Essays in Political Economy

Author:
Abelardo De Anda Casas

Supervisors:
Dr Sanjay Jain
Dr Paola Conconi

Department of Economics
University of Oxford

Thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy in Economics

August 2025

A mi familia: mis papás, mi hermano, y mis abuelos.
Y a mis amigos Francesco y Sydelle, que son como familia.
Con todo mi cariño.

Declaration of Authorship

I, Abelardo De Anda Casas, declare that the work comprising the thesis I am submitting is entirely my own. I also declare that I am aware of the University's regulations on plagiarism; that this work has not been submitted, either wholly or substantially, for another degree of this University, or for a degree at any other institution. I have clearly signalled the presence of quoted or paraphrased material and referenced all sources. I have not sought assistance from any professional agency. I have acknowledged appropriately any assistance I have received in addition to that provided by my supervisors. Details of any such assistance are specified in the acknowledgements of each chapter.

Abelardo De Anda Casas
Oxford, August 2025

Thesis word count: Approximately 47,000 words

Essays in Political Economy

Abelardo De Anda Casas

Thesis Abstract

This thesis is comprised of three independent chapters in the field of political economy. The first paper explores whether politicians can use means of direct communication with the citizenry, in order to consolidate their power. I study the case of Hugo Chávez's in Venezuela, focusing on the influence of the radio broadcasts of *Aló Presidente*. I rely on two sources of variation: incidental coverage of the radio station that carried *Aló Presidente* attributable to the natural ruggedness of the terrain; and variation over time in the content broadcast by this station, including the introduction of *Aló Presidente*. I identify an increase in Chávez's vote share in his first re-election attributable to the influence of *Aló Presidente*. Underlying mechanisms include the appeal of a narrative of democracy, and a deeply personalised style of communication.

The second chapter explores the potential influence of colonial enslavement on contemporary political preferences. I study the economic and political legacies of enslavement in colonial Venezuela. When enslaved people were brought into the continent they were registered and given the name of their ethnic group or place of origin as surnames. These surnames have persisted through time, allowing me to identify individuals with enslaved ancestry in contemporary Venezuela. I find a negative effect of enslaved ancestry on development outcomes and a positive effect in support for Hugo Chávez – an anti-system politician. The mechanisms of historical persistence are low human capital and low-skill occupational specialisation. Individuals with enslaved ancestry do not live in locations where enslaved labour was used during the colonial period. These locations are not economically worse-off today and were not more likely to support Chávez. These findings emphasise differences between mobile and place-based factors of historical persistence, explore a historical explanation to factors contributing to the rise of *Chavismo*, and highlight the long shadow of enslavement even in peripheral areas of a colonial empire.

The third chapter explores the economic dividends of peace. It is an empirical evaluation of a package of policies adopted in 2013 by the state of Coahuila in Mexico with the objective of reducing the presence of drug cartels. Two crucial dimensions included strengthening local police enforcement, and banning businesses that provided liquidity to cartels. Using synthetic controls, I test the effectiveness of these policies against all crime-types and high-frequency economic outcomes. I find a modest reduction in crime-types related with cartels, and no impact on uncorrelated crime-types. I find no statistical evidence of transformative economic effects beyond mechanical artefacts parsimoniously explained

by the policies adopted, such as a reduction in informality.

Acknowledgements

Thank you to my supervisors Sanjay Jain and Paola Conconi for all your support and advice. All my admiration, respect, and gratitude goes to you. Thank you for your patience and encouragement; for helping me remember the big picture while focusing on the details; for helping me prioritize clarity. Thank you for enabling me to explore new and exciting ideas while keeping them grounded with pragmatism, integrity, and humility. Thank you for helping me grow as an independent researcher. I look up to your wisdom, kindness, and generosity. These are qualities that I strive to embody in my professional and personal life. I could not have chosen a better team of supervisors for my DPhil.

I am thankful to the faculty from Oxford's Department of Economics, as well as members and convenors of the Political Economy, Development, and Applied Micro groups: Noam Yuchtman, Jean-Paul Carvalho, Dennis Egger, Barbara Petrongolo, Hamish Low, Simon Cowan, Andrea Bernini. Thank you for your encouragement, advice, support, and for nurturing an ever-stimulating academic environment during my time in Oxford as a student. Thank you as well to academics outside Oxford's Department of Economics: Toke Aidt, Gabriel León-Ablán, Simon Quinn, Sharun Mukand, David Doyle, Noam Angrist, Kamalini Ramdas. You have all been an invaluable network of support, knowledge, and inspiration.

Thank you to my many friends and peers: Ed, Wenzhang, Hannah, Momo, Giulio, Verena, Binta, Sanghamitra, Vatsal, Owen, Joe, Stan, Afonso, Juliette, Jean-Guillaume, Karthik, Jordan, Alena, Manos, Inbar, Nan, Anya, Carlos, Michael, Amelie, Lovisa, Jasmin, Nithya, Antonia, Niccolò, Rocco, Wahed, Nick, Thiago, Krittika, Gaurav, Prabmheet, Sam K., Philipp, Andrés, Sam M., Karl. Thank you as well to the team at the What Works Hub for Global Education: Pedro, Julio, Rute, Anaya, Abraham, Veerangna, Sam H., Sharnic, Erad, Madhuri, Heather, Yue-Yi, Guillermo, Natasha. You have all shaped my thoughts and work in ways for which I am grateful.

My appreciation also goes to St Hugh's College for providing a welcoming, open, and diverse community outside my Department. The garden and library at St Hugh's remain two of my favourite places in Oxford after all these years. I also want to thank all the students I had the fortune of tutoring during the past 4 years, as well as the Colleges where I taught: Queen's, St Catherine's, University, and St Anne's. To my students: you have taught me more lessons than you can imagine. You are a source of pride and inspiration.

A bit *ad hoc*, but I want to express my gratitude to the music that kept me company during my postgraduate studies. The music of Max Richter provided a space for calm reflection, and solace during the pandemic. I worked for hours listening to Anna Christiane Neumann's piano transcriptions of Bach's choral pieces, as well as Víkingur Ólaffson's

interpretations of Glass, Bach, Rameau, and Mozart.

In my memories, Oxford also sounds like Syd singing Radiohead's *Fake Plastic Trees*. Lanterns on the Lake's *When It All Comes True* reminds me of cycling in Jericho at night. Cassandra Jenkins' *An Overview of Phenomenal Nature* reminds me of spring and summer in Oxford – of the Thames, Christ Church Meadow, and being outside with friends; of sparkling water on my face. More music that marked my time as a student includes Beach House's *Bloom*, Talk Talk's *Spirit of Eden*, Weyes Blood's *Titanic Rising*, Jóhann Jóhannsson's *Orphée*.

Finally, I want to thank my family: my mother, father, brother, and grandparents for their never-ending love and support. I would not be who I am without you. My gratitude and love also goes to Francesco Feriozzi and Sydelle de Souza. You know me better than I (sometimes) know myself. You have all been an unconditional and endless source of support, advice, laughter; understanding, friendship, and love. And for that I will always be grateful. You are the ones that give meaning to it all. I dedicate my thesis to all of you.

Contents

Front Matter

Declaration of Authorship . . . iii
Thesis Abstract . . . v
Acknowledgements . . . vii

Introduction **1**

1 *Aló Presidente*: Direct Political Communication and Electoral Outcomes in Chávez's Venezuela **5**

1.1 Introduction . . . 6
1.2 Historical Context . . . 10
1.3 Data Sources . . . 14
1.4 Main Results . . . 18
1.4.1 Effect of Exposure to RNV . . . 18
1.4.2 Are These Results an Artefact of Political Mobilisation? . . . 22
1.4.3 Is this Effect Due to Persuasion? . . . 22
1.4.4 Summary of Main Results . . . 24
1.5 Summary of Robustness Checks . . . 25
1.6 Exploring Mechanisms . . . 28
1.7 Conclusions and Limitations . . . 30
Appendix . . . 32
1.A Additional Figures and Tables . . . 32
1.A.1 Additional Figures . . . 32
1.A.2 Additional Tables . . . 34
1.B Details on Data Sources and Empirical Strategy . . . 39
1.B.1 Documents Used to Infer List of RNV Towers . . . 39
1.B.2 Simulating Radio Signals . . . 47
1.B.3 Field Strength Calculations . . . 48
1.B.4 Details on Mediation Analysis . . . 51

1.C Details of Heterogeneities Analysis . . . 56
1.C.1 Demographic Summary Statistics and Variable Creation . . . 56
1.C.2 Tree-based Analysis . . . 56
1.D Text Analyses . . . 60
1.D.1 Pre-processing Data . . . 60
1.D.2 Data Analyses . . . 60
1.E Robustness . . . 63
1.E.1 Conley Standard Errors . . . 63
1.E.2 Family-wise Error Rates . . . 63
1.E.3 Changing Percentage of Population Covered . . . 64
1.E.4 Changing Signal Strength Threshold . . . 65
1.E.5 Percentage of Covered Area, Rather than Population . . . 66
1.E.6 Changing Tower Network Composition . . . 68
1.E.7 Exposure to a Placebo Network . . . 69
1.E.8 Subsampling . . . 70
1.E.9 Sharing a Border with Colombia . . . 72
1.E.10 Correct for Non-Random Response to Random Shocks . . . 73
1.F Alternative Determinants of Electoral Turnout . . . 82
1.G Details of Survey Analysis . . . 88
1.H Historical Context Details . . . 93

2 Poverty and Low Human Capital: The Legacy of African Slavery in Venezuela and the Deep Historical Roots of *Chavismo* 95

2.1 Introduction . . . 96
2.2 Historical Context . . . 101
2.3 Data Sources . . . 104
2.4 Empirical Strategy . . . 107
2.4.1 Wheat Suitability as a Valid Instrument . . . 107
2.4.2 Instrument's Relevance and First Stage Results . . . 108
2.5 Main Results . . . 109
2.5.1 Enslaved Ancestry and Development Outcomes . . . 109
2.5.2 Persistence Mechanism . . . 110
2.5.3 Enslaved Ancestry and Votes for Chávez . . . 112
2.5.4 Locations Exposed to the Practice of Slavery . . . 117
2.6 Additional Robustness Checks . . . 119
2.7 Conclusions and Limitations . . . 122

Appendix . . . 125
2.A Additional Figures and Tables . . . 125
2.A.1 Appendix Figures . . . 125
2.A.2 Appendix Tables . . . 131
2.B Additional Details on Data Sources . . . 151
2.B.1 Historical Maps of Venezuela . . . 151
2.B.2 Trans-Atlantic and Intra-American Slave Trade Databases . . . 155
2.B.3 African Slave Surnames in Acosta Saignés (1984) . . . 156
2.B.4 The Tascón-Maisanta Database . . . 157

3 The Economic Consequences of Peace: The Economic Effects of Successfully Combating Drug Cartels in Mexico 160
3.1 Introduction . . . 161
3.2 Context . . . 164
3.3 An Overview of the Synthetic Controls Methodology . . . 167
3.4 Data Sources and Implementation . . . 171
3.5 Main Results . . . 173
3.5.1 'Intentional Homicides' . . . 173
3.5.2 Other Crime-Types . . . 173
3.5.3 The Economic Effects of Peace . . . 175
3.6 Robustness Checks . . . 177
3.6.1 Placebo Policy Dates . . . 178
3.6.2 Swapping Treated and Donor Units . . . 178
3.6.3 Spillovers to Donor Units . . . 179
3.6.4 Comparison to Canonical Synthetic Controls . . . 180
3.7 Conclusions and Limitations . . . 183
Appendix . . . 184
3.A Additional Figures . . . 184
3.B Additional Tables . . . 193
3.C Code . . . 196

Concluding Remarks 203

List of References 220

Introduction

This thesis is comprised of three independent chapters in the field of political economy. The first chapter focuses on the role of direct political communication. It asks whether politicians can use means of communication to directly address voters in order to consolidate their power. The paper tackles this question by studying the case of Hugo Chávez in Venezuela, focusing on the influence of *Aló Presidente*, a weekly unscripted radio programme that featured call-ins from citizens, which began broadcasting through Venezuela's national broadcasting agency, *Radio Nacional de Venezuela* (RNV), in 1999 shortly after Chávez's presidency started, and was widely popular.

I rely on two key sources of variation to quantify the influence of *Aló Presidente*. First, I exploit variation in RNV's coverage attributable to the natural ruggedness of the terrain (e.g., Adena, Enikolopov, Petrova, Santarosa, & Zhuravskaya, 2015; Blouin & Mukand, 2019; Olken, 2009). The second source of variation is changes in RNV's content over time: I measure the effect of exposure to RNV on three sets of high-stakes political outcomes measured before Chávez was president, during Chávez's presidency but before *Aló Presidente* began broadcasting, and after the introduction of *Aló Presidente*.

I identify a strong and robust increase in Chávez's vote share in his first re-election attributable to the influence of *Aló Presidente*. Placebo tests rule out that results are driven by exposure to RNV in general, or RNV during Chávez's presidency. Evidence suggests the mechanisms underpinning this effect are a narrative of democracy and a deeply personalised style of communication. I am able to rule out other potential mechanisms such as a dis-intermediation of political parties, or by-passing traditional means of communication.

This paper contributes to the literature in two ways. First, it is the first paper at the intersection of the literatures on mass-media (e.g., DellaVigna & Gentzkow, 2010; Zhuravskaya, Petrova, & Enikolopov, 2020) and populism (Guriev & Papaioannou, 2022) that studies the direct use of means of communication by a left-wing populist leader. The closest work would be Adena et al. (2015) which studies the impact of radio in Nazi Germany. This chapter innovates by studying the electoral returns of the introduction of a direct channel of communication between politicians and their electorate.

