.023)	-0.029* (0.016)	-0.077 (0.055)
Drought 2010 (≤ -2) (SPI6)xPost(2013)	-0.064 (0.044)	0.080 (0.046)	0.016 (0.017)	0.026* (0.012)	-0.342*** (0.038)
Drought 2010&2013 (≤ -1) (SPI12)xPost(2016)	0.142*** (0.037)	-0.078*** (0.023)	0.053*** (0.016)	0.095*** (0.008)	0.028 (0.024)
Drought 2010&2013 (≤ -1) (SPI6)xPost(2016)	0.027 (0.059)	0.013 (0.044)	0.033 (0.020)	-0.011 (0.014)	-0.009 (0.034)
N	7902	7902	7902	7902	7902
r2	0.78	0.44	0.60	0.47	0.59
r2_a	0.64	0.08	0.35	0.12	0.32
Standard errors in parentheses					
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$					

Note: ELCA panel survey 2010-2013-2016. Standard errors (parentheses) clustered at the municipality level m . Using inverse weights of 2010. Includes *region*, *year*, *municipality* and *household* fixed effects as well as the same controls used in previous estimations (see section 3.4.2). Each row corresponds to the interaction of post-treatment and Drought in 2010 using the SPI on the outcome variable in the column. Each row is part of different estimations. N, r2 and r2_a in each column from the first drought measure estimations.

3.5.5 Heterogeneous effects and alternative strategies

Access to credit

Table 3.19 in Appendix 3G shows the results interacting the drought measures with the control variable of having access to credit (the *short-term* and *medium-term adaptation*). It shows that moderate droughts increase the land dedicated to crops and livestock for those with credit compared to those without it (the interaction term of columns 1 and 2 of panel a), but extreme droughts change the behaviour of the households. For those without credit, an extreme drought increases the hectares of land dedicated to crops but reduces the ones for livestock. Those with credit reduce the land dedicated to crops after an extreme drought, compared to households without a credit (the interaction term), but the net effect is an increase in the land used for crops and a reduction of livestock (the sum of both coefficients). Regarding the value of total production, there are statistically significant reductions for rural households with credit. The net effect after an extreme drought for households with a credit is a decrease in the value of the total production. Also note that the households without credit access benefit from droughts and present increases in the gross agricultural productivity of cereals and coffee, which could be due to higher production in yields or higher prices or both. Unfortunately, the way that the variables of value of the crops production are constructed does not allow me to disentangle exactly which effect dominates. On the contrary, for those with credit access, the extreme droughts only increased the productivity of vegetables while reducing the productivity of coffee compared to those without credit.

For the work variables, in both cases (having access to credit or not) I observe an increase the probability of the household head or partner to be employed and work in the farm, but once the droughts are extreme, those with access to credit tend to reduce labor participation in the farm or in off-farm activities with respect to household without credit access. Being exposed to an extreme drought in 2010 increases the probability of doing on-farm work for

those without credit access by 42.3 percentage points. However, extreme drought reduces the probability of doing on-farm work by 10.4 percentage points for household with credit access with respect to those without it. Both effects are statistically significant at the 1% and 5% level. The net effect is still positive for those with credit. Being exposed to an extreme drought increases the probability of doing on-farm work by 31.9 percentage points for the households with credit access. In conclusion, there seems to be some significant differences depending on having access or not to credit, when households face extreme events. Having access to credit might reduce the labor participation, as these households have extra sources of income. It could be that those without credit might not have another choice than working, in their own farms or doing off-farm work.

Another important mechanism relates to the variables for land with access to water. On the one hand, being exposed to a drought and having a credit increases the hectares of land with access to water (any and external) compared to those without credit. In line with the findings for moderate droughts, the interaction in the table shows an increase in the hectares of land for those with credit, compared to those without it. The net effect of droughts for households with credit, column 4 and 6 of panel b) in Table 3.19, is positive across the water sources (any or external). In this sense, having a credit could facilitate households getting access to external sources of water for their farms. On the other hand, only for extreme droughts there is a reduction in the access to water for households without a credit. When households face extreme droughts, there is a reduction by 26% to 36% in the hectares of land with access to water across all the sources (any, own or external) for those without credit. In terms of investment excluding housing, column 7 of panel b), droughts have a negative impact for those without credit. It is only for extreme droughts, having a credit increases the chances of making investments (excluding housing) for the households exposed to droughts, compared to those without credit. However, the net impact for those with credit is close to zero (the sum of the two coefficients of column 7 of panel b) in Table 3.19).

Migration as an alternative adaptation strategy

Finally, I explore an alternative way in which rural households can cope with droughts. As has been shown in the literature, household can respond to droughts on the extensive margin through migration (see Cattaneo et al. (2019)). I develop this by constructing two different measures: 1) "Migrating outside Municipality" for cases when the household moved outside of the initial municipality of 2010. This accounts for 5% of the ELCA sample in Appendix 3C and those were excluded in all the previous estimations as explained in the data section 3.2. And 2) "Moving inside Municipality" for cases when the coping mechanism is through movements to another community inside the same initial municipality. Compared to migration outside of the municipality, these movements are not very far and should be considered as re-location inside the same initial municipality. As they respond to different channels, they are analyzed separated by using the variable "Migrating outside Municipality" and are mutually excluded. These movements inside the municipality account for 11.2% of the ELCA sample of table 3.13 in Appendix 3C and were included in all the previous estimations.

Table 3.18 in Appendix 3G shows the results of the estimations for type of movement. The first three rows correspond to the estimations for the drought measures used in the *short-term* and *medium-term adaptation* analysis, while the next two rows are for the drought measures used in the *consecutive droughts adaptation* analysis. Column 1) of the table indicates a reduction in the probability of migrating following the drought in 2010 between 3 to 11 percentage points, statistically significant at the 10% and 1% level, respectively. However, it is important to take into account that the estimations of column 1) do not control for household fixed effects as they remove all the variability explained by the variability in the explanatory variables (estimated coefficients close to zero in all the variables). Results of column 1) are thus not totally comparable with the estimations in previous sections. However, column 2) of Table 3.18 uses the same controls and fixed effects and are comparable to the results of previous sections. It shows an increase in the probability of moving (re-

locating) to another community inside the same municipality by 4.9 to 11.9 percentage points following a moderate drought in 2010 (not for extreme drought), which is statistically significant at conventional levels. The behaviour is similar for the measures of moderate droughts in the analysis of *short-term* and *medium-term adaptation* and in the *consecutive droughts adaptation* analysis. Interestingly, extreme drought in 2010 reduced the probability of re-locating to another community inside the municipality. Moderate droughts might encourage households to move to other areas inside the same municipality, perhaps looking for job options or other alternative ways to get income. But for extreme droughts, it is harder to find sources of income in the municipality. This corresponds to the findings of the theoretical framework according to which the rural households reduce the production of the crops and the use of adaptive input as climate becomes too extreme. However, it is difficult to assert that the rural households in the sample try to migrate to other municipalities as the estimations of column 1) are not totally comparable.

3.6 Conclusion

This article analyses the effects of weather factors on gross agricultural productivity for a country in development - Colombia - as the literature on agricultural adaptation has mainly focused on developed countries such as U.S. (see [Deschênes and Greenstone \(2007\)](#) and [Schlenker and Roberts \(2009\)](#)). It combines the Colombian panel survey (ELCA) conducted in 2010, 2013 and 2016 with weather information from satellite images to explore how rural households adapt. I start by using similar identification frameworks as the ones already explored in the literature, in order to have a comparison. Then, I propose an alternative way to measure the impact of climate variability on agriculture, by constructing a measure of droughts using the SPI in a DID framework. I divide the DID analysis in two, to see how rural households adapt in the short-term and the medium-term and then, with respect to consecutive droughts.

I separate the effect of climate from yearly weather deviations for the Colombian rural households following Kelly et al. (2005), Deschênes and Kolstad (2011), and in particular Bento et al. (2020). First, the long-run mean of rainfall has a positive effect on gross agricultural productivity and while temperature deviations affect positively gross agricultural productivity, there is no effect from the long-run temperature mean. Also, the analysis of the marginal effects show that the higher temperature affects positively gross agricultural productivity in low-temperature municipalities but negatively in high-temperature municipalities. This shows the potential winners and losers from climate trends and weather shocks and gives evidence of adaptation of rural households in high-temperature municipalities. With respect to the marginal effects by crop, cereals and coffee benefit from higher temperature (long-run and deviations), mostly in low-temperature municipalities; vegetables are mostly affected negatively by the increasing temperature trends; fruits are negatively affected by temperature shocks and climate in low-temperature municipalities while positively affected in high-temperature municipalities; gross productivity of coffee seems to be only affected negatively by temperature shocks in high-temperature municipalities.

Regarding nonlinear effects, the GDD-HDD measures explain less the gross agricultural productivity of the rural households studied, compared to the climate and weather deviation variables. With respect to the marginal effects by crop, the HDD affect negatively the gross agricultural productivity of cereals and coffee in high-temperature municipalities, while they benefit low-temperature municipalities. On the contrary, HDD increase the gross agricultural productivity of vegetables and fruits of high-temperature municipalities while decreasing it in low-temperature municipalities.

Finally, section 3.5 studies if those affected by previous droughts adapt better to those not affected by those shocks. When facing droughts, as *short-term* and *medium-term adaptation*,

the rural households re-allocate land by reducing the type of land left fallow and assigning it to production of crops and for livestock. This captures some of the trade-offs proposed in the theoretical model of section 3.3. It also goes in line with higher values of the total production, particularly for cereals and coffee. Following consecutive droughts (in 2010 and in 2013), rural households also expand the land area dedicated to crops and livestock, with an increase in the gross agricultural productivity of cereals and fruits, but with a negative effect on vegetables. Using more land available in the farm leads the household head or partner to do more agricultural work in the short-term and medium-term. This is similar to Aragón et al. (2021) who find that high-temperatures reduced gross agricultural productivity, and that rural households attenuate the effect on output by increasing the planted area and by mixing crops.

Across the analysis, the marginal effect on perennial crops (fruits and coffee) depends on the high versus low-temperature municipalities and the type of crop. In low-temperature municipalities, higher temperature benefits coffee while it reduces the gross agricultural productivity of fruits; in high-temperature municipalities the marginal effect is the opposite and benefits fruits while reducing the gross agricultural productivity of coffee. With respect to the marginal effects on annual crops, the gross agricultural productivity of vegetables is more affected while cereals tend to benefit from droughts and weather variation. The analysis of the climate-yearly weather deviations and the DID gave qualitatively similar findings, with a positive effect from temperature weather shocks and droughts (short-term and medium-term) on gross agricultural productivity.

In the short-term and medium-term after consecutive droughts, rural households smooth consumption by reducing non-housing investments and increasing the hectares of land with investments (excluding housing). As droughts become more frequent, rural households make investments in the farm to reduce the impact of potential future droughts. Rural households

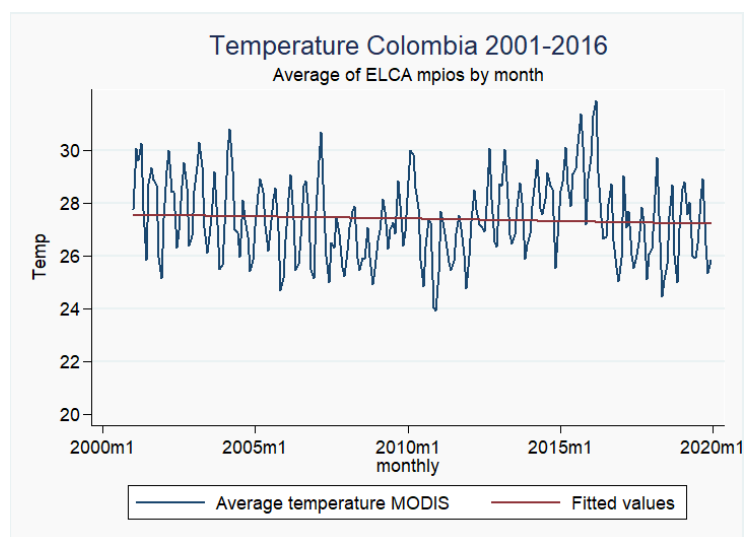
increase the land used, work more in the plots and postpone investments, but as droughts start to be more frequent, it might become more difficult to implement those measures of adaptation. Additionally, the drought in 2010 decreased the access to own water while external access to water increased in the short-term and medium-term. However, when facing consecutive droughts, it becomes more difficult to get access to any water (own and external sources).

A future research agenda is to check the behavior of the farmers who experienced consecutive droughts also in later waves of the ELCA panel (when they become available). Also, explore more in detail the effect of prices and effect of volume in the gross agricultural productivity analyzed here, as well as analyse the potential channels playing a role in the low-temperature and high-temperature municipalities. This deserves attention and it has not been extensively explored in low and middle income countries. Finally, additional heterogeneous effects can be further explored with respect to commercial versus subsistence rural households. Regarding policy implications of the responses to climate change in agriculture, policymakers should take into account the winners and losers of extreme heat, and probably focus more on high-temperature areas. Adaptation to extreme temperature in areas with already high temperature can be more difficult as rural households have fewer possibilities to implement measures against droughts. While low-temperature areas could benefit from higher temperature, some crops might also benefit and have the possibility of being produced there. Food security policies might have to consider these aspects too, switching the production from some crops to others, moving the production of specific crops from some areas to others, or providing safety nets in the most affected high-temperature areas.

Appendices

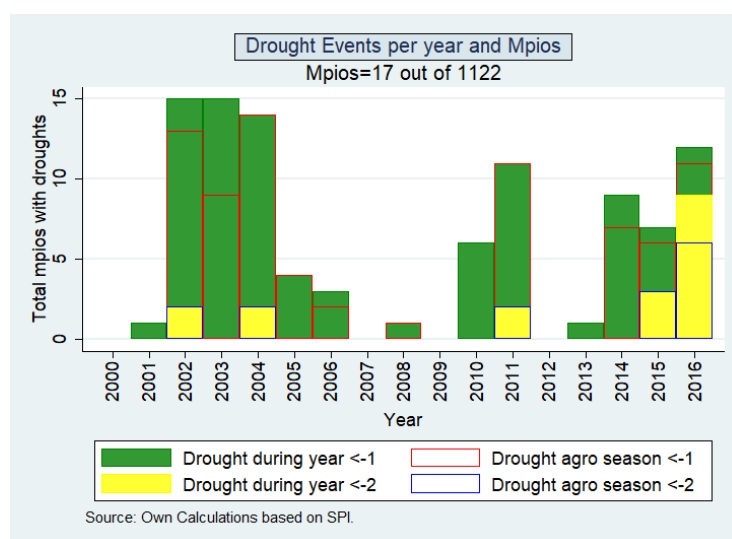
3A Average temperature - municipalities of the sample

Figure 3.2: Average Temperature Colombia - ELCA municipalities



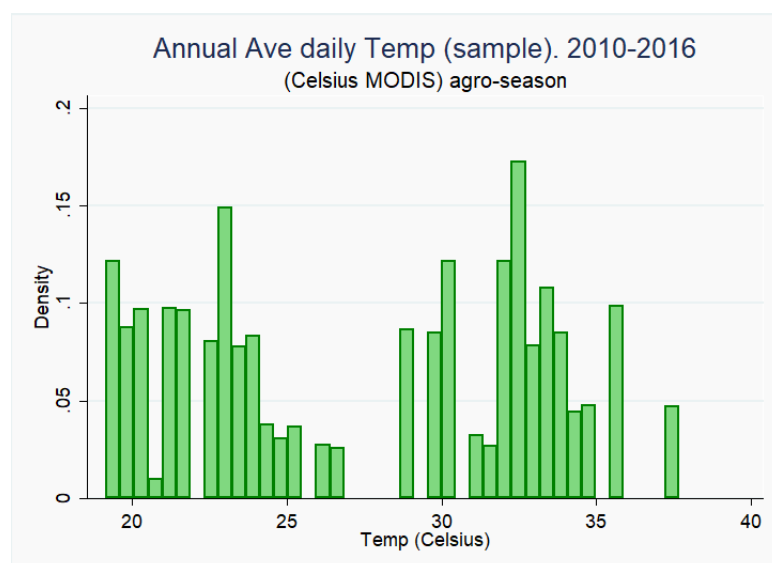
Note: Based on MODIS Land Surface Temperature data imputed to each municipality. Satellite images MOD11A1.006 at 1km of resolution.

Figure 3.3: Droughts by year in the ELCA municipalities



Note: SPI refers to the number of standard deviations by which observed anomaly deviates from long-term mean. The figure uses a monthly scale of 12 and corresponds to the number of months over which water deficits accumulate.

Figure 3.4: Histogram Temperature Agricultural Growing Season



Note: Based on MODIS Land Surface Temperature data imputed to each municipality. Satellite images MOD11A1.006 at 1km of resolution.

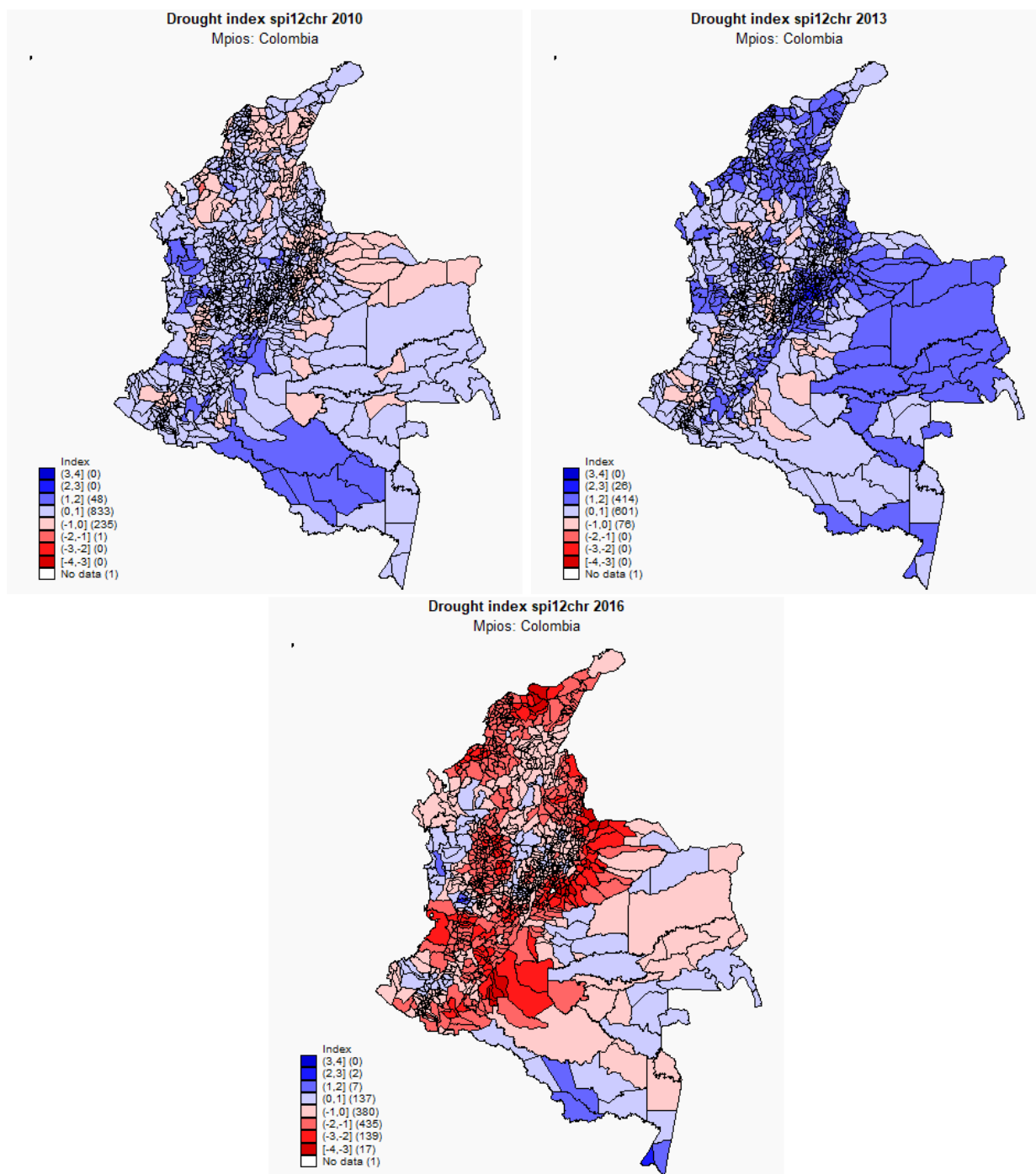
Table 3.12: SPI variables by ELCA year - Average