The second contribution is related to the literature that focuses on Hugo Chávez's regime and its legacy. Here, the closest work would be Knight and Tribin (2019, 2022), which examine 'tuning-out' of opposition voters when exposed to Chávez's propaganda, as well as censorship and backlash. Importantly, these papers study Venezuela post-2006 under a fully autocratic regime. In contrast, this chapter studies the introduction of a

means of direct political communication during a key transition phase at the start of Chávez's presidency between 1998 and 2000, when he was able to consolidate his regime by creating a narrative around democracy. More broadly, this new set of findings links the work of Dornbusch and Edwards (1991) and Edwards (2010) on Latin-American populism with the recent work of Funke, Schularick, and Trebesch (2023), and the literature on democratic backsliding (e.g., Levitsky & Ziblatt, 2018).

The second article explores the potential influence of colonial enslavement on contemporary political preferences. I tackle this question by studying the use of enslaved African people in colonial Venezuela, its effect on contemporary economic and political outcomes, and the reasons for the historical persistence of those outcomes. When brought to Venezuela, enslaved individuals were registered and given the name of their ethnic group or place or origin as surnames (Acosta Saignes, 1984; Ortega, 1992). This unique set of surnames has persisted through time and serves as an indicator of enslaved ancestry. I am able to identify individuals with enslaved ancestry by combining historical sources with a contemporary roster of voters in the country – the so-called Maisanta list, which I describe in detail in the chapter.

As a complement, I digitise new historical maps of Venezuela which allow me to identify the location of the cocoa *haciendas* where enslaved people were brought during the colonial period. In combination with the Trans-Atlantic and Intra-American Slave Trade Databases (Eltis, 1999), these sources provide a place-based measure of exposure to the colonial use of enslaved labour. I explore the effect of enslaved ancestry and exposure to the use of enslaved labour on a battery of development outcomes and support for Hugo Chávez, and rely on census data to explore their historical persistence.

To push causality, I adopt a two-stage least squares identification strategy instrumenting for the independent variables with crop suitability (akin to Acharya, Blackwell, & Sen, 2015, 2016b). Following the abolition of slavery, many formerly enslaved people migrated to unoccupied land at the fringe of estates and became subsistence farmers (Brito Figueroa, 1966, 1973). I use the suitability of the soil for producing wheat to predict the location of individuals with enslaved ancestry, given that wheat was a staple in subsistence farmers' diets, it was widely harvested, but not for exports; and it was easy to grow in land that was not already used for growing other crops (Codazzi, 1960; Codazzi, Baralt, & Díaz, 1940). Conversely, I use cocoa soil-suitability to predict the location of colonial cocoa *haciendas*. In both approaches, I argue for the validity of the instruments leveraging a comprehensive list of covariates, and test for robustness under alternative specifications.

I identify a negative effect of enslaved ancestry on development outcomes: night-light density, literacy, and child mortality. I find correlational evidence suggesting the

mechanisms for persistence are low human capital and low-productivity occupational specialisation. I observe a positive effect on Hugo Chávez's vote share in 1998. I explain this finding through Chávez's anti-system rhetoric, since enslaved ancestry is correlated with dissatisfaction with politics, distrust of politicians, and zero-sum thinking prior to Chávez's presidency. I confirm that individuals with enslaved ancestry are not located close to cocoa *haciendas*. I find no evidence of a place-based effect of exposure to the colonial practice of enslavement on development outcomes or Chávez's vote share.

This paper contributes to the literature in three ways. First, to the best of my knowledge, it is the first paper in the persistence literature to consider historical factors that contributed to the rise of a left-wing populist politician. The closest work would be Althoff and Reichardt (2024) or Nunn and Wantchekon (2011), which explore the negative long-term effects of enslavement on human capital accumulation in the United States, and inter-personal trust across Africa, respectively. In a recent working paper, Becker and Voth (2023) explore the persistence of culture and (absence of) religion as factors that might have historically contributed to the rise of Nazism in Germany. The mechanisms of persistence explored in this chapter point to low human capital and low-productivity occupational specialisation as factors historically contributing to poverty and the rise of *Chavismo*.

Second, by testing the effects of enslaved ancestry and exposure to the practice of enslavement, this paper differentiates between mobile and place-based factors of historical persistence. Conceptually, the first group includes culture and human capital, the second group includes laws or property rights. Identifying the persistence or lack thereof, of these factors in the same context represents a unique and new set of findings. Third, the results emphasise the long shadow cast by enslavement even in peripheral areas of a colonial empire. The experience of Venezuela stands in contrast to other Latin American areas colonised earlier and under comparatively more extractive institutional arrangements (Mahoney, 2010), such as Peru (Dell, 2010) or Colombia (Acemoglu, García-Jimeno, & Robinson, 2012).

The third chapter is a policy evaluation that explores the economic dividends of peace. I evaluate a set of policies that were adopted in 2013 by the state government of Coahuila in Mexico with the objective of reducing the presence of drug cartels. The complex policy-package had two crucial dimensions: strengthening local police-enforcement, and banning businesses that provided liquidity to cartels. These businesses included brothels, junk-yards, or casinos, for instance. I evaluate the effectiveness of these policies against all reported crime-types, and explore whether they had any transformative economic effects.

Coahuila's government claims the package of policies was highly successful by comparing levels of intentional homicides before and after the adoption of the policies. I test this

claim formally by constructing valid counterfactuals via synthetic controls (Abadie, 2021; Abadie & Gardeazabal, 2003), and evaluate the policies' effect on all crime-types and a set of high-frequency economic indicators. For inference, I apply the recent advances in Chernozhukov, Wüthrich, and Zhu (2021) to construct confidence intervals for all observations in the post-treatment period. I complement these analyses by adopting a synthetic differences-in-differences strategy (Arkhangelsky, Athey, Hirshberg, Imbens, & Wager, 2021) that yields similar results. Overall, I confirm a modest reduction in crime-types associated with the presence of cartels and no effect on crime-types unrelated to cartels. I observe no transformative impacts on economic indicators, beyond mechanical artefacts plausibly caused by the adoption of the policies themselves. These include, for instance, a reduction in informality.

The paper innovates in two ways. First, it contributes to the literature on conflict and violence by examining the economic consequences of peace. The economic drivers of conflict have been well-documented in the past (see e.g., Ray & Esteban, 2017), as well as conflict's economic consequences (see e.g., Rohner & Thoenig, 2021). Yet, identifying the economic dividends of peace has remained elusive with the occasional exception – for instance, of the work of Besley and Mueller (2012). Second, the paper offers an empirical contribution, by applying tools of causal inference to evaluate an original policy with implications for policy-makers interested in effectively fighting organised-crime and reducing violence in their contexts. From a practical standpoint, the paper contributes to applied researchers by incorporating recent methodological advances in econometrics (Chernozhukov et al., 2021) and providing code to implement these techniques more easily.

The rest of the thesis is organised as follows. Each paper is presented as an independent chapter. Each chapter includes an appendix with additional figures, tables, details on data sources, analyses, code, or contextual information. The thesis ends with a section of concluding remarks, which discusses limitations of the current papers and identifies productive areas for future research. The list of references cited throughout the thesis is provided at the end of the document.

Chapter 1

Aló Presidente: Direct Political Communication and Electoral Outcomes in Chávez's Venezuela

Abstract

Can political leaders use means of communication directly to consolidate their power? After the electoral victory of Hugo Chávez in 1998, and in a span of six years, Venezuela transitioned from a democracy into an autocracy. State propaganda used by Chávez in the shape of *Aló Presidente*, a weekly unscripted radio talk show broadcast through *Radio Nacional de Venezuela* (RNV), played a crucial role in this process. I analyse the impact of exposure to RNV, relying on the natural ruggedness of the terrain as a source of plausibly exogenous variation. I identify an increase in Chávez's vote share in his first re-election, which cannot be fully accounted for by changes in electoral turnout. Underlying mechanisms include the appeal of a narrative of democracy as well as a deeply personalised style of communication. Additional variation over time in RNV's content allows me to rule out that results are due to exposure to radio in general or radio under Chávez's regime: results appear only after *Aló Presidente* was introduced.

Acknowledgements: I want to thank Sanjay Jain and Paola Conconi for their support throughout this project. Thank you to Sharun Mukand, Noam Yuchtman, Ruben Enikolopov, Ekaterina Zhuravskaya, Toke Aidt, David Doyle, Gabriel León-Ablán, Simon Quinn, Hamish Low, Andrea Bernini, Jean-Paul Carvalho, Anandi Mani, Pramila Krishnan, Dennis Egger, Nicolas Cerkez, Niccolò Merigi, Stefan Dercon, Christian Meyer, Noam Angrist, Debraj Ray, Claudia Rei, James Fenske, Krisztina Kis-Katos, Nicola Limodio, Sahar Parsa, Clement Imbert, Valeria Rueda, Clement Minaudier, Daniela Horta, Bjorn Brey, Yannik Schenk, Marcel Caessmann, Eduard Krkoska, Giulio Schinaia, Shihang Hou, Wenzhang Zhang, Liyu Zheng, Rocco Zizzamia, Michael Eldar, Sanghamitra Mukherjee, Verena Wiedemann, Karl Welzel and Sydelle de Souza who provided valuable comments, questions, and feedback at various stages of the project. Thank you to participants and discussants at the CSAE Workshop in January 2023, the CAGE Summer School 2023 at the University of Warwick, Oxford's Political Economy Seminar in 2023 and 2024, Oxford's CSAE Conference 2024, the Royal Economic Society Annual Conference 2024, the European Public Choice Society Meeting 2024, Oxford's Economics Research Jamboree 2024, OxDev 2024, and POLECONUK 2025 for all their valuable comments, questions and suggestions. Thank you to Dorothy Kronick, who provided me with the electoral outcomes data. I am thankful for the support that I received from the Social Sciences Library (SSL), part of the Bodleian Library at the University of Oxford, in digitising the data from the 1990 Venezuelan Census. Specifically, I want to thank Craig Finlay, Rachel Scanlon, Ana Da Silva and the rest of the weekend staff at SSL who carefully helped me scan the books comprising the census results. All mistakes and omissions are my own responsibility.

1.1 Introduction

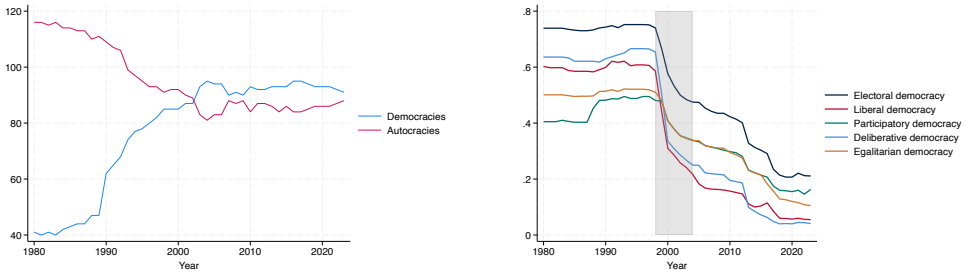
> *Every government leader uses the media to justify and persuade, project and burnish, but none like Chávez. (...) His word was de facto law (...). The man was a consummate performer. He would sing, dance, rap; (...) act the fool, the statesman, the patriarch. There was a freewheeling, improvised air to it all. Suspense came from not knowing what would happen.*
>
> *- Rory Carroll, Comandante: The life and legacy of Hugo Chávez*

This paper quantifies the influence of direct political communication from a populist leader in the consolidation of his political power. The unconventional use of media is an ubiquitous characteristic of populist politicians (see e.g., Eichengreen, 2018; Zhuravskaya et al., 2020). Donald Trump preferred Twitter and now Truth; Jair Bolsonaro, Facebook. For the case of Hugo Chávez, it started with radio – specifically with *Aló Presidente*. Venezuela represents an ideal setting for studying the role of direct communication by a populist leader given its quick transition into an autocratic regime, and Chávez's unique use of means of communication.

While the recent literature on populism shows that economic performance and the quality of bureaucrats worsen following the election of a populist candidate (Bellodi, Morelli, & Vannoni, 2024; Funke et al., 2023), democracy across the world has been mostly resilient to the recent populist wave. This, however, need not be the case everywhere (e.g., Enikolopov, Petrova, & Zhuravskaya, 2011). The world is currently at a crossroads where the spread of democracy has stagnated and possibly even begun to decline (see Panel (a) of Figure 1.1). Further, as populist rhetoric spreads across the world (Guriev & Papaioannou, 2022; Lewis et al., 2019), we should look for lessons of how democracy has been lost at the hands of populist leaders in the past, specially when they use mass media directly.

I study the influence of *Aló Presidente* – Hugo Chávez's weekly unscripted talk show, broadcast through *Radio Nacional de Venezuela* (RNV) – during a key period of transition in Venezuela following Chávez's election in 1998 (see Panel (b) of Figure 1.1). For identification, I rely on two sources of variation: incidental variation in the coverage of radio attributable to the natural ruggedness of the terrain (see Blouin & Mukand, 2019; Enikolopov et al., 2011; Olken, 2009); and, variation over time in content broadcast by RNV that allows me to perform a series of placebo tests to sustain that propaganda[1] only comes into effect after

[1] I use the term propaganda as employed by Jowett and O'Donnell (2015), wherein propaganda is a subcategory of persuasion, and is defined as 'the deliberate, systematic attempt to shape perceptions, manipulate



(a) Democracies and Autocracies in the World

(b) V-Dem Main Indices Following Chávez's Election in Venezuela

Notes: Panel (a) shows the number of countries classified as liberal or electoral democracies, and electoral or closed autocracies by Herre (2022). The figure in Panel (b) shows the five main indices of the V-Dem database for Venezuela from 1980 until 2022. In grey, the figure highlights the years this paper focuses on: 1998-2004.