	2010 mean	2013 mean	2016 mean
SPI Index 12 months CHIRPS(opposite)	-0.77	-1.03	1.31
SPI Index 12 months CHIRPS(opposite)-agro season	-1.23	-1.37	1.14
Dummy 1 or more months drought (≤ -1)(SPI12)	0.18	0.03	0.79
Dummy 1 or more months drought (≤ -1)(SPI12)-agro season	0.00	0.00	0.75
Dummy 1 or more months drought (≤ -2)(SPI12)	0.00	0.00	0.68
Dummy 1 or more months drought (≤ -2)(SPI12)-agro season	0.00	0.00	0.47
Drought 2010 (≤ -1) (SPI12)	0.18	0.18	0.17
Drought 2010-13 (≤ -1) (SPI12)	0.03	0.03	0.03
SPI Index 6 months CHIRPS(opposite)	-0.17	-0.44	1.21
SPI Index 6 months CHIRPS(opposite)-agro season	-0.78	-1.04	1.12
Dummy 1 or more months drought (≤ -1)(SPI6)	0.62	0.56	0.79
Dummy 1 or more months drought (≤ -1)(SPI6)-agro season	0.00	0.05	0.77
Dummy 1 or more months drought (≤ -2)(SPI6)	0.01	0.07	0.73
Dummy 1 or more months drought (≤ -2)(SPI6)-agro season	0.00	0.00	0.60
Drought 2010 (≤ -1) (SPI6)	0.62	0.61	0.60
Drought 2010-13 (≤ -1) (SPI6)	0.55	0.56	0.54

Note: based on the ELCA panel survey 2010-2013-2016 and weather data, using expansion factors for 2010 as recommended by the ELCA team.

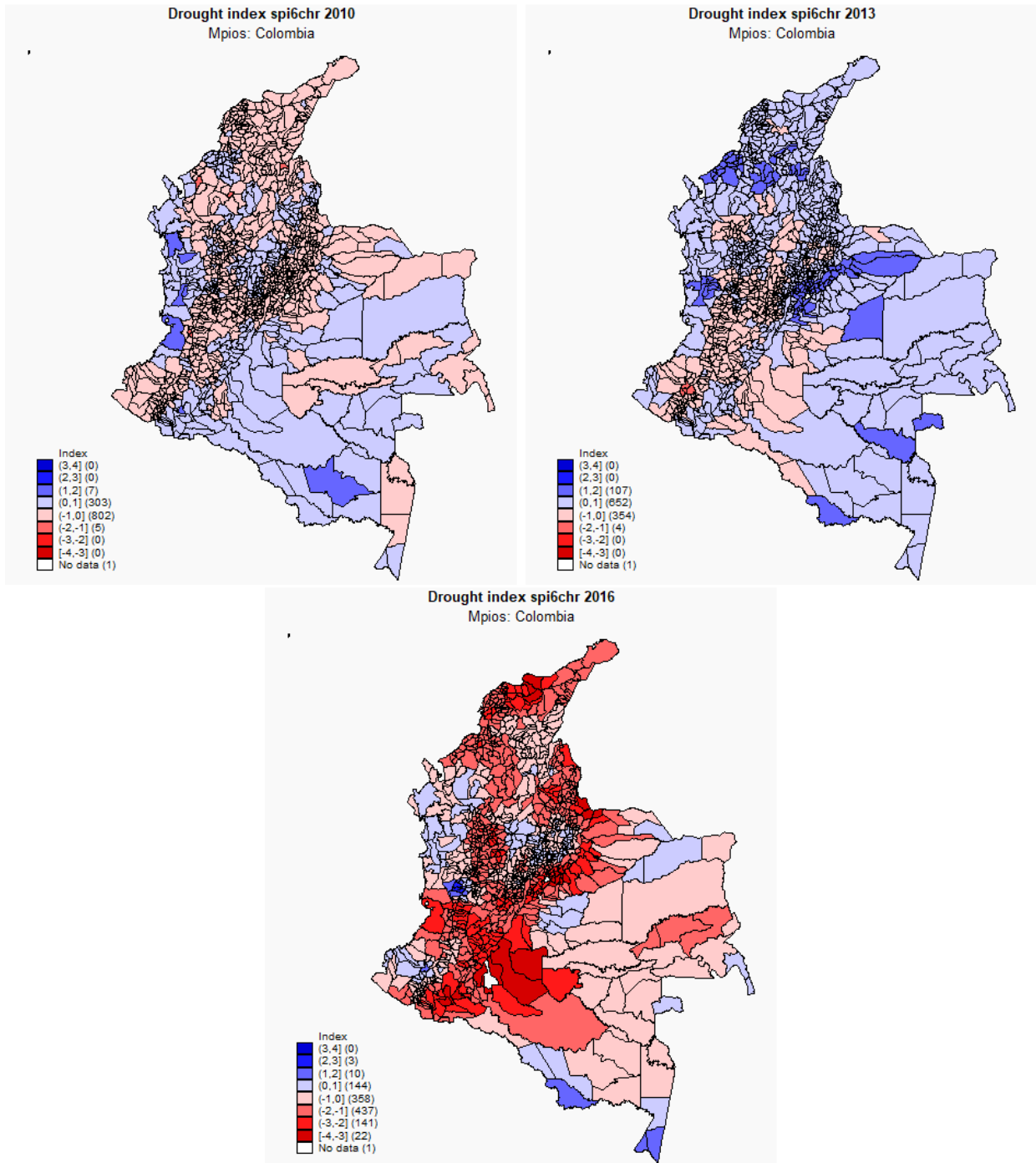
3B Maps of SPI drought measures by municipality

Scale=12 months



Note: SPI number of standard deviations by which observed anomaly deviates from long-term mean. Monthly scale (12-6) as number of months over which water deficits accumulate. Calculated based on CHIRPS data.

Scale=6 months



Note: SPI number of standard deviations by which observed anomaly deviates from long-term mean. Monthly scale (12-6) as number of months over which water deficits accumulate. Calculated based on CHIRPS data.

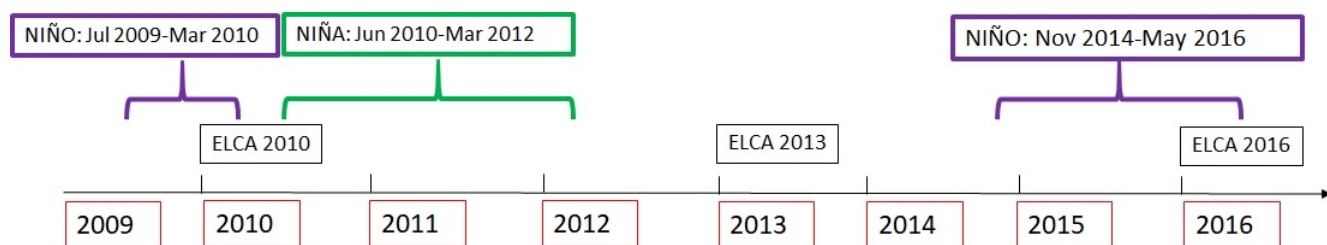
3C ELCA sample

Table 3.13: ELCA sample by municipality and year

Municipalities	2010	2013	2016	Total
	Observ.	Observ.	Observ.	Observ.
Saboya	435	428	385	1248
Cerete	318	312	293	923
Chinu	180	177	164	521
Cienaga de Oro	288	287	274	849
Sahagun	300	291	277	868
Simijaca	181	175	140	496
Susa	206	201	179	586
Tocaima	155	153	128	436
Circasia	217	212	158	587
Cordoba	76	71	57	204
Filandia	193	178	148	519
Belen de Umbria	557	543	450	1550
Puente Nacional	325	318	290	933
Sampues	130	127	109	366
Natagaima	187	185	163	535
Ortega	502	492	441	1435
Purificacion	261	257	230	748
Total	4511	4407	3886	12804

Note: created based on the ELCA

Figure 3.5: Time-line of ELCA surveys and El Niño phenomenon.



Note: created based on the ELCA and El Niño periods from the NOAA.

3D Test of climate and weather deviations with temperature squared

Table 3.14: Climate and Yearly weather deviations

	Value prod. total (Millions Col. Pesos/Ha) ln					
	MEAN TEMPERATURE			MAXIMUM TEMPERATURE		
	AGRO YEAR			AGRO YEAR		
	HIGH TEMP.	TEMP. SQUARED	BOTH	HIGH TEMP.	TEMP. SQUARED	BOTH
(1)	(2)	(3)	(4)	(5)	(6)	
Rainfall Trend 30 years (MEAN) (AGRO YEAR)	4.054** (1.798)	5.236*** (1.581)	4.972** (1.764)	4.358*** (1.446)	5.042*** (1.239)	4.547*** (1.452)
Rainfall Shock (MEAN) (AGRO YEAR)	0.250** (0.094)	0.265** (0.099)	0.276*** (0.087)	0.285*** (0.086)	0.309*** (0.091)	0.282*** (0.087)
Temperature Trend 15 years (MEAN) (AGRO YEAR)	2.208*** (0.645)	4.558*** (1.248)	0.000 (.)			
Temperature Shock (MEAN) (AGRO YEAR)	0.091 (0.083)	0.060 (0.071)	0.000 (.)			
High-Temp=1 × Temperature Trend 15 years (MEAN) (AGRO YEAR)	-1.202 (0.811)		2.185 (1.492)			
High-Temp=1 × Temperature Shock (MEAN) (AGRO YEAR)	-0.204*** (0.061)		-0.150 (0.103)			
Temperature Trend 15 years (MAX) (AGRO YEAR)				1.652*** (0.520)	2.607 (1.513)	0.000 (.)
Temperature Shock (MAX) (AGRO YEAR)				0.104** (0.036)	0.059 (0.036)	0.000 (.)
High-Temp=1 × Temperature Trend 15 years (MAX) (AGRO YEAR)				-0.695 (0.619)		0.932 (1.089)
High-Temp=1 × Temperature Shock (MAX) (AGRO YEAR)				-0.169*** (0.057)		-0.158 (0.109)
Temperature Trend 15 years squared (MEAN) (AGRO YEAR)		-0.057** (0.021)	-0.144** (0.051)			
Temperature Shock squared (MEAN) (AGRO YEAR)		-0.089** (0.034)	-0.021 (0.055)			
Temperature Trend 15 years squared (MAX) (AGRO YEAR)					-0.025 (0.023)	-0.080* (0.038)
Temperature Shock squared (MAX) (AGRO YEAR)					-0.039** (0.017)	0.001 (0.027)
Observations	1770	1770	1770	1770	1770	1770
r2	0.63	0.63	0.63	0.63	0.63	0.63
r2_a	0.30	0.30	0.30	0.30	0.30	0.30

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: based on the ELCA panel survey 2010-2013-2016. The standard errors (in parentheses) are clustered at the municipality level m to account for the correlation in treatment within municipality and the estimations use the inverse weights of 2010. Includes *region*, *year*, *municipality* and *household* fixed effects as well as agricultural problems of the household (except drought or other weather-related problems). It also includes controls for the dependency ratio in 2010, the age of the household head, a dummy for whether or not the household head is a woman, a dummy for whether or not the household has a credit in that year, a dummy at the community level for whether or not the access to the cabecera (center of the town) is reachable and in good shape, and the minutes to reach the cabecera for the community. High-temp is a dummy equal to one for the municipalities with average temperature higher than 27 Celsius degrees, and zero for the other municipalities. Sample restricted to rural households producing uniquely crops (excludes livestock producers or those producing livestock and crops).

3E Identification strategy and results using the SPI

I estimate the next equation at the level of the rural household h as:

$$Y_{h,m,r,t} = \eta + \delta \times \text{SPI}_{m,t} + \gamma_x \times X_{h,m,r,t} + \tau_r + \tau_t + \tau_m + \tau_h + \epsilon_{h,m,r,t} \quad (12)$$

$Y_{h,m,r,t}$ agriculture outcomes for rural household h (or agricultural unit), of municipality m , region r in year t ;

$\text{SPI}_{m,t}$ represents the Standard Precipitation Index;

$X_{h,m,r,t}$ vector of controls for household and community characteristics, as in equation 3.9;

$\tau_r, \tau_t, \tau_m, \tau_h$ dummies by *region*, by *year*, by *municipality* and by *household*.

The standard errors are clustered at the municipality level m to account for the correlation in treatment within municipality and the estimations use the inverse weights of 2010. The coefficients of interest are δ .

Table 3.15 shows the results of the effect of SPI on gross agricultural productivity, to compare with section 3.4.2 and section 3.4.4, including other agricultural outcomes. The "SPI 12 months (opposite)" captures the cumulative water deficit over twelve months during the calendar year (February of the previous year to January of the year) or agricultural growing season (February of the previous year to August of the previous year) (see section 3.2). The opposite of the measure is the SPI multiplied by minus one, so higher values correspond to droughts. I also consider the SPI for six months.

Panel a) of Table 3.15 shows no effect of the SPI on the total land used in hectares, or land used for crops but it reduced the land used for livestock, with no effect on the total value of gross agricultural productivity (in value terms). By crops, the SPI affected negatively the gross agricultural productivity (in value terms) of fruits, coffee, and to a lesser extent,

vegetables. Additionally, the negative impact on livestock is larger from the SPI at twelve months rather than six months, as the SPI 12 months captures more serious droughts. An increase in one unit of the SPI (values larger than one define droughts and values larger than two an extreme drought) reduces the value of the gross agricultural productivity of fruits by 13.9%-16.9%, (statistically significant at the 10% level). Panel b) of Table 3.15 measures adaptation in different forms, either in the labor market, or by investments. A one unit increase in the SPI indices reduce the probability of doing agricultural work by 6.1 to 7.6 percentage points for the household at the 10% significance level, without affecting the other labor market outcomes. The SPI increases the land plots with access to own sources of water by 4.4%-6.4%, but not to external water. Once households face drought problems, they might try to find their own sources of water for the land and depend less on external sources, that could be more affected by droughts. Additionally, there are increases of investments (excluding housing) of the rural household on the plots by 9.3%-12.1%.

Table 3.15: Impact of SPI on main agricultural outcomes

a) Production							
	Land (Ha) ln		Value prod. total (Millions Col. Pesos/Ha) ln				
	Crops (1)	Livestock (2)	Total (3)	Vegetables (4)	Fruits (5)	Cereals (6)	Coffee (7)
SPI 12 months (opposite)	-0.014 (0.022)	-0.076** (0.027)	-0.104 (0.076)	-0.080 (0.057)	-0.094 (0.055)	0.102 (0.090)	-0.095* (0.048)
SPI 12 months (opposite)-agro season	-0.007 (0.029)	-0.100** (0.036)	-0.099 (0.078)	-0.049 (0.059)	-0.169* (0.081)	0.141 (0.094)	-0.094* (0.048)
SPI 6 months (opposite)	-0.013 (0.014)	-0.043* (0.024)	-0.107 (0.069)	-0.113* (0.063)	-0.018 (0.035)	0.062 (0.085)	-0.088* (0.047)
SPI 6 months (opposite)-agro season	-0.024 (0.022)	-0.057 (0.039)	-0.128 (0.092)	-0.028 (0.081)	-0.139* (0.073)	0.101 (0.116)	-0.115** (0.049)
N	7902	7902	1770	1770	1770	1770	1770
r2	0.73	0.78	0.62	0.62	0.49	0.70	0.72
r2_a	0.56	0.64	0.29	0.29	0.04	0.45	0.49
Standard errors in parentheses							
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$							
b) Employment and Investment							
	HH Head or partner work			Land Water (ha) ln			Land INV(ha) ln
	Employed (1)	Agro. (2)	Off-farm (3)	Any (4)	Own (5)	Ext. (6)	ANY no house (7)
SPI 12 months (opposite)	-0.025 (0.030)	-0.076* (0.041)	0.033 (0.025)	0.024 (0.032)	0.044** (0.020)	-0.012 (0.028)	0.093** (0.037)
SPI 12 months (opposite)-agro season	-0.014 (0.038)	-0.068 (0.050)	0.049 (0.035)	0.038 (0.057)	0.064** (0.023)	-0.007 (0.051)	0.121** (0.049)
SPI 6 months (opposite)	-0.025 (0.021)	-0.061* (0.031)	0.015 (0.018)	0.010 (0.017)	0.023 (0.017)	-0.014 (0.019)	0.054 (0.044)
SPI 6 months (opposite)-agro season	-0.022 (0.040)	-0.074 (0.060)	0.025 (0.028)	0.027 (0.032)	0.031 (0.031)	-0.000 (0.034)	0.063 (0.060)
N	7902	7902	7902	7902	7902	7902	7902
r2	0.66	0.64	0.49	0.76	0.71	0.61	0.55
r2_a	0.44	0.40	0.17	0.61	0.53	0.36	0.27
Standard errors in parentheses							
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$							

Note: based on the ELCA panel survey 2010-2013. The standard errors (in parentheses) are clustered at the municipality level m to account for the correlation in treatment within municipality and the estimations use the inverse weights of 2010. Includes *region*, *year*, *municipality* and *household* fixed effects as well as agricultural problems of the household (except drought or other weather-related problems). It also includes controls for the dependency ratio in 2010, the age of the household head, a dummy for whether or not the household head is a woman, a dummy for whether or not the household has a credit in that year, a dummy at the community level for whether or not the access to the cabecera (center of the town) is reachable and in good shape, and the minutes to reach the cabecera for the community. High-temp is a dummy equal to one for the municipalities with average temperature higher than 27 Celsius degrees, and zero for the other municipalities. Each row corresponds to the coefficient estimated of the effect of the SPI index on the outcome variable in the column and are part of different estimations. The N, r2 and r2_a in each column corresponds to the estimations of the SPI in the first row but the explanatory power of the other estimations were almost the same when using the SPI agro season. In columns 3)-7) of Panel a), sample restricted to rural households producing uniquely crops (excludes livestock producers or those producing livestock and crops).

3F Balance tables Drought vs No Drought 2010

Table 3.16: Socioeconomic variables 2010 - Balance tables

Variable	Sample: Drought-No drought 2010 as T vs C			
	(1) No Drought	(2) Drought	(3) Diff	(4) N
Dependency ratio ($\leq 14 + \geq 65$)/(15-64) Per. 2010	82.710 (12.323)	78.270 (12.985)	6.784 (0.000)	17
Age HH head	46.752 (2.894)	47.323 (2.528)	-4.045 (0.000)	17
Dummy women HH head	0.186 (0.076)	0.184 (0.084)	0.007 (0.000)	17
Percentage crops (DROUGHT problem)	0.234 (0.187)	0.358 (0.241)	0.271 (0.000)	17
Percentage crops (PEST problem)	0.282 (0.086)	0.284 (0.088)	-0.006 (0.000)	17
Percentage crops (BRUSH problem)	0.075 (0.042)	0.045 (0.027)	0.014 (0.000)	17
Percentage crops (RAIN problem)	0.041 (0.025)	0.034 (0.023)	-0.034 (0.000)	17
Percentage crops (SEEDS problem)	0.011 (0.008)	0.013 (0.013)	-0.004 (0.000)	17
Percentage crops (VANDALISM problem)	0.010 (0.013)	0.003 (0.003)	-0.004 (0.000)	17
Percentage crops (OTHER problem)	0.069 (0.107)	0.029 (0.023)	-0.012 (0.000)	17
Percentage crops (NONE problem)	0.506 (0.188)	0.462 (0.189)	-0.067 (0.000)	17
Land (ha) (OWNED)	2.337 (1.330)	1.831 (0.600)	0.712 (0.000)	17
Land (ha) (OWNED WITH TITLE)	1.830 (1.051)	1.256 (0.384)	0.702 (0.000)	17
PCA 1: HH WEALTH	0.177 (0.078)	0.190 (0.051)	0.003 (0.000)	17
PCA 1: HH (LIVESTOCK ASSETS)	0.063 (0.441)	-0.089 (0.458)	0.322 (0.000)	17
PCA 1: HH (AGR. ASSETS)	-0.186 (0.113)	-0.135 (0.102)	0.108 (0.000)	17
Dummy HH has credit	0.502 (0.075)	0.445 (0.120)	0.090 (0.000)	17
Dummy access cabecera is ok (community)	0.283 (0.305)	0.518 (0.177)	-0.378 (0.000)	17
Minutes to reach cabecera (community)	44.384 (27.041)	40.517 (12.831)	5.430 (0.000)	17
Observations	11	6	17	