Figure 1.1: Democracy Across the World and in Venezuela

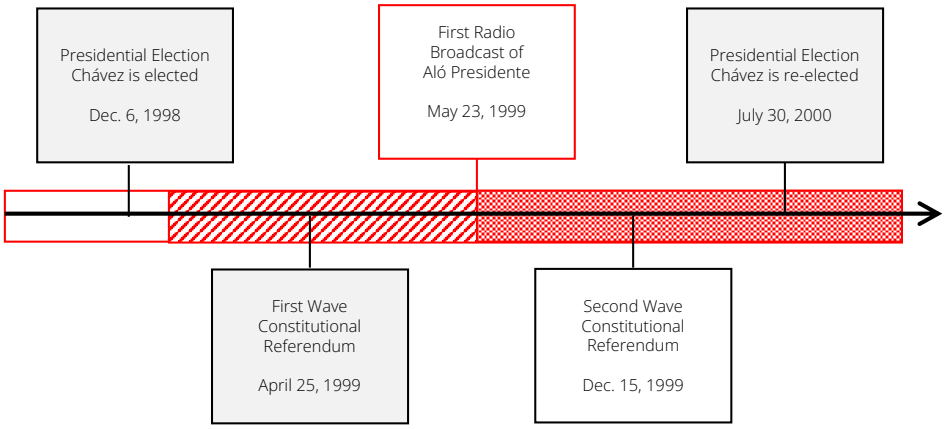
Aló Presidente was introduced.

Two key political changes took place in Venezuela during this period of time: the Constitutional Referendum of 1999 and the 'Mega-Election' of 2000 (see Figure 1.2). The adoption of the new Constitution as a result of Chávez's election meant a deep institutional transformation for the country. It eroded democracy by strengthening Chávez's regime: concentrating power in the president, removing political counter-weights, allowing the re-election of the president, extending the presidential mandate to 6 years, allowing recall referenda, among other changes. The new Constitution also called for the re-election of all politicians in office. This episode known as the 'Mega-Election' of 2000 was the first of several of Chávez's re-elections, allowing him to fully consolidate his power.

I find that exposure to RNV had a strong impact in favour of Chávez during the 'Mega-Election' of 2000. Exposure to RNV had no impact on Chávez's first election in 1998. It also had no impact on the first wave of the Constitutional Referendum of 1999 – which took place just two months after Chávez was inaugurated (see Figure 1.2).[2] These two results nuance the impact found on Chávez's share in 2000. First, finding no effect on the election of 1998 suggests results are not driven simply by being exposed to RNV's coverage. Second, finding no effect in the first wave of the Constitutional Referendum suggests results are not driven by being exposed to media under Chávez's regime either. The key difference in the content broadcast through RNV between the first wave of the

cognitions, and direct behaviour to achieve a response that furthers the desired intent of the of the propagandist.'

[2] The Constitutional Referendum of 1999 took place across two waves: the first wave in April 1999 asked Venezuelans whether they wanted to summon a Constitutional Assembly to draft a new constitution, and whether they agreed with a set of rules defined by Chávez to do this. The second wave took place in December 1999 and was a vote for the draft of the new constitution. More details are provided in Section 1.2. Unfortunately, to the best of my knowledge, the disaggregated results of the second wave of the Referendum are not available.



Notes: Study timeline. Gray boxes indicate available data. The red area around the arrow represents how I interpret exposure to RNV. The first, hollow, section represents exposure to RNV before Chávez's government. The second section with diagonal filling represents exposure to RNV under Chávez's government. The third section with a cross pattern represents exposure to RNV after the introduction of *Aló Presidente*.

Figure 1.2: Timeline

Constitutional Referendum and the 'Mega-Election' of 2000 was the introduction of *Aló Presidente*. These findings are highly robust, including correcting for potential omitted variables akin to correcting for the non-random response to exogenous shocks (Borusyak & Hull, 2023) by assigning counterfactual topographies to the country.

I find that RNV also decreased electoral turnout in the 'Mega-Election' of 2000.[3] I probe whether the results on Chávez's share are due to political mobilisation or a change in who individuals vote for – what I call persuasion.[4] Overall, the increase in Chávez's share cannot be fully accounted for by changes in turnout. I complement these findings by analysing non-monotonic heterogeneous effects of propaganda and calculating persuasion rates. Effects are stronger among marginal voters. Exposure to RNV was highly persuasive compared to other examples of exposure to mass media (see Jun & Lee, 2023). Qualitatively, these findings also speak to the large appeal of Chávez's rhetoric (e.g., Frajman, 2014; Marcano & Barrera Tyszka, 2007).

I explore the mechanisms underpinning these results and find qualitative and quantitative evidence suggesting that Chávez's appeal can be attributed to a narrative around democracy as well as a deeply personalised style of communication. I find correlational ev-

[3] In additional to exposure to media, I also test classical determinants of electoral turnout such as closeness in the previous election (Downs, 1957; Riker & Ordeshook, 1968; Tullock, 1968), or incumbent's advantage (Cox & Katz, 1996; Erikson & Titiunik, 2015). While I find correlational evidence in support for both, I test the latter via a regression continuity design (Cattaneo, Idrobo, & Titiunik, 2019) and fail to find a statistically significant effect at the margin. By decomposing the influence of these determinants, I find that exposure to propaganda is a better predictor of turnout than these alternative variables.

[4] The standard persuasion framework in economics would not distinguish between going to vote and deciding who to vote for (Jun & Lee, 2023; Kamenica, 2019; Kamenica & Gentzkow, 2011). In this framework, both are choices that lead to utility maximisation. While individuals could certainly be persuaded to go out and vote, throughout the paper I refer to persuasion exclusively as changes in who people vote for.

idence of an increase in the share of Venezuelan citizens reporting living in a democracy as a consequence of being exposed to RNV. Applying natural language processing techniques on the transcripts of *Aló Presidente* and Chávez's speeches off the radio, I can document how his content adapted topically while centring on expressions of direct democracy. The personalisation of his style of communication is higher when he is on *Aló Presidente* than when he gives speeches off the radio. Taken together, these findings highlight a tension between liberal democracy and direct democracy e.g., removing political-counterweights and allowing recall referenda or re-electing everyone in office in the 'Mega-Election' of 2000, respectively. Finally, I consider other mechanisms that help explain Chávez's rise to power, but that fail to account for the paper's main findings. These mechanisms include interpreting exposure to Chávez on the radio as a technology that enables the dis-intermediation of political parties, and by-passing traditional means of communication.

This paper contributes to the literature in three specific ways. First, it focuses on exposure to a means of mass-communication directly used by a populist leader. Chávez defined a trend emulated by several other populist leaders,[5] and was a highly effective and charismatic communicator. Empirically, it is challenging to attribute any effects directly to Chávez. In this case, I rely on two key placebo tests that rule out results being driven by exposure to RNV, or exposure to RNV under Chávez's regime. Results only appear after *Aló Presidente* was introduced – highlighting the importance of the direct use of means of communication by a populist leader, even if I am unable to quantitatively attribute the effect solely to Chávez. This result resonates with qualitative evidence that documents how pervasive Chávez's influence was (Frajman, 2014).[6]

Chávez's rhetoric[7] allowed him to strengthen his regime, while undermining democracy in the country. Being able to map the specific rhetoric used by a populist leader to achieve this dual purpose represents a new set of findings in the literature. Third, this is the first paper to measure the impact of media under a left-wing populist leader from Latin America. The results hark back to the classic work of Dornbusch and Edwards (1991) and Edwards (2010). Chávez's rhetoric helped erode the democratic institutions of the country. The erosion of democratic institutions contributed to the economic instability and perpetuation of poverty in Venezuela. These findings link the work of Dornbusch and Edwards (1991) with Funke et al. (2023), emphasising institutional deterioration as a key mechanism for poor economic performance following the election of a populist candidate.

This work interacts with the vast literature on the impact of media (e.g., Adena et al.

[5] For instance, Nicolás Maduro, Andrés Manuel López Obrador in his *Mañaneras*, Narendra Modi in his *Maan Ki Baat*, Rafael Correa in his *Conversando con Correa*, Mauricio Funes, Mario Abdo Benítez, Evo Morales, etc.

[6] Recent work that pushes this margin includes studying exposure to rallies of a specific political leader (Assouad, 2020). Similarly, Esposito, Rotesi, Saia, and Thoenig (2023) study exposure to a specific movie.

[7] See Frajman (2014) for an overview.

(2015); Bernini (2023); Bursztyn et al. (2024); Caesmann, Caprettini, Voth, and Yanagizawa-Drott (2021); DellaVigna and Kaplan (2007); Enikolopov et al. (2011); Esposito et al. (2023); Gagliarducci, Onorato, Sobbrio, and Tabellini (2020); Gentzkow (2006); Olken (2009); Wang (2021); Yanagizawa-Drott (2014), or see DellaVigna and Gentzkow (2010) for a review of traditional means of communication, and Zhuravskaya et al. (2020) for a review of the impact of the internet), as well as the literature on populism, and their intersection (e.g., Guriev & Papaioannou, 2022). The paper also highlights the role of government propaganda as a means of persuasion (Jun & Lee, 2023; Kamenica, 2019; Kamenica & Gentzkow, 2011).[8] Finally, the paper contributes to the body of academic work that focuses on Chávez's regime in Venezuela, and its legacy (e.g., De Anda Casas, 2025b; Frajman, 2014; Knight & Tribin, 2019, 2022; Mallén & García Guadilla, 2017). Specifically, the paper complements the recent work of Knight and Tribin (2019, 2022) which focused on the impact of censorship on individuals' preferences in Venezuela post 2006 – once the autocracy had already been consolidated.

The rest of the paper is organised as follows. Section 1.2 describes the historical setting, focusing on Chávez's rise to power, the consolidation of his autocracy, and his use of means of communication. Section 1.3 describes the various data sources. Further details on data, simulations, and empirical strategies are provided in Appendix 1.B. Section 1.4, presents the paper's main results: the impact of RNV on the referendum to summon the Constitutional Assembly of 1999, and the 'Mega-Election' of 2000. Section 1.5 summarises all robustness tests and is complemented by Appendix 1.E. Section 1.6 analyses the potential mechanisms underpinning the results. Section 1.7 concludes.

1.2 Historical Context

The rise of a populist leader From the 1950s to the 1970s, Venezuela was one of the richest countries in Latin America due to its strong dependence on oil production. Political stability was guaranteed due to an agreement between three of the country's biggest political parties – the *Pacto de Puntofijo* – which alternated their positions in power. This system created a politically apathetic society (Wilpert, 2006), where party identification was eventually replaced by individual political leaderships, and an active civil society (Penfold Becerra, 2001).

The oil boom ended, leading to a decrease in real incomes and a rise in poverty from

[8] Some of the paper's additional results also speak to the role of propaganda and its potential role in influencing electoral turnout (see e.g., Barzel & Silberberg, 1973; Downs, 1957; Franklin, 2004; Górecki, 2013; Herrmann, Jong-A-Pin, & Schoonbeek, 2019; Matsusaka, 1993; Matsusaka & Palda, 1993; Plutzer, 2002; Riker & Ordeshook, 1968; Tullock, 1968)

17% to 65% (Wilpert, 2006). In 1983 Venezuela lived an exchange crisis, but social unrest peaked in 1989 after president Carlo Andrés Pérez enacted a strong set of neoliberal reforms (most notably, liberating petrol prices) in exchange for an IMF loan. The backlash against this reform led to a riot that was later called *El Caracazgo*, where hundreds of civilians were killed. This unrest was capitalised by a young Hugo Chávez and a group of like-minded soldiers (known as *La Causa R*), who were later involved in two attempted coups in 1992. Chávez made his first famous appearance on television acknowledging the failure of the coup 'for the time being' before turning himself in. This gave Hugo Chávez incredible popularity among Venezuelans, which was but a symptom of the collapse of the party system in the country.

In 1994 president Rafael Caldera gave amnesty to individuals involved in the coups of 1992. Upon leaving prison, Chávez formed a new party, *Movimiento por la Quinta República* (MVR). His campaign for the presidency officially started in 1997, when MVR received support from other left-wing parties such as *Movimiento al Socialismo* (MAS), *Patria Para Todos* (a radical wing of *La Causa R*), the Communist Party and five other small parties (Gott, 2005).

Hugo Chávez was elected president in 1998 with strong support from Venezuela's middle and lower socioeconomic strata. His electoral victory can be understood in light of the dissolution of Venezuela's party system; reflected in the rallying of small, satellite, parties around single charismatic leaders – a system described as 'unstable bipartisanship' (Molina V, 2001; Sonntag & Maingón, 2001). His election also catalysed a generalised discontent with the results of the neoliberal programme that previous governments had implemented, and as a response to rising economic and racial inequalities rooted in Venezuela's colonial past (Cannon, 2013; De Anda Casas, 2025b).