Note: uses ELCA panel survey only for 2010. Standard errors (in parentheses) clustered at the municipality level m to account for the correlation in treatment within municipality. The estimations use the inverse weights of 2010 as recommended in the documentation of the ELCA survey. It controls for *region*, *year* and *municipality* fixed effects. Treatment and control are determined based on whether there was a drought in 2010 in the municipality of the household. It uses the variable "Dummy 1 or more months drought (≤ -1)(SPI12)" to define a drought in 2010. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$

Table 3.17: Agricultural variables in 2010 - Drought vs No Drought in 2010

Variable	Sample: Drought-No drought 2010 as T vs C			
	(1) No Drought	(2) Drought	(3) Diff	(4) N
Agro sale (Millions Col. Pesos/Ha) ln	0.343 (0.192)	0.488 (0.234)	-0.341 (0.000)	17
Total Agro costs (Millions Col. Pesos/Ha) ln	0.425 (0.154)	0.484 (0.145)	-0.203 (0.000)	17
Percentage crops (YEARLY)	0.758 (0.210)	0.379 (0.086)	-0.219 (0.000)	17
Percentage crops (SEMIANNUAL)	0.132 (0.121)	0.392 (0.193)	0.261 (0.000)	17
Percentage crops (QUARTERLY)	0.037 (0.038)	0.146 (0.070)	-0.037 (0.000)	17
Percentage crops (MONTHLY)	0.034 (0.051)	0.066 (0.062)	-0.004 (0.000)	17
Percentage crops (BIMONTHLY-OTHER)	0.039 (0.054)	0.017 (0.012)	-0.000 (0.000)	17
Land size HH (Ha) ln	0.974 (0.269)	0.935 (0.254)	0.382 (0.000)	17
Total Area planted (Ha) (perman.+trans.+mixed) ln	0.380 (0.114)	0.441 (0.192)	0.313 (0.000)	17
Land PERMANENT crops (Ha) ln	0.178 (0.122)	0.200 (0.160)	0.145 (0.000)	17
Land TRANSITIONAL crops (Ha) ln	0.204 (0.138)	0.122 (0.073)	0.023 (0.000)	17
Land MIXED crops (Ha) ln	0.092 (0.091)	0.267 (0.205)	0.170 (0.000)	17
Land LIVESTOCK (Ha) ln	0.307 (0.246)	0.152 (0.103)	-0.056 (0.000)	17
Land PASTURE (Ha) ln	0.183 (0.119)	0.160 (0.102)	0.096 (0.000)	17
Land FOREST (Ha) ln	0.059 (0.086)	0.059 (0.022)	0.038 (0.000)	17
Land OTHER USES (Ha) ln	0.081 (0.028)	0.045 (0.029)	0.006 (0.000)	17
Land NO USED (Ha) ln	0.070 (0.029)	0.153 (0.107)	0.114 (0.000)	17
vr prodtot (Millions Col. Pesos/Ha of crops) ln	1.579 (0.504)	1.617 (0.554)	-0.582 (0.000)	17
vr prodtot-vegetable (Millions Col. Pesos/Ha of crops) ln	0.309 (0.361)	0.544 (0.387)	-0.527 (0.000)	17
vr prodtot-fruit (Millions Col. Pesos/Ha of crops) ln	0.248 (0.229)	0.179 (0.125)	0.144 (0.000)	17
vr prodtot-cereal (Millions Col. Pesos/Ha of crops) ln	1.053 (0.658)	0.824 (0.756)	-0.258 (0.000)	17
vr prodtot-coffee (Millions Col. Pesos/Ha of crops) ln	0.149 (0.280)	0.327 (0.388)	0.000 (0.000)	17
Tropical Livestock Units (FAO reference) ln	0.392 (0.266)	0.347 (0.105)	0.089 (0.000)	17

continue next page ...

... continuation

Variable	(1) No Drought	(2) Drought	(3) Diff	(4) N
HH head-partner employed	0.362 (0.110)	0.347 (0.185)	0.036 (0.000)	17
HH head-partner look for job	0.242 (0.109)	0.234 (0.098)	0.135 (0.000)	17
HH head-partner agro work	0.207 (0.106)	0.254 (0.186)	0.069 (0.000)	17
HH head-partner no agro work	0.238 (0.067)	0.208 (0.065)	0.039 (0.000)	17
HH head-partner Ave. wage (Millions Col. Pesos) ln	-1.291 (0.348)	-1.091 (0.304)	-0.145 (0.000)	17
HH head-partner Ave. hours worked month	44.097 (6.394)	48.029 (5.347)	5.581 (0.000)	17
Land (ha) (ANY WATER) ln	0.617 (0.281)	0.497 (0.270)	0.019 (0.000)	17
Land (ha) (OWN WATER) ln	0.449 (0.204)	0.380 (0.228)	0.081 (0.000)	17
Land (ha) (EXTERNAL WATER) ln	0.206 (0.217)	0.147 (0.096)	-0.063 (0.000)	17
Land (ha) ANY INVEST ln	0.299 (0.191)	0.365 (0.238)	0.065 (0.000)	17
Land (ha) IRRIGATION INVEST ln	0.009 (0.012)	0.005 (0.007)	-0.009 (0.000)	17
Land (ha) STRUCTURES INVEST ln	0.051 (0.052)	0.067 (0.083)	0.010 (0.000)	17
Land (ha) CONSERVATION INVEST ln	0.052 (0.057)	0.019 (0.021)	-0.012 (0.000)	17
Land (ha) FRUITS INVEST ln	0.026 (0.028)	0.084 (0.089)	0.077 (0.000)	17
Land (ha) WOOD INVEST ln	0.010 (0.009)	0.039 (0.041)	0.000 (0.000)	17
Land (ha) COMMERC. INVEST ln	0.006 (0.009)	0.115 (0.125)	-0.004 (0.000)	17
Land (ha) HOUSING INVEST ln	0.117 (0.158)	0.047 (0.016)	-0.004 (0.000)	17
Land (ha) OTHER INVEST ln	0.063 (0.046)	0.056 (0.026)	0.008 (0.000)	17
Land (ha) ANY NO HOUSING INVEST ln	0.199 (0.115)	0.331 (0.243)	0.063 (0.000)	17
Dummy after moving to another comm inside mpio 2010	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	17
Observations	11	6	17	

Note: based on the ELCA panel survey 2010. The standard errors (in parentheses) are clustered at the municipality level m to account for the correlation in treatment within municipality and the estimations use the inverse weights of 2010 as recommended in the documentation of the ELCA survey. It controls for *region*, *year* and *municipality* fixed effects. Treatment and control are determined based on whether there was a drought in 2010 in the municipality of the household. The statistics correspond to year 2010. It uses the variable "Dummy 1 or more months drought(≤ -1)(SPI12)" to define a drought in 2010. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

3G Effects on migration and heterogeneity with credit

Table 3.18: DID for 2010-2013-2016 for type of movement in the municipality

	Type of movement	
	Migrating outside Municipality (1)	Moving inside Municipality (2)
Drought 2010 (≤ -1) (SPI12)xPost(2013)	-0.036 (0.025)	0.050 (0.049)
Drought 2010 (≤ -1) (SPI6)xPost(2013)	-0.027* (0.013)	0.059** (0.025)
Drought 2010 (≤ -2) (SPI6)xPost(2013)	-0.111*** (0.009)	-0.057*** (0.017)
Drought 2010&2013 (≤ -1) (SPI12)xPost(2016)	-0.001 (0.010)	0.119*** (0.017)
Drought 2010&2013 (≤ -1) (SPI6)xPost(2016)	-0.003 (0.007)	0.049* (0.023)
FE for Rural Household	No	Yes
N	8845	7902
r2	0.06	0.66
r2_a	0.06	0.44

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: based on the ELCA panel survey 2010-2013-2016. The standard errors (in parentheses) are clustered at the municipality level m to account for the correlation in treatment within municipality and the estimations use the inverse weights of 2010. Includes *region*, *year* and *municipality* fixed effects as well as agricultural problems of the household (except drought or other weather-related problems). The fixed effects for rural household are not included in the estimations of migration (column 1) as they washed out all the variability of the explanatory variables. It also includes controls for the dependency ratio in 2010, the age of the household head, a dummy for whether or not the household head is a woman, a dummy for whether or not the household has a credit in that year, a dummy at the community level for whether or not the access to the cabecera (center of the town) is reachable and in good shape, and the minutes to reach the cabecera for the community. Each row corresponds to the interaction of post-treatment and Drought using the SPI on the outcome variable in the column. Each row is part of different estimations. The N, r2 and r2_a in each column corresponds to the estimations of the drought measure in the first row but the explanatory power of the other estimations were almost the same when using the other ones.

Table 3.19: DID for Drought in 2010 on main agricultural outcomes - Heterogeneity having credit
a) Production

	Land (Ha) ln		Value prod. total (Millions Col. Pesos/Ha) ln				
	Crops (1)	Livestock (2)	Total (3)	Vegetables (4)	Fruits (5)	Cereals (6)	Coffee (7)
Drought 2010 (≤ -1) (SPI12)xPost(2013)	-0.032 (0.052)	-0.004 (0.106)	0.148 (0.151)	-0.018 (0.263)	0.213 (0.200)	-0.308 (0.358)	0.177* (0.099)
Drought 2010 (≤ -1) (SPI12)xPost(2013)xCredit	0.188** (0.069)	0.124 (0.076)	-0.759** (0.263)	-0.540* (0.285)	0.022 (0.133)	-0.180 (0.254)	-0.090 (0.130)
Drought 2010 (≤ -1) (SPI6)xPost(2013)	0.021 (0.048)	-0.009 (0.100)	0.583 (0.583)	0.205 (0.591)	0.264 (0.260)	0.338* (0.181)	0.156** (0.057)
Drought 2010 (≤ -1) (SPI6)xPost(2013)xCredit	0.164*** (0.046)	0.143* (0.074)	-0.390 (1.018)	-0.845 (0.700)	-0.014 (0.145)	-0.281 (0.786)	0.020 (0.057)
Drought 2010 (≤ -2) (SPI6)xPost(2013)	0.229*** (0.032)	-0.133*** (0.044)	0.071 (0.153)	-0.524*** (0.126)	-0.148** (0.054)	0.431*** (0.130)	0.016 (0.106)
Drought 2010 (≤ -2) (SPI6)xPost(2013)xCredit	-0.121** (0.055)	0.078 (0.046)	-0.256 (0.150)	1.051*** (0.165)	0.083 (0.104)	-0.125 (0.128)	-0.514*** (0.065)
N	7902	7902	1770	1770	1770	1770	1770
r2	0.73	0.78	0.62	0.62	0.49	0.71	0.72
r2_a	0.56	0.64	0.29	0.29	0.04	0.45	0.48

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

b) Employment and Investment

	HH Head or partner work			Land Water (ha) ln			Land INV(ha) ln
	Employed (1)	Agro. (2)	Off-farm (3)	Any (4)	Own (5)	Ext. (6)	ANY no house (7)
Drought 2010 (≤ -1) (SPI12)xPost(2013)	0.052 (0.058)	0.073 (0.072)	-0.046 (0.074)	0.015 (0.090)	-0.104 (0.065)	0.110 (0.086)	-0.198** (0.080)
Drought 2010 (≤ -1) (SPI12)xPost(2013)xCredit	0.120 (0.078)	0.123* (0.059)	0.062 (0.109)	0.354*** (0.081)	0.150 (0.091)	0.246** (0.100)	0.161 (0.189)
Drought 2010 (≤ -1) (SPI6)xPost(2013)	0.125** (0.053)	0.221*** (0.044)	0.050 (0.105)	-0.018 (0.078)	-0.058 (0.077)	0.027 (0.080)	-0.200** (0.094)
Drought 2010 (≤ -1) (SPI6)xPost(2013)xCredit	0.090* (0.048)	0.154*** (0.041)	-0.036 (0.087)	0.159** (0.066)	0.024 (0.070)	0.182** (0.078)	0.029 (0.158)
Drought 2010 (≤ -2) (SPI6)xPost(2013)	0.242*** (0.043)	0.423*** (0.058)	0.110** (0.038)	-0.377*** (0.068)	-0.235*** (0.076)	-0.381*** (0.048)	-0.554*** (0.071)
Drought 2010 (≤ -2) (SPI6)xPost(2013)xCredit	0.071 (0.045)	-0.104** (0.043)	-0.150*** (0.030)	0.654*** (0.082)	0.055 (0.080)	0.937*** (0.076)	0.506*** (0.067)
N	7902	7902	7902	7902	7902	7902	7902
r2	0.66	0.63	0.49	0.76	0.71	0.61	0.55
r2_a	0.44	0.40	0.17	0.61	0.52	0.37	0.26

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: ELCA panel survey 2010-2013-2016. Standard errors (parentheses) clustered at the municipality level m . Using inverse weights of 2010. Includes *region*, *year*, *municipality* and *household* fixed effects as well as the same controls used in previous estimations (see section 3.4.2). Each row corresponds to the interaction of post-treatment and Drought in 2010 using the SPI on the outcome variable in the column. Each row is part of different estimations. N, r2 and r2_a in each column from the first drought measure estimations. In columns 3)-7) of Panel a), sample restricted to rural households producing uniquely crops (excludes livestock producers or those producing livestock and crops).

Bibliography

- Abadie, A., Athey, S., Guido W., I., and Wooldridge, J. (2017). When should you adjust standard errors for clustering? *NBER National Bureau of Economic Research, Technical Report*, 1(1):555–561.
- Abril-Salcedo, D. S., Melo-Velandia, L. F., and Parra-Amado., D. (2020). Nonlinear relationship between the weather phenomenon El Niño and Colombian food prices. *The Australian Journal of Agricultural and Resource Economics*, doi: 10.1111/1467-8489.12394(59):1–28.
- Almond, D., Chay, K., and Lee, D. S. (2005). The costs of low birth weight. *Quarterly Journal of Economics*, 120(3):1031–1083.
- Almond, D. and Currie, J. (2011a). Human capital development before age five. In Card, D. and Ashenfelter, O., editors, *Handbook of Labor Economics*, volume 4 Part B Chapter 15, pages 1315–1486. Elsevier.
- Almond, D. and Currie, J. (2011b). Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives*, 25(3):153–172.
- Anderson, M. (2014). Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion. *American Economic Review*, 104(9):2763–2796.
- Aragón, F. M., Oteiza, F., and Rud, J. P. (2021). Climate change and agriculture: Subsistence farmers’ response to extreme heat. *American Economic Journal: Economic Policy*, 13(1):1–44.
- Arceo, E., Hanna, R., and Oliva, P. (2015). Does the effect of pollution on infant mortality differ between developing and developed countries? Evidence from Mexico City. *The Economic Journal*, 126(1):257–280.

- Asfaw, S. and Maggio, G. (2017). Gender, weather shocks and welfare: Evidence from Malawi. *The Journal of Development Studies*, 54(2):271–291.
- Ashraf, N., Bau, N., Nunn, N., and Voena, A. (2020). Bride price and female education. *Journal of Political Economy*, 128(2).
- Auffhammer, M. and Schlenker, W. (2014). Empirical studies on agricultural impacts and adaptation. *Energy Economics*, 46(1):555–561.
- Baez, J., Caruso, G., Mueller, V., and Niu, C. (2017a). Droughts augment youth migration in Northern Latin America and the Caribbean. *Climatic Change*, 140:423–435.
- Baez, J., Caruso, G., Mueller, V., and Niu, C. (2017b). Heat exposure and youth migration in Central America and the Caribbean. *American Economic Review (Papers and Proceedings)*, 107(5):446–540.
- Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica*, 77(4):1229–1279.
- Barreca, A., Clay, K., Deschênes, O., Greenstone, M., and Shapiro, J. J. (2015). Convergence in adaptation to climate change: Evidence from high temperatures and mortality, 1900-2004. *American Economic Review: Papers and Proceedings*, 105(5):247–251.
- Barreca, A. and Schaller, J. (2020). The impact of high ambient temperatures on delivery timing and gestational length. *Nature Climate Change*, 10:77–82.
- Barrett, K., Lynch, M. J., Long, M., and Stretesky, P. B. (2017). Monetary penalties and noncompliance with environmental laws: a mediation analysis. *American Journal of Criminal Justice*, 43(3):530–550.
- Bazzi, S. (2017). Wealth heterogeneity and the income elasticity of migration. *American Economic Journal: Applied Economics*, 9(2):219–55.
- Beegle, K., Galasso, E., and Goldberg, J. (2017). Direct and indirect effects of Malawi’s public works program on food security. *Journal of Development Economics*, 128:1–23.
- Beegle, K. and Poulin, M. (2013). Migration and the transition to adulthood in contemporary Malawi. *The Annals of the American Academy of Political and Social Sciences*, 648(July):38–51.

- Beguera, S., Vicente-Serrano, S. M., F., R., and B., L. (2014). Standardized Precipitation Evapotranspiration Index - SPEI revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *International Journal of Climatology*, 34(10):3001–3023.
- Behrentz, E., Benavides, J. M., Bocarejo, J. P., Canal, M., Espinosa, M., Franco, N., Lobo, I. D., Pardo, O. A., and Sanchez, M. (2010). Decadal plan for air decontamination for Bogotá. *Report for the Secretary of Environment of Bogotá*, 1(1):1–324.
- Beine, M. and Jeusette, L. (2018). A meta-analysis of the literature on climate change and migration. CESifo Working Paper 7417. Retrieved from http://www.cesifo.org/DocDL/cesifo1_wp7417.pdf.
- Bento, A., Miller, N. S., Mookerjee, M., and Severnini, E. R. (2020). A unifying approach to measuring climate change impacts and adaptation. *NBER Working Paper Series*, 1(27247):1–50.
- Berlemann, M. and Steinhardt, M. F. (2017). Climate Change, Natural Disasters, and Migration - A Survey of the Empirical Evidence. *CESifo Economic Studies*, 63(4):353–385.
- Blanc, E. and Schlenker, W. (2017). The use of panel models in assessments of climate impacts on agriculture. *Review of Environmental Economics and Policy*, 11(2):258–279.
- Bound, J. and Krueger, A. (1991). The extent of measurement error in longitudinal earnings data: Do two wrongs make a right? *Journal of Labor Economics, University of Chicago Press*, 9(1):1–24.
- Branco, D. and Féres, J. (2020). Weather shocks and labor allocation: Evidence from rural Brazil. *forthcoming American Journal of Agricultural Economics*, pages 1–19.
- Brando, J. F. and Santos, R. J. (2015). La niña y los niños: Effects of an unexpected winter on early life human capital and family responses. *Documents CEDE-Los Andes University*, 1(013316):1–60.
- Burke, M. and Emerick, K. (2016). Adaptation to climate change: Evidence from US agriculture. *American Economic Journal: Economic Policy*, 8(3):106–140.
- Caminade, C., Turner, J., Metelmann, S., Hesson, J. C., Blagrove, M. S. C., Solomon, T., Morse, A. P., and Baylisa, M. (2016). Global risk model for vector-borne transmission of Zika virus