Transitioning into an autocracy. The first two years of Chávez's government (1998 - 2000) were devoted to consolidating his power. There were few changes in terms of policy, but there was a radical shift in the government's discourse (Carrasquero & Welsch, 2001). The first two years of Chávez's government were spent creating a narrative around his government's programme, represented in the Constitution of 1999, and building the institutional scaffolding for the new regime. Changing the country's constitution was a key element of Chávez's presidential campaign. In April 1999, just five months after being elected, he called for a referendum to summon the National Assembly and draft this new constitution. The referendum passed with an overwhelming majority, and the Constitution of 1999 was adopted in another referendum in November 1999.

The change in the constitution brought about some profound changes in Venezuelan political life. Some of these changes include giving the president the power to call refer-

enda, naming an Executive Vice President, dissolving the National Assembly; extending presidential terms from four to six years; replacing the bicameral National Congress with an unicameral National Assembly; rescinding recent labour laws, guaranteeing universal health care, education, pensions, and employment (Garcia-Serra, 2001). Wilpert (2006) discusses these changes in detail. A more thorough summary is provided in Appendix 1.H. Even though many of the changes implemented in the Constitution never came to full fruition, the Constitution became a political banner and a project to work towards (Wilpert, 2006)[9] – something that Chávez's opposition completely lacked. The new Constitution also called for all elected members of the government to be 're-legitimised' in what became the 'Mega-Election' of 2000 (Carrasquero & Welsch, 2001). This would be Chávez's first re-election.

Despite its ambitious reforms, the institutional changes that Venezuela underwent resulted in a high concentration of power (Brewer Carías, 2015; Garcia-Serra, 2001). The pace at which reforms were being implemented did not match the speed of the creation of all the laws needed for a smooth transition. The constitution had to be amended several times to reflect the country's reality and the government's constraints. These amendments often granted more power to the government, and violated Venezuela's supra-constitutional norms (see Brewer Carías (2015) for a detailed account). Ultimately, this created a culture of presidentialism and clientelism. The economic reforms also prevented structural transformations from happening, and increased the country's dependence on oil revenues. Support for Chávez waned among middle class Venezuelans and concentrated among the country's destitute.

There was no clear opposition to Chávez, except from members of a small elite, whose political power was ousted as a consequence of Chávez's election (Wilpert, 2006). Opposition to Chávez mainly came from the newspaper *El Mundo* – owned by a former member of MAS – and from capitalist representatives of commerce chambers who owned private media outlets (Gott, 2005). He referred to the owners of *Venevisión*, *Radio Caracas Televisión*, *Globovisión*, and *Televen* as the Four Horsemen of the Apocalypse. Apart from *El Mundo*, he faced opposition from other major newspapers published mainly in Caracas such as *El Universal* and *El Nacional*. But outside the capital, newspapers adopted a neutral tone towards his government (Gott, 2005). Tensions with private media peaked in 2002 due to the coverage that was given to the failed military coup, and culminated with his government nationalising several media outlets.

The years that followed a failed coup were characterised by Ellner (2008) as a stage

[9] Often printed in pocket-sized editions with a blue cover, the Constitution of 1999 bears a clear resemblance to Mao's little red book.

of anti-neoliberal radicalisation, which led to a change in economic model in 2005, and Chávez's posterior reelection in 2006 (also see Cannon, 2013). Wilpert (2006) describes how Chávez had not used the term 'socialism' to describe his programme until 2005, and compares the radicalisation that occurred between 2002 and 2004 to the Cuban and Nicaraguan revolutions – with the difference that by 2006 it seemed that Chávez's government had defeated all real political opposition in the country. Steadily, Venezuela's government not only increased its political control, but also its participation in the economy by taking over businesses in oil production, agriculture, finance, mining, power generation, tourism and telecommunications (Reuters Staff, 2012). After Chávez's second re-election, and with the passing of further 'enabling laws' in November 2000 which granted even more direct power to the president (Garcia-Serra, 2001), Venezuela slowly became an autocracy.

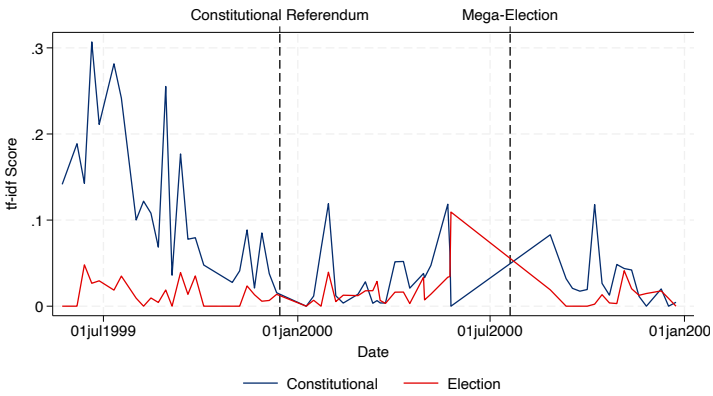
Chávez and his use of state media. State media played an instrumental part in Chávez's government since 1998. Most notably, in May 1999, just five months after his election, he started the radio programme *Aló Presidente*. Broadcasting every Sunday through *Radio Nacional de Venezuela* (RNV), the show featured Chávez speaking directly to his country's population. He talked about his government's programme, the challenges they faced and their successes, the show included interviews and call-ins from supporters.

Chávez was incredibly effective at harnessing this means of communication. Given the popularity of the programme, he could easily adapt its contents to match the interests of his government. This is captured in Figure 1.3 which shows the term frequency-inverse document frequency (tf-idf) scores for two key words: 'constitutional' and 'election'. The relative importance of the word 'constitutional' is high in the aftermath of the Constitutional Assembly Referendum of April 1999 and remains relevant until after the Constitutional Referendum took place in November 1999. The word 'election' only increases in its relative importance until just before the 'Mega-Election' of 2000.

The show was so successful that they decided to turn it into a television show. Its first broadcast was on August 27, 2000 – just a month after Chávez was re-elected. *Aló Presidente* became a model of communication that has been adopted by other populist politicians in the world, such as Narendra Modi in India, or Andrés Manuel López Obrador in Mexico. For almost a decade, the model was even adopted by his opposition in a programme called *Aló Ciudadano*. Even after Chávez's death, his successor, Nicolás Maduro has starred in three similar shows *En Contacto con Maduro*, *Los domingos con Maduro*, and most recently *Con Maduro +*.[10]

Aló Presidente was not the only content produced by RNV, but Chávez's weekly programmes did set the tone for the rest of the content produced by RNV. Being Venezuela's

[10] The long-term influence of *Aló Presidente* is studied in Frajman (2014).



Notes: The plot shows the term frequency-inverse document frequency (tf-idf) scores for the terms 'constitutional' and 'election' in the transcripts of *Aló Presidente* between its first broadcast on May 1999 and 2001, along with Chávez's speeches during these months. This measure assigns higher values to words that are relevant for a specific document (at the cross sectional level), but penalises terms that are common across documents (at the time series level). High tf-idf scores indicate frequent terms that are unique to specific documents. The series shows how the word 'constitutional' was frequently used in the aftermath of the National Assembly Referendum from April 1998 and the lead-up to the Constitutional Referendum that took place in November 1999. Afterwards, the term was much less frequently used. The series also shows the term 'election' only became relevant immediately before the 'Mega-Election' of June 2000 took place. All transcripts were scraped from todochavez.gob.ve. All analyses were carried out in Spanish. See Appendix 1.D for more detail on text analyses.

Figure 1.3: tf-idf Scores for Key Words in *Aló Presidente*

official means of communication, it constantly broadcast content that was specifically aligned with Chávez's agenda. This became official in 2002 when RNV was formally incorporated into the Venezuelan government under the leadership of Jesse Chacón, who was Chávez's Minster of Communications at the time. Under his tenure, RNV underwent a massive expansion process, embracing the adoption of FM radio as a tool of political propaganda. This is documented in Appendix 1.B.1. To this day, RNV is still the radio station with most coverage in Venezuela.

This paper takes advantage of a unique window of opportunity between Chávez's first election in 1998, and the changes that it brought to the content broadcast by RNV, including the creation of *Aló Presidente*, and the Constitutional Referendum of 1999 and the 'Mega-Election' of 2000. During this short period of time there was no response from his opposition, RNV did not expand, the same type of content was not broadcast through television, and there were no local elections. This particular set of conditions allows for a clean identification of the effect of state media on electoral outcomes.

1.3 Data Sources

Electoral outcomes. I retrieve the data from the 1998 and 2000 Venezuelan presidential elections, as well as the referendum to summon the Constitutional Assembly of 1999 from

ESDATA via the Wayback Machine.[11] I aggregate all results at the level of the civil parish – Venezuela's third administrative level. Overall, the electoral outcomes data comprises a panel for 1,098 civil parishes which includes the share of votes that Chávez and his opposition received in 1998 and 2000, voters' choices to the two questions asked in the 1999 Referendum,[12] turnout, the share of invalid votes, and the number of registered voters for each parish-year.

Radio simulations. I retrieve the list of 10 radio towers comprising RNV's network in 1998 from RNV's website. Specifically, I rely on two sources reproduced and translated in Appendix 1.B.1.[13] Following Blouin and Mukand (2019), I approximate RNV's coverage, first by retrieving the geographic location of these 10 towers from FMScan.org. I uploaded the locations of these towers to the desktop version of the open source application RadioMobile, which allows users to simulate the strength and reception of radio signal by using publicly available topographic data.[14] The simulation generates about 3 million data points, which cover the entire country. Each data point represents approximately 0.83 km^2. Panel (a) of Figure 1.4 shows the political division of Venezuela and the location of the 10 RNV towers. Panel (b) of Figure 1.4 shows the raw output of the simulation.

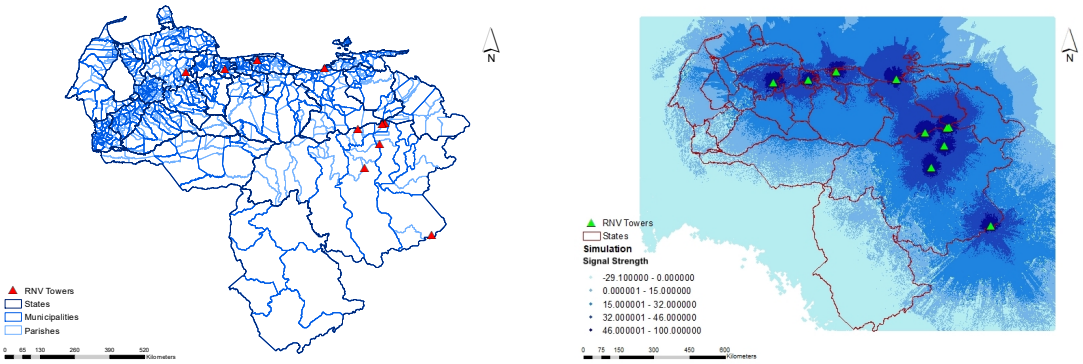
The strength of a radio signal is measured in decibels relative to 1 microvolt per meter ($dB\mu V/m$). I adopt a threshold of 32 $dB\mu V/m$ as the minimum strength for a data point to

[11] I want to thank Dorothy Kronick for helping me find ESDATA's website. The Wayback Machine is an internet service launched in 2001 that stores copies of websites. As a user, one can browse past versions of websites; access websites that are no longer available; or save copies of current websites for future access. As of 2020, the Wayback Machine had stored over 70 petabytes of data.

[12] The specific wording of the questions is: Q1: 'Do you convene a National Constituent Assembly with the purpose of transforming the State and creating a new legal system that allows the functioning of a Social and Participatory Democracy?', and Q2: 'Do you agree with the bases proposed by the National Executive for the convocation of the National Constituent Assembly, examined and modified by the National Electoral Council in session dated March 24, 1999 and published in its full text, in the Official Gazette of the Republic of Venezuela No. 36.669 dated March, 25 of 1999?'

[13] One of these sources is a news article from RNV dated on April 2004 reporting the inauguration of 21 FM radio frequencies following RNV's aggressive expansion process. The article states that RNV broadcast in 30 stations at the time: the 21 newly inaugurated FM frequencies, and 9 AM frequencies. The second source is an official list of RNV frequencies from September 2004. I matched the frequencies inaugurated in April to the list from September and found that there were 10 AM frequencies remaining plus a handful of frequencies devoted to music. I take this list of 10 AM frequencies as the list of frequencies, and I assume that they were active in 1998. This is a strong assumption. One concern would be whether I missed any frequencies. Anecdotal evidence supports that FM radio was only used for political talk shows, such as *Aló Presidente*, well into the 21st century – way after Chávez's presidency started. This was discussed by Polo Troconis, a Venezuelan radio host, in an interview documented by Velásquez Malavé (2009). The other concern is whether I am including frequencies that were not active in 1998. I found no evidence of an expansion process of RNV prior to the start of Chávez's presidency and before the well-documented expansion that began in 2004. To address this concern I perform leave-one-out simulations and test the sensitivity of the results to alternative definitions of the RNV network.

[14] RadioMobile runs on the Irregular Terrain Model – a radio propagation model designed to represent frequencies in the range of 40 MHz to 100 GHz. This does not include the frequencies used in AM radio broadcasting. Forcing the model to predict coverage of these frequencies can introduce inaccuracies, which I assume as measurement error (see Bernini, 2023, for a similar approach). In section 1.5, I report tests that allay the concern that any results are driven by this measurement error. Details of the simulation process for replicability are available in Appendix 1.B.2 including a list of the assumptions I adopted.