- reveals the role of El Niño 2015. *PNAS-Proceedings of the National Academy of Sciences of the United States of America*, 114(1):119–124.
- Cattaneo, C., Beine, M., Frohlich, C., Kniveton, D., Martinez-Zarzoso, I., Mastrorillo, M., Millock, K., Piguët, E., and Schraven, B. (2019). Human migration in an era of climate change. *Review of Environmental Economics and Policy*, 13(2):189–206.
- Chindarkar, N. (2012). Gender and climate change-induced migration: proposing a framework for analysis. *Environmental Research Letters*, 7(025601).
- Conley, D., Strully, K. W., and Bennett, N. G. (2001). Birth weight and income: interactions across generations. *Journal of Health and Social Behavior*, 42(4):450–465.
- Conley, D., Strully, K. W., and Bennett, N. G. (2006). Twin differences in birth weight: the effects of genotype and prenatal environment on neonatal and post-natal mortality. *Economics and Human Biology*, 4(2):151–183.
- Corno, L., Hildebrandt, N., and Voena, A. (2020). Age of marriage, weather shocks, and the direction of marriage payments. *Econometrica*, 88(3):879–915.
- Corno, L. and Voena, A. (2017). Selling daughters: age of marriage, income shocks and the bride price tradition. Unpublished manuscript.
- Costinot, A., Donaldson, D., and Smith, C. (2016). Evolving comparative advantage and the impact of climate change on agricultural markets: Evidence from 1.7 million fields around the world. *Journal of Political Economy*, 124(1):205–248.
- Crespin-Boucaud, J. (2020). Interethnic and interfaith marriages in sub-Saharan Africa. *World Development*, 125:1–19.
- Cui, X. (2020). Climate change and adaptation in agriculture: Evidence from US cropping patterns. *Journal of Environmental Economics and Management*, 101(102306):1–52.
- Currie, J., Neidell, M., and Schmieder, J. F. (2009). Air pollution and infant health: Lessons from New Jersey. *Journal of Health Economics*, 28:688–703.

- D'Agostino, A. L. and Schlenker, W. (2016). Recent weather fluctuations and agricultural yields: implications for climate change. *Journal of Agricultural Economics*, 47(1):159–171.
- de la Mata, D. and Gaviria, C. F. (2019). Exposure to pollution and infant health: evidence from Colombia. *Working Paper CINCH Competence in competition + health*, 1(1):1–50.
- Dell, M., Jones, B. F., and Olken, B. A. (2012). Temperature shocks and economic growth: evidence from the last half century. *American Economic Journal: Macroeconomics*, 4(3):66–95.
- Dell, M., Jones, B. F., and Olken, B. A. (2014). What do we learn from the weather? the new climate-economy literature. *Journal of Economic Literature*, 52(3):740–798.
- Dercon, S. (2002). Income risk, coping strategies, and safety nets. *World Bank Research Observer*, 17(2):141–166.
- Dercon, S. (2005). *Insurance against poverty*. Oxford University Press.
- Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., and Reif, J. (2019). The mortality and medical costs of air pollution: evidence from changes in wind direction. *American Economic Review*, 109(12):4178–4219.
- Deschênes, O. and Greenstone, M. (2007). The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather. *The American Economic Review*, 97(1):354–385.
- Deschênes, O. and Greenstone, M. (2011). Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the US. *American Economic Journal: Applied Economics*, 3(4):152–185.
- Deschênes, O., Greenstone, M., and Guryan, J. (2009). Climate change and birth weight. *American Economic Review: Papers and Proceedings*, 99(2):211–217.
- Deschênes, O., Greenstone, M., and Shapiro, J. J. (2017). Defensive investments and the demand for air quality: evidence from the NOx budget program. *American Economic Review*, 107(10):2958–2989.

- Deschênes, O. and Kolstad, C. (2011). Economic impacts of climate change on California agriculture. *Climate Change*, 109(1):365–386.
- Dillon, A., Mueller, V., and Salau, S. (2011). Migratory responses to agricultural risk in northern Nigeria. *American Journal of Agricultural Economics*, 93(4):1048–1061.
- Dingel, J. I., Hsiang, S., and Meng, K. (2020). Spatial correlation, trade, and inequality: Evidence from the global climate. *NBER WP*, DOI 10.3386/w25447(25447).
- DNP (2017). Costs on health associated with the environmental degradation in Colombia. *Report of the National Department of Planning of Colombia*, 1(1).
- Donaldson, D. and Storeygard, A. (2016). The view from above: applications of satellite data in economics. *Journal of Economic Perspectives*, 30(4):171–198.
- Driscoll, J. and Kraay, A. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *The Review of Economics and Statistics*, 80(4):549–560.
- Edmonds, E. and Shrestha, M. (2013). Independent child labor migrants. In Constant, A. and Zimmermann, K., editors, *International Handbook of the Economics of Migration*, pages 98–120. Edward Elgar Press.
- Eleftherios, G. and Ozgur, O. (2016). Structural equation modelling and the causal effect of permanent income on life satisfaction: the case of air pollution in Switzerland. *Journal of Economic Survey*, 30(3):430–459.
- Elorreaga, O. A., Huicho, L., and Lescano, A. G. (2020). El niño/southern oscillation (enso) and stunting in children under 5 years in Peru: a double-difference analysis. *The Lancet*, 8 Special Issue(29):29–29.
- Englund, H. (2002). The village in the city, the city in the village: Migrants in Lilongwe. *Journal of Southern African Studies*, 28(1):137–154.
- Fafchamps, M. and Shilpi, F. (2013). Determinants of the choice of migration destination. *Oxford Bulletin of Economics and Statistics*, 75(3):388–409.

- Fafchamps, M., Udry, C., and Czukas, K. (1998). Drought and saving in West Africa: are livestock a buffer stock? *Journal of Development Economics*, 55(1):273–305.
- FAO (2016). Evidence on internal and international migration patterns in selected African countries. Technical report, Rome. Also available as <http://www.fao.org/3/a-i7468e.pdf>.
- Fuertes, N., Tibavisco, M., Galeano, M. A., and Castaño, L. M. (2017). A descriptive analysis of two rural micro-regions of the ELCA: Cundiboyacense and the Eje Cafetero. *ELCA Reports*, October(9).
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., and Michaelsen, J. (2015). The climate hazards infrared precipitation with stations – A new environmental record for monitoring extremes. *Scientific Data* 2, doi:10.1038/sdata.2015.66(150066).
- Graff Zivin, J. and Neidell, M. (2012). The impact of pollution on worker productivity. *American Economic Review*, 102(7):3652–3673.
- Graff Zivin, J. and Neidell, M. (2013). Environment, health and human capital. *Journal of Economic Literature*, 51(3):689–730.
- Gray, C. and Mueller, V. (2012). Drought and population mobility in rural Ethiopia. *World Development*, 40(1):134–145.
- Grundström, M., Linderholm, H. W., Klingberg, J., and Pleijel, H. (2011). Urban NO₂ and NO pollution in relation to the North Atlantic oscillation NAO. *Atmospheric Environment*, 45:883–888.
- Guo, L.-C., Zhang, Y., Lin, H., Zeng, W., Liu, T., Xiao, J., Rutherford, S., You, J., and Ma, W. (2016). The washout effects of rainfall on atmospheric particulate pollution in two Chinese cities. *Environmental Pollution*, 215(1):195–202.
- Hampel, R., Lepeule, J., Schneider, A., Bottagisi, S., Charles, M.-A., Ducimetière, P., Peters, A., and Slamab, R. (2011). Short-term impact of ambient air pollution and air temperature on blood pressure among pregnant women. *Journal of Epidemiology*, 22:671–679.

- Hanna, R. and Oliva, P. (2015). The effect of pollution on labor supply: Evidence from a natural experiment in Mexico City. *Journal of Public Economics*, 122:68–79.
- Harris, I. and Jones, P. (2017). CRU TS4.01 Climatic Research Unit: Time-series version 4.01 of high-resolution gridded data of month-by-month variation in climate. *University of East Anglia-Centre for Environmental Data Analysis*, 04 December 2017.
- Hatton, T. J. and Williamson, J. G. (2003). Demographic and economic pressure on emigration out of Africa. *The Scandinavian Journal of Economics*, 105(3):465–486.
- Henry, S., Piche, V., Ouedrago, D., and Lambin, E. (2004a). Descriptive analysis of the individual migratory pathways. *Population and Environment*, 25(5):397–422.
- Henry, S., Schoumaker, B., and Beauchemin, C. (2004b). The impact of rainfall on the first out-migration: A multi-level event history analysis in Burkina Faso. *Population and Environment*, 25(5):423–460.
- Hertel, T. W. and de Lima, C. Z. (2020). Viewpoint: Climate impacts on agriculture: Searching for keys under the streetlight. *Food Policy*, 95(101954):1–12.
- Hirvonen, K. (2016). Temperature changes, household consumption, and internal migration: Evidence from Tanzania. *American Journal of Agricultural Economics*, 98(4):1230–1249.
- Hoffmann, R., Dimitrova, A., Muttarak, R., Crespo Cuaresma, J., and Peisker, J. (2020). A meta-analysis of country-level studies on environmental change and migration. *Nature Climate Change*. <https://doi.org/10.1038/s41558-020-0898-60>.
- Hoogeveen, J., van der Klaauw Bas, and van Lomwel, G. (2011). On the timing of marriage, cattle and shocks. *Economic Development and Cultural Change*, 60(1):121–154.
- Hotte, R. and Marazyan, K. (2019). Demand for insurance and within-kin-group marriages: Evidence from a West-African country. Working Paper.
- Hsiang, S., Meng, K., and Cane, M. (2011). Civil conflicts are associated with the global climate. *Letter Research Nature*, 476:438–441.

- Jayachandran, S. (2006). Selling labor low: Wage responses to productivity shocks in developing countries. *Journal of Political Economy*, 114(3):538–575.
- Jovanovic, N., Alwang, A., and Ricker-Gilbert, J. (2019). Short-term returns to agricultural household migration decisions: Evidence from a tracking panel data study in Malawi. 2019 Annual Meeting, July 21-23, Atlanta, Georgia, Agricultural and Applied Economics Association.
- Kaminski, J., Kan, I., and Fleischer, A. (2013). A structural land-use analysis of agricultural adaptation to climate change: A proactive approach. *American Journal of Agricultural Economics*, 95(1):70–93.
- Kazianga, H. and Udry, C. (2006). Consumption smoothing ? Livestock, insurance and drought in rural Burkina Faso. *Journal of Development Economics*, 79(2):413–446.
- Kelly, D. L., Kolstad, C. D., and Mitchell, G. T. (2005). Adjustment costs from environmental change. *Journal of Environmental Economics and Management*, 50(3):468–495.
- Knittel, C., Miller, D., and Sanders, N. (2016). Caution, drivers! children present: traffic, pollution, and infant health. *Review of Economics and Statistics*, 98(2):350–366.
- Ko, J., Myung, C.-L., and Park, S. (2019). Impacts of ambient temperature, DPF regeneration, and traffic congestion on NO_x emissions from a Euro 6-compliant diesel vehicle equipped with an LNT under real-world driving conditions. *Atmospheric Environment*, 200(1):1–14.
- Kolstad, C. D. and Moore, F. C. (2020). Estimating the economic impacts of climate change using weather observations. *Review of Environmental Economics and Policy*, 14(1):1–24.
- Kovats, S., Bouma, M. J., Hajat, S., Worrall, E., and Haines, A. (2003). El Niño and health. *The Lancet*, 362(9394):1481–1489.
- Kubik, Z. and Maurel, M. (2016). Weather shocks, agricultural production, and migration: Evidence from Tanzania. *Journal of Development Studies*, 52(5):665–680.
- Laffan, K. (2018). Every breath you take, every move you make: Visits to the outdoors and physical activity help to explain the relationship between air pollution and subjective wellbeing. *Ecological Economics*, 147:96–113.

- Lavaine, E. and Neidell, M. (2017). Energy production and health externalities: Evidence from oil refinery strikes in France. *Journal of the Association of Environmental and Resource Economists*, 4(2):1–31.
- Lewin, P. A., Fisher, M., and Weber, B. (2012). Do rainfall conditions push or pull rural migrants: Evidence from Malawi. *Agricultural Economics*, 43:191–204.
- Lichtenfels, A. J. F. C., Gomes, J. B., Pieri, P. C., Miraglia, S. G. E. K., Hallak, J., and Saldiva, P. H. N. (2007). Increased levels of air pollution and a decrease in the human and mouse male-to-female ratio in São Paulo, Brazil. *Fertility and Sterility*, 87(1):230–232.
- Lowes, S. and Nunn, P. (2018). Bride price and the wellbeing of women. In Anderson, S., Beaman, L., and Platteau, J.-P., editors, *Towards gender equity in development*. Oxford University Press.
- Makwemba, M., Chinsinga, B., Kantukule, C. T., Munthali, A., Woldegorgis, M., Haenni, S., and Lin, Q. (2019). Traditional practices in Malawi. Technical report, Center for Child Well-being and Development at the University of Zurich.
- McKee, T. B., Doesken, N. J., and Kleist, J. (1993). The relationship of drought frequency and duration of time scales. *Eighth Conference on Applied Climatology, American Meteorological Society*, Anaheim CA:179–186.
- Millock, K. (2015). Migration and environment. *Annual Review of Resource Economics*, 7(1):35–60.
- Mobarak, A. M., Kuhn, R., and Peters, C. (2013). Consanguinity and other marriage market effects of a wealth shock in Bangladesh. *Demography*, 50:1845–1871.
- Mueller, V., Gray, C., and Kosec, K. (2014). Heat stress increases long-term human migration in rural Pakistan. *Nature Climate Change*. DOI: 10.1038/NCLIMATE2103.
- Nichols, A. (2003). Vincenty: Stata module to calculate distances on the earth’s surface. *Statistical Software Components S456815, Boston College Department of Economics*, (revised 16 Feb 2007).
- Ortiz-Bobea, A., Ault, T. R., Carrillo, C. M., Chambers, R. G., and Lobell, D. B. (2021). Anthropogenic climate change has slowed global agricultural productivity growth. *Nature Climate Change*, 1(1):1–14.

- Ortiz-Bobea, A. and Just, R. E. (2013). Modeling the structure of adaptation in climate change impact assessment. *Journal of Agricultural Economics*, 95(2):244–251.
- Palmer, W. C. (1965). Meteorologic drought. *U.S Weather Bureau Research Paper*, 45(58):1–65.
- Robinson, A. L. and Gottlieb, J. (2019). How to close the gender gap in political participation: Lessons from matrilineal societies in Africa. *British Journal of Political Science*, pages 1–25. doi:10.1017/S0007123418000650.
- Rosenzweig, M. R. and Stark, O. (1989). Consumption smoothing, migration, and marriage: Evidence from rural India. *Journal of Political Economy*, 97(4):905–926.
- Rosenzweig, M. R. and Wolpin, K. I. (1993). Credit market constraints, consumption smoothing, and the accumulation of durable production assets in low-income countries: Investments in bullocks in India. *Journal of Political Economy*, 101(2):223–244.
- Schlenker, W. and Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *PNAS-Proceedings of the National Academy of Sciences*, 106(37):15594–15598.
- Schlenker, W. and Walker, W. R. (2016). Airports, air pollution, and contemporaneous health. *The Review of Economic Studies*, 83(2):768–809.
- Secretary-Environment, D. (2015). Annual report of air quality in Bogotá. *Report RMCAB*, 1:1–181.
- Sesmero, J., Ricker-Gilbert, J., and Cook, A. (2018). How do African farm households respond to changes in current and past weather patterns? A structural panel data analysis from Malawi. *American Journal of Agricultural Economics*, 100(1):115–144.
- Spinoni, J., Naumann, G., Carrao, H., Barbosa, P., and Vogt, J. (2014). World drought frequency, duration, and severity for 1951-2010. *International Journal of Climatology*, 34(1):2792–2804.
- Staiger, D. and Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65(3):557–586.
- Stenseth, N. C., Ottersen, G., Hurrell, J. W., Mysterud, A., Lima, M., Chan, K.-S., Yoccoz, N. G., and Alandsvik, B. (2003). Studying climate effects on ecology through the use of climate indices:

- the North Atlantic Oscillation, El Niño Southern Oscillation and beyond. *The Royal Society Review Paper*, 270:2087–2096.
- Thiede, B. and Gray, C. (2017). Erratum to: Heterogeneous climate effects on human migration in Indonesia. *Population and Environment*, 39:173–195.
- Vicente-Serrano, S. M., Beguería, S., and López-Moreno, J. I. (2010). A multi-scalar drought index sensitive to global warming: The Standardized Precipitation Evapotranspiration Index - SPEI. *Journal of Climate*, 23:1696–1718.
- Wan, Z., Hook, S., and Hulley, G. (2015). Mod11a1 modis/terra land surface temperature/emissivity daily 13 global 1km sin grid v006. *Distributed by NASA EOSDIS Land Processes DAAC*, <https://doi.org/10.5067/MODIS/MOD11A1.006>.
- Wang, G., Cai, W., Gan, B., Wu, L., Santoso, A., Lin, X., Chen, Z., and McPhaden, M. J. (2017). Continued increase of extreme El Niño frequency long after 1.5 Celsius degrees warming stabilization. *Nature Climate Change*, 7:568–572.
- Watson, A., Bates, R., and (Editors), D. K. (1988). *Air Pollution, the Automobile, and Public Health*. National Academies Press (US).

Summary

This dissertation analyzes different ways in which climate change and climatic phenomena can impact economic outcomes in low and middle income countries. The first chapter studies the migration responses following droughts in Malawi according to gender and stated motive of migration (for marriage and work-related reasons); in particular, how marriage-related institutions affect such migration patterns. The second chapter contributes to the literature on air pollution and health by assessing an additional channel, the effect of El Niño Southern Oscillation (ENSO or El Niño-La Niña events) on health. Unlike previous studies, it jointly investigates the effects of ENSO, air pollution and local weather on health at birth for the case of Bogotá. The final chapter analyzes how rural households in Colombia adapt to droughts and extreme heat, by exploring standard approaches and proposing an alternative way to capture climate and weather shocks on agricultural productivity. The different chapters explore the consequences of weather variability and climatic phenomena for rural households and individuals in different countries. The dissertation shows not only the different effects in rural and urban contexts of climatic variability, but also, that the relationships can be very complex across different domains, and context-dependent. The dissertation addresses empirically the different questions by merging information of surveys and administrative data with remote sensing information.

Keywords: climate change, droughts, gender migration, health at birth, ENSO events, weather, air pollution, agricultural adaptation.

Résumé

Cette thèse analyse les différentes manières dont le changement climatique et les phénomènes climatiques peuvent avoir un impact sur les résultats économiques dans les pays à revenu faible et intermédiaire. Le premier chapitre étudie les réponses migratoires suite aux sécheresses au Malawi en fonction du sexe et du motif déclaré de la migration (pour le mariage et pour des raisons de travail) ; en particulier, comment les institutions liées au mariage affectent ces schémas de migration. Le deuxième chapitre contribue à la littérature sur la pollution atmosphérique et la santé en évaluant un phénomène additionnel, l'effet de l'oscillation australe El Niño (ENSO ou événements El Niño-La Niña) sur la santé. Contrairement aux études précédentes, le chapitre étudie conjointement les effets de l'ENSO, de la pollution atmosphérique et du climat local sur la santé à la naissance dans le cas de Bogotá. Le dernier chapitre analyse la manière dont les ménages ruraux de Colombie s'adaptent aux sécheresses et aux chaleurs extrêmes, en explorant les approches conventionnelles et en proposant une manière alternative pour évaluer les effets des chocs climatiques et météorologiques sur la productivité agricole. Les différents chapitres explorent les conséquences de la variabilité météorologique et des phénomènes climatiques sur les ménages ruraux et les individus dans différents pays. La thèse montre non seulement les effets différents de la variabilité climatique dans les contextes ruraux et urbains, mais aussi que les relations peuvent être très complexes dans différents domaines, et dépendantes du contexte. La thèse met en oeuvre différentes méthodes en utilisant des données fusionnées issues de télédétection, d'enquête et de sources administratives.

Mots-clés: changement climatique, sécheresses, genre et migration, santé à la naissance, événements ENSO, météo, pollution atmosphérique, adaptation agricole.