(a) Venezuela's States, Municipalities and Parishes, with RNV's Tower Locations

(b) RNV Signal Strength Simulation

Notes: Panel (a): Shapefile of Venezuela with administrative divisions and location of RNV towers. Panel (b): Simulations carried out using RadioMobile, imported into ArcMap.

Figure 1.4: Political Divisions of Venezuela, RNV Tower Placement, and Signal Strength Simulation

be considered as receiving a strong enough signal to be picked by a conventional radio. To define this threshold I followed EBU (2013).[15] Full details on the calculation of this threshold including parameter assumptions are included in Appendix 1.B.3.[16]

Following Guriev, Melnikov, and Zhuravskaya (2021), I define the main independent variables by combining the radio signal simulations with population density data. I use the Global Population Density Grid Time Series for 1990 from NASA's Socioeconomic Data and Applications Center (SEDAC) and the Center for International Earth Science Information Network (CIESIN) (2011; 2017).[17] First, I intersect the Population Density Grid ($\mathscr{I}$) with the shapefile that includes Venezuela's parishes ($\mathscr{P}$). I then take each pixel (i) from the Population Density Grid located in parish p with value $\text{pop}_{i,p}$ and spatial join[18] it to the nearest data point (j) from the radio simulation ($\mathscr{J}$) with strength value signal_j. This creates

[15] For instance, Blouin and Mukand (2019) relied on the methodology from Hoeg and Lauterbach (2003) (see Table 7.3 in the latter), who follow Hunt et al. (1995), CEPT (2007) and EBU (2003) – later superseded by EBU (2013).

[16] It is worth emphasising that the simulations from RadioMobile only account for topographic variation. That is, they omit the potential interference from buildings, trees, billboards, vehicles, walls and other radio signals, not to mention other propagation effects such as signal attenuation, shadowing, multipath effects, scattering and diffraction (Figueiras & Frattasi, 2010). Adjustments need to be made to account for these factors. Further details are on the simulations are in Appendix 1.B.3. I report robustness of the main results to changes on the construction of the network,a s well as signal strength assumptions in Section 1.5.

[17] The resolution of this source is approximately 1 km^2 per pixel.

[18] All of these operations were carried out in ArcGIS Pro. Using the GIS metalanguage, this process was equivalent to intersecting the (point) features that comprise the Population Grid with the (line) features extracted from the parishes shapefile. I then spatial joined the population features to the (point) features corresponding to the raster generated in the simulation. The values pop and signal are two fields assigned to the population feature.

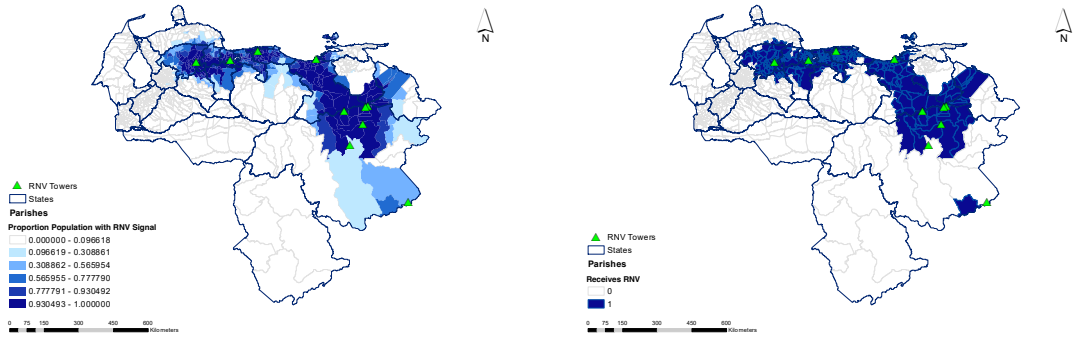
a tuple $(\text{pop}, \text{signal})_{i,p}$ for each point $i \in \mathcal{I} \cap \mathscr{P}$. For each parish $p \in \mathscr{P}$, I then calculate:

$$\text{PropRNV}_p := \frac{\sum_{i=1} \text{pop}_{i,p} \mathbb{1}\{\text{signal}_{i,p} \geq 32\}}{\sum_{i=1} \text{pop}_{i,p}}.$$

I also define a coverage indicator as:

$$\text{RNV}_p := \mathbb{1}\{\text{PropRNV}_p \geq .50\}.$$

This variable takes a value of 1 when at least half the population living in a parish lives in an area where radio signal is strong enough to be picked by a conventional radio, and 0 otherwise. I explore robustness of the results to alternative definitions of these variables in Section 1.5 and Appendix 1.E. I plot these two variables in Figure 1.5.[19]



(a) Proportion of Population Exposed to RNV per Parish

(b) Parishes Exposed to RNV

Notes: The maps on the figure illustrate the two main definitions of the independent variable used in the paper. Panel (a) shows the continuous version. Panel (b) shows the discretized version.

Figure 1.5: Independent Variables

Demographics. I rely on the results from the 1990 census published in a series of 24 volumes – 23 for each state in the country (OCEI, 1993b) and a general volume aggregating data for the entirety of Venezuela (OCEI, 1993c), which I digitise and include in the analyses.[20] I rely on this digitised information as most data is reported at the level of the civil parish – giving me more variation than currently existing demographic sources. The variables I obtained include the gender composition of each parish, the workforce

[19] I show the proportion of parishes coded as receiving the signal when I make changes to either threshold in Figure 1.A.1. Changes along the field strength axis have a greater gradient than along the population axis.

[20] The entire collection is available at the Bodleian Library at the University of Oxford, and with support from the staff of the Social Sciences Library I was able to scan the tables that showcased the relevant results for this project. The scans were then transformed into documents using Kofax PowerPDF and the data was extracted directly from these documents. A couple of alternative ways of extracting data from scanned documents include using ABBYY Fine Reader PDF or LayOut Parser (Shen et al., 2021), an open source deep learning tool. A recently created alternative is the OcrPy library (maxentlabs, 2022).

composition of each parish,[21] literacy,[22] information about houses at the level of the parish,[23] and physical house characteristics at the level of the municipality.[24] Table 1.C.1 in Appendix 1.C shows summary statistics of these variables.

Other key data sources. I use the Intercalibrated average DMSP-OLS night-time luminosity series for 1998 (Elvidge et al., 1999; Hsu, Baugh, Ghosh, Zhizhin, & Elvidge, 2015) to account for gross urban/rural differences and local levels of economic development (see Henderson, Storeygard, & Weil, 2011, 2012).[25] Figure 1.A.3 shows this data along with the location of the 10 radio towers included in the simulations. Transcripts of *Aló Presidente* and Chávez's speeches were scraped from todochavez.gob.ve. Finally, I include all available rounds of Latinobarometer before Chávez became president and until 2004. The years I include in this cross-sectional panel are 1996, 1997, and from 2000 to 2004. There were no waves in 1998 and 1999.[26]

1.4 Main Results

1.4.1 Effect of Exposure to RNV

Throughout the paper, the econometric models take the form of simple least squares specifications:

$$y_{p,2000} = \beta \text{PropRNV}_p + X_p\delta + \epsilon_p, \tag{1.1}$$

where $y_{p,t}$ can be the share of votes for each of the options of the referendum in 1999, Chávez's vote share in 2000, or turnout in 1999 or 2000, for parish p. Whenever appropriate, I include the measure of each outcome in 1998, $y_{p,1998}$, as a control. These specifications

[21] This includes the number of men and women in the legal working age, and the number of employed and unemployed men and women.

[22] Including the number of individuals over the age of 5, and the number of literate and illiterate men and women

[23] This includes the number of households, number of occupied and unoccupied houses, occupants per house, number of houses under construction and number of empty houses

[24] This includes the number of houses by type of outer walls, type of roof and type of flooring materials

[25] It is evident that the location of the towers is not entirely random. They were built near urban areas, specially in the north of the country. As explained in section 1.2, support for Hugo Chávez in urban areas gradually waned after 1998, and most strongly after 2000. Thus, in this specific setting, including this variable in the regression models is fundamental as night-time luminosity is correlated with the location of urban areas. Night-time luminosity will therefore improve the estimation of the effect of being exposed to the signal from an RNV tower. It will account for gross urban/rural differences, and will prevent us from misattributing them to the main independent variables.

[26] I choose 2004 as the cut-off point given RNV's aforementioned aggressive expansion process. 2004 was also the year when the Tascón-Maisanta database was released (see Hsieh, Miguel, Ortega, & Rodriguez, 2011), potentially altering citizens' incentives for political participation. Further, the post-2004 period has been studied by Knight and Tribin (2019, 2022).

are akin to ANCOVAs.[27]

X_p is a vector of controls that includes the average luminosity in parish p in 1998, the geodesic distance from the centre of each parish to the nearest radio tower (as in Bernini, 2023; Blouin & Mukand, 2019; Olken, 2009, etc.), the change in the number of registered voters in each parish relative to those in 1998,[28] the population in parish p as a share of the country's total population in 1990, an indicator for possible corruption that takes the value of 1 when the number of null votes in the parish exceeds the long-term national average of 5% and the majority of tables in the parish used an electronic voting system in 1998 (following Kornblith, 2001),[29] as well as demographics such as the parish's literacy rate, proportion of men, unemployment rate, an Anderson (2008) index aggregating precarious living conditions;[30] and state indicators.

The main coefficient of interest is β. When the lag of the outcome is included, the interpretation of β corresponds to the change on Chávez's share between parishes receiving RNV's signal and not as the proportion of the population receiving RNV's signal increases from 0 to 100%. The interpretation of this coefficient is akin to an Intent to Treat (ITT).

Of the covariates I include in the regressions, luminosity and the distance to the nearest tower are needed for identification of the effect associated with exposure to RNV: their purpose is accounting for the endogeneity stemming from the non-random placement of radio towers, and has been a feature of this literature since Olken (2009) (see also Bernini, 2023; Blouin & Mukand, 2019; Enikolopov et al., 2011; Gagliarducci et al., 2020, etc.).[31] Other covariates that help with identification are state indicators, the poverty index, and the literacy rate.[32] The rest of the covariates are there to increase the precision of the estimates.

[27] I use ANCOVA specifications rather than differences in differences due to the gain in power from using the former. McKenzie (2012) pointed out that the ratio of the differences-in-differences variance relative to the ANCOVA variance is given by $2/(1+\rho)$, where ρ is the autocorrelation of the outcome of interest. In this setting, the autocorrelation of Chávez's share is $\rho = 0.65$, which would yield a relative power loss of 21% from using differencess-in-differences. See also Özler (2015) for a blog post on this decision.

[28] Winsorised at 1 and 99% to reduce measurement error.

[29] The 1998 election in Venezuela introduced the use of electronic voting systems wherein voters marked a sheet of paper with their vote. The marked sheets were later fed into a scanner that read and counted the votes. According to Kornblith (2001), having a percentage of null votes higher than the long-term country average of 5% suggests the possibility that the sheets were pre-marked, as the electronic system would invalidate a vote when more than one candidate was selected.

[30] More details about the construction of this index in section 1.4.3

[31] Papers in the literature have also overcome this issue using instrumental variables to measure exposure to AM radio by relying on the soil's ion composition (Adena et al., 2015), or relying on ionospheric propagation (Gagliarducci et al., 2020). Given the open source tool I use to generate the radio simulations, the measure of exposure abstracts from these channels of propagation and relies exclusively on the natural ruggedness of the terrain as in Bernini (2023). Instead, I address the potential endogeneity issue by relying on the correction proposed by Borusyak and Hull (2023) as discussed in section 1.5. I instrument PropRNV$_p$ with the average proportion of individuals that would have received RNV in each parish under a series of counterfactual topographies.

[32] See Table for 1.A.1 balance on other demographic characteristics and political variables on the residualised measure of exposure.

All of the models are calculated using robust standard errors and without clustering (see Abadie, Athey, Imbens, & Wooldridge, 2022). I consider alternative ways of clustering, and calculating Conley (1999) standard errors in the robustness section. For hypothesis testing, I report sharpened p-values (q-values) to control for the False Discovery Rate (FDR), following Benjamini, Krieger, and Yekutieli (2006). I also consider adjusting for the Family-Wise Error Rate (FWER) using the stricter Westfall and Young (1993) algorithm as a robustness test.

Finally, there are some states for which no parishes are receiving RNV's signal. This means state indicators could be collinear with the measures of coverage, which would lower the precision of the estimates. To address this concern I include an additional column where I penalise the state indicators via Lasso.

Table 1.1: RNV's Signal on the 'Mega-Election' of 2000

	Chávez's Share 2000				
	(1)	(2)	(3)	(4)	(5)
RNV Coverage	0.051***	0.034***	0.048***	0.032***	0.034***
	(0.016)	(0.012)	(0.015)	(0.012)	(0.011)
	{0.006}	{0.018}	{0.006}	{0.018}	{0.006}
Chávez's Share 1998		0.771***		0.755***	0.732***
		(0.032)		(0.032)	(0.032)
		{0.000}		{0.000}	{0.000}
Turnout Rate 1998			-0.308***	-0.216***	-0.212***
			(0.055)	(0.043)	(0.042)
			{0.000}	{0.000}	{0.000}
Observations	1096	1092	1096	1092	1092
Control Mean	0.583	0.583	0.583	0.583	0.583
LASSO	No	No	No	No	Yes
R^2	0.365	0.683	0.390	0.695	0.674

Notes: Effect of RNV's Signal on electoral outcomes from the Venezuelan 'Mega-Election' of 2000. RNV's Signal (Proportion) is the proportion of each parish's population covered with a signal with a field strength of at least 32 dBμV/m. All models include the distance to the nearest RNV tower, the parish's population as a share of the country's population, the average luminosity in the parish in 1998, the change in registered voters at each parish between 1998 and 2000, demographics, a dummy variable that takes a value of 1 when the number of null votes in a parish exceeds 5% and most of the polling stations in the parish were electronic, and state indicators. Robust standard errors in parentheses. Sharpened p-values in curly brackets. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

Table 1.1 shows the effect of exposure to RNV on the Venezuelan 'Mega-Election' of 2000. I observe a strong and consistent increase in the share of votes that Chávez received. The results are robust to the inclusion of the turnout rate in the previous election as a proxy for participation in 2000.

I perform two key placebo tests. Table 1.2 shows that RNV had a null impact on the share of votes that Chávez received in 1998 when he first ran for president. The lack of

Table 1.2: Placebo Test 1: The Presidential Election of 1998

	Chávez Share 1998		
	(1)	(2)	(3)
RNV Coverage	0.021	0.019	0.017
	(0.015)	(0.015)	(0.012)
	{0.357}	{0.357}	{0.357}
Turnout Rate (1998)		-0.160***	-0.168***
		(0.050)	(0.049)
		{0.002}	{0.002}
Observations	1092	1092	1092
Control Mean	0.498	0.498	0.599
LASSO	No	No	Yes
R^2	0.416	0.423	0.420

Notes: Effect of RNV's Signal on the Presidential Election of 1998. RNV's Signal (Proportion) is the proportion of each parish's population covered with a signal with a field strength of at least 32 dBμV/m. All models include the distance to the nearest RNV tower, the parish's population as a share of the country's population, the average luminosity in the parish in 1998, the registered voters at each parish between 1998, demographics, a dummy variable that takes a value of 1 when the number of null votes in a parish exceeds 5% and most of the polling stations in the parish were electronic, and state indicators. Robust standard errors in parentheses. Sharpened p-values in curly brackets. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

Table 1.3: Placebo Test 2: The First Wave of the Constitutional Referendum of 1999

	Share 'Yes' Q1				
	(1)	(2)	(3)	(4)	(5)
RNV Coverage	0.005	0.005	-0.002	-0.002	0.003
	(0.008)	(0.008)	(0.007)	(0.007)	(0.005)
	{1.000}	{1.000}	{1.000}	{1.000}	{1.000}
Turnout Rate (1998)		-0.023		0.027	0.027
		(0.026)		(0.022)	(0.022)
		{0.692}		{0.622}	{0.622}
Chávez's Vote Share (1998)			0.311***	0.313***	0.292***
			(0.020)	(0.020)	(0.017)
			{0.000}	{0.000}	{0.000}
Observations	1077	1077	1075	1075	1075
Control Mean	0.868	0.868	0.868	0.868	0.868
LASSO	No	No	No	No	Yes
R^2	0.234	0.235	0.445	0.446	0.436

Notes: Effect of RNV's Signal on the first wave of the Constitutional Referendum of 1999. The outcome is the share that votes 'Yes'. RNV's Signal (Proportion) is the proportion of each parish's population covered with a signal with a field strength of at least 32 dBμV/m. All models include the distance to the nearest RNV tower, the parish's population as a share of the country's population, the average luminosity in the parish in 1998, the change in registered voters at each parish between 1998 and 1999, demographics, a dummy variable that takes a value of 1 when the number of null votes in a parish exceeds 5% and most of the polling stations in the parish were electronic, and state indicators. Robust standard errors in parentheses. Sharpened p-values in curly brackets. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

an effect on this outcome suggests that results are not driven by exposure to RNV *per se.* The second placebo test is looking for an effect on the first wave of the Constitutional Referendum. Table 1.3 confirms that there is no effect of RNV on this outcome either. This nuances the findings from Table 1.1 further, as these results suggest the findings are not driven by media under Chávez's regime either. Altogether, these results emphasize the influence of *Aló Presidente* even if it is not possible to specifically attribute the findings to exposure to *Aló Presidente* on its own.

1.4.2 Are These Results an Artefact of Political Mobilisation?

I find that the effect of RNV on Chávez's share is driven by locations where Chávez originally won in 1998 (Table 1.A.4), i.e., Chávez's share deepened. I also find that RNV also decreased turnout (Table 1.A.2).[33] I probe whether the increase in Chávez's share is entirely driven by the reduction in electoral turnout and do not find evidence in support for this hypothesis.

Results on Chávez's share (Table 1.1) are robust to controlling for a proxy for turnout in 2000. Further, the reduction in turnout occurred across the sample – not only in locations where Chávez won in 1998 (Table 1.A.5). In fact, in locations where turnout decreased, the magnitude of RNV's effect on Chávez's share was lower (Table 1.A.6). Turnout in 2000 would be a 'bad control' to include in a regression on Chávez's share in 2000. But I can push the analysis further by considering turnout in 2000 as a mediator for Chávez's share. First, I am able to quantify the change on the magnitude of RNV's effect on Chávez's share by including turnout in a regression via a Gelbach (2016) decomposition. I push the analysis even further at the cost of stronger assumptions via a mediation analysis in Appendix 1.B.4. Regardless of the methodology, I obtain the same results: the increase in Chávez's share cannot be fully accounted for by changes in electoral turnout.

1.4.3 Is this Effect Due to Persuasion?

If the effect on Chávez's share is not fully accounted for by changes in turnout, persuasion could have played a role. I probe this further by exploring heterogeneities among those covered by RNV's signal, and by calculating persuasion rates as in DellaVigna and Kaplan (2007); Enikolopov et al. (2011); Jun and Lee (2023).

Heterogeneous response to exposure to RNV. To explore heterogeneities on demographics I use the data digitised from the 1990 census. Summary statistics of all these

[33] Appendix 1.F is devoted to exploring this effect relative to other determinants of turnout such as closeness and incumbency. The effect of exposure to RNV on turnout is robust to considering alternative determinants to turnout, and is a better predictor of turnout than closeness or incumbency.

variables are presented in Table 1.C.1 in Appendix 1.C. To complement this data and in order to increase statistical power, I aggregate the variables associated with precarious building materials (floors, walls and roofs) into indices following the methodology from Anderson (2008). Finally, I create another index, which I call 'poverty index' that aggregates the components of the 'precarious' variables defined above, but not the indices I defined *ex post*. I provide further details on the construction of these indices, and the heterogeneities analysis in Appendix 1.C.

To explore heterogeneities among those receiving RNV, I use a causal forest framework. This method has been highlighted by Athey and Imbens (2017), and recently applied in experimental settings by Davis and Heller (2017) and Riley (2022). I apply a causal forest algorithm[34] to estimate the effects associated with exposure to RNV using the coverage indicator.[35,36] I then compare the demographic characteristics of parishes at the bottom and top quartiles of the distribution of effects for those that were exposed to RNV. Table 1.4 summarises the results I find for Chavez's share. I present the full results for Chavez's share and the turnout rate in 2000 in Tables 1.C.2 and 1.C.3, respectively.

Table 1.4: Random Forest Analysis: Key demographic heterogeneities

	1Q		4Q		4Q - 1Q	
Variable	Mean	SD	Mean	SD	Difference	p-value
Δ Chávez's Share (Pred.)	-0.001	(0.005)	0.037	(0.006)	0.038***	(0.000)
Illiteracy Rate	0.226	(0.115)	0.143	(0.050)	-0.084***	(0.000)
Precarious Wall Index	-0.071	(0.405)	-0.102	(0.321)	-0.031	(0.566)
Precarious Roof Index	-0.244	(0.517)	-0.288	(0.456)	-0.044	(0.546)
Precarious Floor Index	-0.087	(0.506)	-0.275	(0.228)	-0.188***	(0.001)
Poverty Index	-0.184	(0.419)	-0.132	(0.382)	0.052	(0.382)
Total Houses	8,030.322	(13,262.197)	12,711.036	(15,238.875)	4,680.714**	(0.028)
Occupied Houses (Prop.)	0.769	(0.117)	0.855	(0.097)	0.086***	(0.000)
Parish Area (km^2)	567.726	(1,258.383)	198.561	(328.693)	-369.165***	(0.007)
Men (Prop.)	0.521	(0.035)	0.502	(0.015)	-0.019***	(0.000)
Observations	91		91		182	

Notes: Random forest analysis applying a causal forest algorithm to estimate the impact of exposure to RNV on Chávez's share in 2000 using the base specification described in section 1.4.1 controlling for turnout. Implemented in R (R Core Team, 2020) with the grf package. Quartiles are defined on the impact of RNV's reception on Chávez's share in 2000 for parishes that were exposed to RNV. Index variables are standardised indices built following the methodology in Anderson (2008). Further details on variable creation and the implementation of the algorithm are provided in Appendix 1.C. Tables reporting all demographic variables are Tables 1.C.2 and 1.C.3 for Chávez's share and turnout in 2000 respectively.

Overall, the results suggest the increase in Chávez's vote share associated with RNV's reception was larger in parishes with a higher literacy rate. Additionally, I also observe higher increases in parishes with better living conditions. The strongest difference comes

[34] Despite the name, these results can only be interpreted as causal if we are convinced that exposure to RNV is conditionally exogenous. I want to thank James Fenske for raising this concern.

[35] I carry this out in R R Core Team (2020) using the grf package.

[36] Further details on this methodology are provided in Appendix 1.C.

from the index that aggregates precarious floors; the indices for precarious walls and roofs point in the same direction but are not statistically significant. Additionally, I observe higher increases in parishes that had a smaller proportion of men.

I also find that parishes where the effect was larger are more densely populated, as represented by higher number of houses and smaller areas. This is also reflected in the statistically significant difference I observe between the number of occupied houses. Mechanically, I also find that parishes at the bottom quartile of the effect distribution have a higher proportion of unoccupied, occasionally occupied houses, and houses under construction. Given that results are robust to including a parish's area in the regressions, I interpret these findings as suggesting a network effect.

Estimating persuasion rates. Estimating persuasion rates has been a feature of the literature on media exposure since DellaVigna and Kaplan (2007). Following the discussion and notation in Jun and Lee (2023), this was originally carried out by computing the (unfeasible) estimand $f = \frac{y_{\mathbb{T}} - y_{\mathbb{C}}}{e_{\mathbb{T}} - e_{\mathbb{C}}} \cdot \frac{1}{1-y_0}$, where $\mathbb{T}$ and $\mathbb{C}$ represent an instrumental treatment assignment, y_j denotes the share of a group taking a specific action, and e_j denotes the share of a group exposed to a message, where $j \in \{\mathbb{T}, \mathbb{C}\}$; and y_0 refers to the unobserved share of a group who would have taken the specific action without being exposed to the transmitted message anyway. In practice, this is calculated using $y_{\mathbb{C}}$ instead of y_0, yielding the (feasible) estimand $\tilde{f}$.

Jun and Lee (2023) show how this estimand might not identify the average persuasion rate (APR) nor the local persuasion rate (LPR), and provide the theory to identify them. In this setting I do not observe individual decisions ($y \in \{0, 1\}$), exposure ($t \in \{0, 1\}$), nor treatment assignment ($z \in \{0, 1\}$), but I can estimate $Pr(y = 1|z)$ and $Pr(t = 1|z)$ with the available data, which allows me to provide bounds on these estimands.

I define z as an indicator when the proportion of people receiving RNV in a given parish is above 50% under placebo topographies as detailed in section 1.5, t as the exposure indicator as defined in section 1.3, and y as voting for Chávez in 2000, or participating in the 2000 elections. Table 1.5 provides the bounds for these estimands, as well as the feasible estimand from DellaVigna and Kaplan (2007) for both actions. Results are in line with the discussion from above: I observe a null effect on political mobilisation, given that zero is included in the APR's and LPR's bounds.

1.4.4 Summary of Main Results

I find that exposure to RNV had a strong impact increasing Chávez's vote share in the 'Mega-Election' of 2000. I see no effect of exposure to RNV on the presidential election of

Table 1.5: Estimating Persuasion Rates

	$\tilde{f}$	APR	LPR
Vote for Chávez in 2000	0.1743	[0.1288, 0.6636]	[0.1288, 1]
Voted in elections in 2000	-0.0771	[-0.0569, 0.4838]	[-0.0323, 1]

Notes: I calculate persuasion rates using the methodology from DellaVigna and Kaplan (2007) and Jun and Lee (2023). I define a treatment indicator for when more than half the population of a civil parish is covered by RNV's signal. As an instrument to predict this indicator, I use a new indicator for when more than half of a parish's population is coded as receiving RNV's signal under placebo topographies as detailed in Section 1.E. As a benchmark, consider the following rates from the literature: Enikolopov et al. (2011): 7%; Chiang and Knight (2011): 2-6%; Gerber, Karlan, and Bergan (2009): 6-20%; Gentzkow (2006): 4%; Chen and Yang (2019): >50%. Naturally, one should be cautious about extrapolating results from one context onto another.

1998, and the Constitutional Referendum of 1999. These results serve as placebo tests to highlight the influence of *Aló Presidente*, as results cannot be attributed to exposure to radio in general, or exposure to radio under Chávez's government. The most important change in RNV's content during this period of time was the introduction of *Aló Presidente*.

I have also tested whether the increase in Chávez's share can be attributable to changes in electoral turnout. The evidence suggests the increase in Chávez's share cannot be fully accounted for by changes in turnout. This motivated whether persuasion played a role. An analysis of heterogeneities that allows for non-monotonicities suggests that the effect was strongest in mostly literate, wealthier, smaller, urban, and denser civil parishes. Estimating persuasion rates yields a comparatively high value for effects on Chávez's share, and null values for effects on turnout. In section 1.5 I test the robustness of the main results. In section 1.6 I explore potential mechanisms underlying these findings.

1.5 Summary of Robustness Checks

I probe the robustness of the results by performing a series of tests that I summarise in this section. Further details on all robustness checks are in Appendix 1.E.

Inference. To test for robustness in terms of inference, I estimate the models with Conley (1999) standard errors.[37] I set a range of distance cut-offs within which errors are allowed to correlate in 10km increments from 20km to 300km. Results are presented in Figure 1.E.1 in Appendix 1.E.1. Panel (a) corresponds to the model without state indicators, and Panel (b) corresponds to the model that includes state indicators. Results from the former are fully robust, and results from the latter become weaker in the range of 80-100 km remaining significant at 10%. Next, I explore adjusting for family-wise error rates applying the Westfall-Young (1993) algorithm as an alternative to calculating sharpened p-values

[37] I implement this in Stata via the ols_spatial_HAC command (Hsiang, 2010).

(Benjamini et al., 2006). All results are robust to this much stricter algorithm. I include these p-values in Table 1.A.3. Further details on this methodology are provided in Appendix 1.E.2.

Percentage of population covered. One of the two independent variables I use is defined as an indicator for parishes where more than half the population is covered by a signal strong enough to be picked by a conventional radio. I test for robustness to different percentages of the population being covered while keeping the field strength threshold fixed. The main results are unchanged to these alternative specifications. Results are presented in Panels (a) and (b) from Figure 1.E.2 and discussed in detail in Appendix 1.E.3.

Changes in signal strength threshold. Similarly, I also test for robustness to changes along the signal strength threshold. Changing this threshold should be interpreted cautiously as increasing this number could translate as adopting unnecessarily conservative assumptions,[38] or as measuring exposure to a different signal altogether. Additionally, increasing the signal strength threshold reduces the number of parishes coded as receiving the signal, which can raise concerns about statistical power. I provide further details on these checks in Appendix 1.E.4 and Figure 1.E.3, along with power calculations. Overall, results are robust to slightly decreasing the field strength threshold, which is in line with how conservative the assumptions used to calculate the threshold were in the first place. I find no statistically significant effects associated to indicators defined over higher field strength thresholds. In these cases, even if some point estimates are positive, the Minimum Detectable Effect increases substantially as the proportion of parishes coded as receiving the signal in the sample decreases. When I calculate the effect sizes for the latter point estimates, I find that their 95% confidence intervals include zero and the MDE.

Percentage of covered area, rather than population. A commonly used approach in the literature is defining coverage of mass media over the proportion of a geographic extent (e.g., Blouin & Mukand, 2019) rather than over population living in a certain place (e.g., Guriev et al., 2021; Mensah, Tafere, & Abay, 2021). The former approach will give equal weight to each squared kilometre in a sample, while the latter will weight them according to the population living in them. Results are robust to defining exposure over the coverage of a parish's area, rather than its population as presented in Table 1.E.1. Further comments are included in Appendix 1.E.5.[39]

Changes in tower network composition. I also test the sensitivity of the results to drop-

[38] See for example the discussion around building penetration loss in Appendix 1.B.3.

[39] Both approaches run the risk of introducing a bias if we do not control for the population or size of the parish, respectively. Urban parishes are more densely populated and smaller in size, so it is easier for them to pass either threshold. The results of both approaches remain the same even if I control for population or the size of the parish, respectively. These results are not shown. I want to thank Verena Wiedemann for raising this concern.

ping one of the towers that comprise RNV's network at a time. I find that dropping the towers located in Ciudad Bolívar, Ciudad Guyana, Puerto Ordáz, Guri, Paragua and Santa Elena de Ushuaia does not alter the results in any way (see Figure 1.E.5). Dropping the towers located in Barcelona, Barquisimieto, Caracas and Valencia does change the coefficients I obtain. For the case of Barcelona I obtain a coefficient with a higher magnitude. For the other three towers, results lose statistical significance depending on the specification though this is mainly due to a reduction in statistical power, as point estimates remain fairly consistent. Further details are provided in Appendix 1.E.6.

Exposure to a placebo network. I test whether the results are driven by an artefact of the simulation process itself by creating a placebo network. I follow Blouin and Mukand (2019) and Björkegren (2019) and 'place' 10 fictitious towers on the 10 tallest mountains in Venezuela as these are the most advantageous positions for radio signal transmission. I then simulate the strength of the network adopting the same assumptions I adopted for RNV's network. This placebo network is shown in Figure 1.A.4. I find no meaningful relationships between exposure to this placebo network and the electoral outcomes, which suggests that the main results are not being spuriously generated by the simulation process itself. This is reassuring as some of the measurement error introduced by the stochastic component of the simulations I perform could have been driving the results. Since I kept all assumptions constant while generating this placebo network, I am confident the simulations approximate the coverage of RNV accurately. I devote Appendix 1.E.7 to discuss this further. Table 1.E.2 provides information on the 10 tallest mountains in Venezuela, and Table 1.E.3 presents the results.

Subsampling. Ideally, we would like to compare the effect of receiving RNV within the same state or municipality. Thus, I restrict the sample conditioning on the distance to the nearest RNV tower. I present the coefficients of proportion of people covered by RNV for regressions restricting the sample from 80km to 270km, which correspond to the 30th and 80th percentiles of the distribution of distances, in Figure 1.E.6 Appendix 1.E.8. I also create 'doughnuts' around the radio towers. I drop parishes using distance ranges (from 10km to 65km) (Figure 1.E.7), or distance percentiles (from 1p. to 10p.) (Figure 1.E.8). Point estimates remain consistent and statistically significant at least at 10% throughout the distribution of distances.

Sharing a border with Colombia. A warranted concern would be that results are driven by the population that lives in states that share a border with Colombia. Possible explanations range from being in contact with people and ideas from Colombia itself; being far from the main urban centres located in the northern part of Venezuela; or being exposed to counter-propaganda from Colombian means of communication. To address this concern,

Table 1.E.4 presents the main results of the paper including an indicator for being in a state that shares a border with Colombia. Results are unchanged.

Control for non-random response to random shocks. Exposure to radio (and means of communication in general) can be understood as a composite variable of the type described by Borusyak and Hull (2021), where exposure is a data process comprised of exogenous and endogeneous data processes. The former being topographic ruggedness, while the latter referring to pre-defined characteristics of each parish in Venezuela (e.g., population, distance to radio towers, location, etc.). The pre-defined component could therefore, influence the response observed for each unit of analysis to the specific realisation of the exogenous process that we observe, biasing it upwards akin to having an omitted variable bias. Borusyak and Hull's (2021) methodology suggests this potential OVB can be corrected by instrumenting for exposure with a variable that captures the average exposure in the parish given a series of counterfactual shocks. I experiment with three different methodologies for generating these shocks. First, following advice set in Borusyak and Hull (2021), I permute the observed realisations of topographic ruggedness across the units of analysis. Second, I approximate the empirical distribution of topographic ruggedness in Venezuela via well-known theoretical distributions. I then allocate random draws from these distributions to each parish and use the average exposure under these counterfactual realisations as an instrument. Finally, I generate realistic counterfactual topographies with Perlin (1985) noise[40] – a type of quasi-randomly generated noise commonly used to create procedurally generated textures in digital media – and use these counterfactual topographies to calculate more candidate instruments. Results from these methods are presented in Tables 1.E.5-1.E.9. Point estimates remain significant albeit weakened. Further details are provided in Appendix 1.E.10.

1.6 Exploring Mechanisms

I explore the mechanisms underpinning the main results. Two compelling explanations are that Chávez created a narrative around democracy that is attractive to the Venezuelan electorate. I find both qualitative and quantitative evidence in support for this hypothesis. Further, Chávez's style of communicating in *Aló Presidente* is highly personalised – a unique and highly appealing characteristic of his rhetoric (see e.g., Frajman, 2014; Marcano & Barrera Tyszka, 2007). Other mechanisms tested that fail to explain these results but that contribute nonetheless to understand Chávez's original rise to power are a dis-intermediation of political parties, and understanding *Aló Presidente* as a way

[40] I want to thank Karl Welzel for this suggestion.

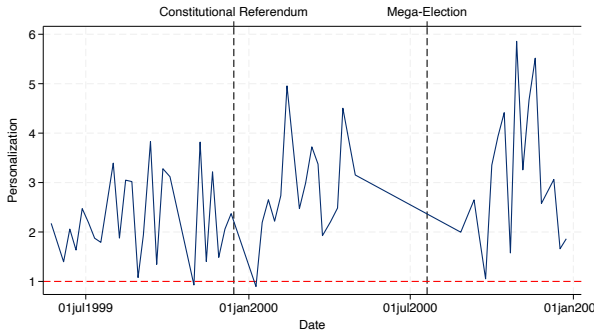
of by-passing traditional means of communication. These analyses mostly rely on data from Latinobarometer. I devote Appendix 1.G to describing this data and providing further details on its analysis.

A narrative of democracy. The Constitution of 1999 introduced massive changes to political life in Venezuela. Some key innovations were giving voting rights to the military, introducing recall referenda, as well as re-legitimising everyone who had a position in power through the 'Mega-Election' of 2000. Clearly, these changes extended the notion of direct democracy for the Venezuelan electorate. I find quantitative evidence that this is the case, as I document correlational evidence of an increase in the share of Venezuelans reporting that their country is an established democracy as a consequence of being exposed to RNV (see Appendix 1.G and Figure 1.G.1). Further, there is a strong correlation between votes for Chávez in 1998 and locations where individuals predominantly believed Venezuela was not yet a democracy as measured in the pre-1998 waves of the survey. This relationship is completely flipped when we look at RNV's effect i.e., changes in 2000 relative to 1998: the increase is coming from locations that did not originally support Chávez (Table 1.A.7). Finally, from a qualitative point of view, the topics covered by Chávez in *Aló Presidente* are closely related to expressions of direct democracy. As documented in Figure 1.3, Chávez openly discussed the new constitution as the referendum approached. Similarly for the 'Mega-Election' of 2000.

Personalised communication. Chávez's style of communication can be characterised by a unique charisma and sense of familiarity that he generated among the Venezuelan electorate (see e.g., Cannon, 2013; Frajman, 2014; Marcano & Barrera Tyszka, 2007). I can quantify the degree of personalisation of *Aló Presidente* relative to Chávez average speech off the radio. Relying on the methodology from Van Aelst, Sheafer, and Stanyer (2012), I create an aggregate index by taking into account three dimensions: 1) visibility i.e., direct mentions of Chávez; 2) mentions of Chávez's characteristics in personal rather than political contexts; 3) mentions of Chávez's personal life. I find the values of this index for all transcripts of *Aló Presidente* and Chávez's speeches off the radio. I then calculate the average degree of personalisation of the latter, and find the ratio between the index for each broadcast of *Aló Presidente* and the average speech. All broadcasts of *Aló Presidente* show a more personalised style of communicating than the average Chávez speech as documented in Figure 1.6.[41]

Other potential mechanisms. Two additional mechanisms that I test, but that fail to explain the results from section 1.4.1 are a dis-intermediation of political parties, and *Aló Presidente* as a way of by-passing traditional means of communication. *Aló Presidente* can

[41] For more details, see Appendix 1.D.



Notes: The plot shows the values of the index created using the methodology from Van Aelst et al. (2012) on the degree of personalisation of each broadcast of *Aló Presidente* relative to the average value of Chávez's speeches off the radio. Values higher than 1 indicate a higher degree of personalisation. See Appendix 1.D for more details.

Figure 1.6: Personalisation of *Aló Presidente*

be understood as realizing the technological potential of radio to eliminate the transaction cost incurred by political parties of intermediating between politicians and the electorate. Essentially, Chávez cutting the middle-man through *Aló Presidente*. I test this hypothesis looking whether the effect of RNV is driven by locations where citizens have reported distrusting political parties. Results are not driven by these locations (Table 1.A.8). Similarly, I test whether *Aló Presidente* could be understood as a way of cutting the messenger i.e., traditional means of communication. Results are not driven either by locations where individuals predominantly distrust traditional means of communication such as the printed press (Table 1.A.9).

1.7 Conclusions and Limitations

Following Hugo Chávez's original electoral victory in 1998, Venezuela's official radio broadcasting station, *Radio Nacional de Venezuela*, became an essential tool of state propaganda. I focus on a key transition period when two key political changes took place following Chávez's election: the adoption of the Constitution of 1999 and the 'Mega-Election' of 2000. This paper highlights the role that propaganda can play in consolidating the political power of a populist candidate.

I find evidence of an increase in the vote share that Hugo Chávez received in the 'Mega-Election' of 2000. I do not find an effect of being exposed to RNV on Chávez's share in 1998 nor an effect on the first wave of the Constitutional Referendum of 1999. These placebo tests nuance the findings on the 'Mega-Election' of 2000 and suggest that effects are not driven just by exposure to RNV, nor by exposure to RNV under Chávez's regime. The main change in content that took place between the first wave of the Constitutional Referendum

and Chávez's re-election was the introduction of *Aló Presidente*.

I test whether this effect can be explained as an artefact of political mobilisation via electoral turnout or persuasion. Even if exposure to RNV also reduced electoral turnout, this reduction cannot fully account for the increase in Chávez's share – suggesting that persuasion might have played a role. If voters are Bayesian, we should observe an increase in Chávez's share among the marginal voters. I find suggestive evidence that this is the case, as I observe greater effects in parishes with above-median living conditions and literacy rates. I also observe stronger effects in denser parishes. These findings follow from a natural concavity in the persuasiveness of media, specially in contexts where voters are clearly sorted, such as two elections held in only two-years time. I probe the reliability of these results with a series of strong robustness tests. I strengthen the claim that these findings are causal by including the correction for a potential OVB akin to that described in Borusyak and Hull (2023).

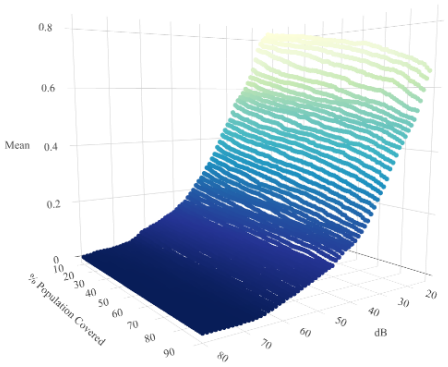
Finally, I explore the mechanisms underlying these results. I find both qualitative and quantitative evidence suggesting that the increase in Chávez's share is coming from voters being persuaded by a strong narrative of democracy, as well as the strong appeal of a personalised style of communication. These mechanisms highlight a tension between liberal and direct democracy exemplified by the deep reforms implemented through the Constitution of 1999. This is exemplified by the elimination of political counter-weights and the introduction of recall referenda, or the paradoxical evidence documented in the V-Dem data and Latinobarometer. Other potential mechanisms such as the dis-intermediation of political parties and by-passing traditional means of communication help explain Chávez's rise to power, but fail to account for the effects documented in the paper's findings.

This is the first contribution to the literature on populism and media from a middle-income country focusing exclusively on a left-wing populist leader. It is also the first paper in the literature that focuses on the direct use of means of communication by a populist leader. This paper is a timely reminder of the fragility of democracy and the power of propaganda.

Appendix

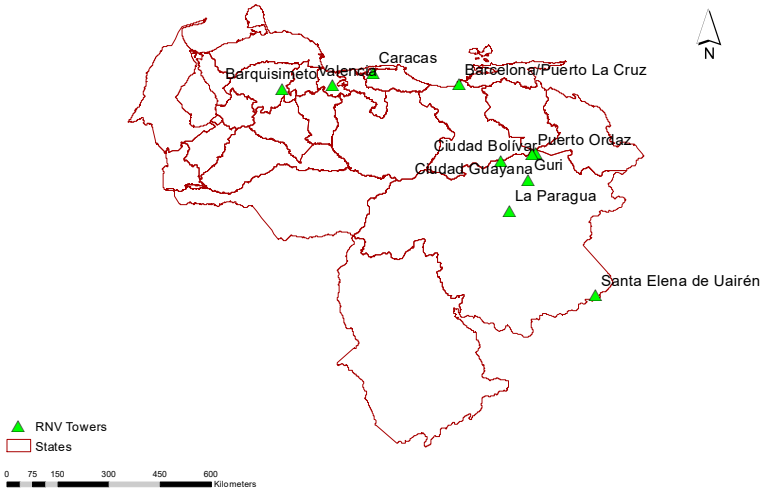
1.A Additional Figures and Tables

1.A.1 Additional Figures



Notes: Figure shows the proportion of parishes coded as receiving RNV's signal (vertical axis) when changes are made to the two thresholds used to define the main independent variable of the paper: the signal strength threshold, and the proportion of a parish's population covered by the signal of a given strength (horizontal axes). Movements along the strength signal generate much larger changes, than movements along the population threshold.

Figure 1.A.1: Proportion of Parishes Coded as Receiving RNV by Population and Field Strength Thresholds



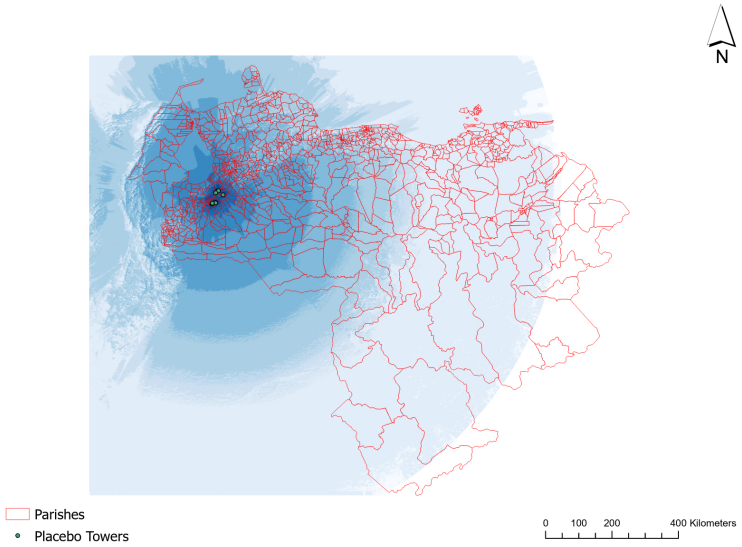
Notes: Map denoting the location of the 10 radio towers used to calculate RNV's signal. The shapefile is from Porto Tapiquén (2020). All maps were processed in ArcGIS Pro.

Figure 1.A.2: Location and Names of RNV Towers in Venezuela



Notes: Map denoting the location of radio towers used to construct the measure of exposure to RNV's signal, and the average luminosity. The map illustrates the non-random placement of radio towers. The shapefile is from Porto Tapiquén (2020). All maps were processed in ArcGIS Pro.

Figure 1.A.3: Night-Time Luminosity and Radio Towers



Notes: Map illustrating the location of the 10 placebo towers, along with the raw simulation results after adopting the same assumptions used to generate the simulations for the main measure of exposured used in the paper. Venezuela's shapefile is from Porto Tapiquén (2020). All maps were processed in ArGIS Pro.

Figure 1.A.4: Simulation of Placebo Network

1.A.2 Additional Tables

Table 1.A.1: Residualised Exposure to RNV on Demographics and Political Variables

	(1)
Unemployment Rate	0.179
	(0.166)
Δ Registered Voters 1998-2000	0.014
	(0.052)
Possible Corruption Dummy	0.013
	(0.017)
Population	8.620*
	(4.968)
Log Parish's Area	-0.009
	(0.005)
Proportion Men	0.332
	(0.318)
Latitude	-0.006
	(0.006)
Longitude	0.001
	(0.002)
TRI	-0.003
	(0.005)
TRI MAD	0.000
	(0.000)
Number Voting Centres	0.002
	(0.002)
Number Automatic Voting Tables	0.003
	(0.026)
Number Active Voting Centres	0.017
	(0.094)
Registered Voters 2000	0.000
	(0.000)
Registered Voters 1999	-0.000
	(0.000)
Men	-0.000
	(0.000)
Women	-0.000
	(0.000)
Number of Houses	-0.000
	(0.000)
Number of Occupied Houses	0.000
	(0.000)
Number of Occupants	0.000
	(0.000)
Occupants per House	-0.015
	(0.013)
Number of Unoccupied Houses	0.000
	(0.000)
Number of Occasional Houses	0.000
	(0.000)
Number of Houses Under Construction	-0.000
	(0.000)
Observations	1096
R^2	0.012

Notes: Residualised exposure on demographics and political variables. RNV's coverage is the proportion of each parish's population covered with a signal with a field strength of at least 32 dBμV/m. Residual exposure was calculated after controlling for distance to closest parish, average luminosity, the poverty index, illiteracy rate, and state indicators. Robust standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

Table 1.A.2: Exposure to RNV on Turnout in the Venezuelan 'Mega-Election' of 2000

	Turnout 2000		
	(1)	(2)	(3)
RNV Coverage	-0.035***	-0.030***	-0.039***
	(0.011)	(0.009)	(0.008)
	{0.002}	{0.002}	{0.000}
Turnout Rate 1998		0.456***	0.458***
		(0.048)	(0.046)
		{0.000}	{0.000}
Observations	1096	1096	1096
Control Mean	0.586	0.586	0.586
LASSO	No	No	Yes
R^2	0.309	0.449	0.444

Notes: Effect of RNV's Signal on turnout in 2000. RNV's coverage is the proportion of each parish's population covered with a signal with a field strength of at least 32 dBμV/m. All models include the distance to the nearest RNV tower, the parish's population as a share of the country's population, the average luminosity in the parish in 1998, the change in registered voters at each parish between 1998 and 2000, demographics, a dummy variable that takes a value of 1 when the number of null votes in a parish exceeds 5% and most of the polling stations in the parish were electronic, and state indicators. Robust standard errors in parentheses. Sharpened p-values in curly brackets. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 1.A.3: Exposure to RNV on the Venezuelan 'Mega-Election' of 2000: Penalised state indicators and Westfall-Young P-Values

	Δ Turnout		Δ Chávez's Share	
	(1)	(2)	(3)	(4)
RNV Signal (Proportion)	-0.030***	-0.035***	0.034***	0.082***
	(0.009)	(0.008)	(0.012)	(0.014)
	{0.001}	{0.000}	{0.007}	{0.000}
Observations	1096	1096	1092	1096
Control Mean	0.599	0.599	0.583	0.599
Covariates	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes
LASSO	No	Yes	No	Yes
F-Statistic	30.515	38.258	70.614	25.904
R^2	0.449	0.447	0.683	0.372

Notes: Effect of RNV's Signal on turnout and Chávez's vote share in 2000. RNV's coverage is the proportion of each parish's population covered with a signal with a field strength of at least 32 dBμV/m. All models include the distance to the nearest RNV tower, the parish's population as a share of the country's population, the average luminosity in the parish in 1998, the change in registered voters at each parish between 1998 and 2000, demographics, a dummy variable that takes a value of 1 when the number of null votes in a parish exceeds 5% and most of the polling stations in the parish were electronic, and state indicators. Indicators penalised via Lasso in even-numbered columns. Robust standard errors in parentheses. Westfall-Young p-values in curly brackets. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 1.A.4: RNV Deepens Chávez's Share

	Chávez's Share (2000)			
	(1)	(2)	(3)	(4)
RNV	-0.014	-0.007	-0.015	-0.009
	(0.018)	(0.018)	(0.018)	(0.017)
Chávez Won (1998)	0.102***	-0.006	0.102***	-0.004
	(0.007)	(0.009)	(0.007)	(0.009)
RNV $\times$ Chávez Won	0.078***	0.055***	0.076***	0.055***
	(0.017)	(0.014)	(0.016)	(0.014)
Observations	1096	1092	1096	1092
Chávez's Share (1998)	No	Yes	No	Yes
Turnout Rate (1998)	No	No	Yes	Yes
R^2	0.549	0.689	0.571	0.701

Notes: Effect of RNV's signal interacted with indicator for parishes in which Chávez won in 1998 on Chávez's vote share in 2000. RNV's coverage is the proportion of each parish's population covered with a signal with a field strength of at least 32 dBμV/m. All models include the distance to the nearest RNV tower, the parish's population as a share of the country's population, the average luminosity in the parish in 1998, the change in registered voters at each parish between 1998 and 2000, demographics, a dummy variable that takes a value of 1 when the number of null votes in a parish exceeds 5% and most of the polling stations in the parish were electronic, and state indicators. Additional controls detailed at the bottom of table. Robust standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

Table 1.A.5: RNV Decreased Turnout Across the Sample

	Turnout Rate (2000)	
	(1)	(2)
RNV	-0.032**	-0.030**
	(0.015)	(0.013)
Chávez Won (1998)	-0.025***	-0.024***
	(0.006)	(0.006)
RNV $\times$ Chávez Won	-0.002	0.001
	(0.013)	(0.011)
Observations	1096	1096
Turnout Rate (1998)	No	Yes
R^2	0.327	0.465

Notes: Effect of RNV's signal interacted with indicator for parishes in which Chávez won in 1998 on turnout in 2000. RNV's coverage is the proportion of each parish's population covered with a signal with a field strength of at least 32 dBμV/m. All models include the distance to the nearest RNV tower, the parish's population as a share of the country's population, the average luminosity in the parish in 1998, the change in registered voters at each parish between 1998 and 2000, demographics, a dummy variable that takes a value of 1 when the number of null votes in a parish exceeds 5% and most of the polling stations in the parish were electronic, and state indicators. Additional controls detailed at the bottom of table. Robust standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

Table 1.A.6: RNV's Effect on Chávez's Share was Lower Where Turnout Decreased

	Chávez's Share (2000)			
	(1)	(2)	(3)	(4)
RNV	0.070***	0.050***	0.069***	0.049***
	(0.020)	(0.016)	(0.019)	(0.015)
Turnout Decreased	0.028***	0.003	0.046***	0.014**
	(0.008)	(0.007)	(0.008)	(0.007)
RNV $\times$ Turnout Decreased	-0.031*	-0.027**	-0.035**	-0.029**
	(0.018)	(0.014)	(0.017)	(0.013)
Observations	1096	1092	1096	1092
Chávez's Share (1998)	No	Yes	No	Yes
Turnout Rate (1998)	No	No	Yes	Yes
R^2	0.371	0.685	0.407	0.697

Notes: Effect of RNV's signal interacted with indicator for parishes in which tunout decreased between 1998 and 2000 on Chávez's vote share in 2000. RNV's coverage is the proportion of each parish's population covered with a signal with a field strength of at least 32 dBμV/m. All models include the distance to the nearest RNV tower, the parish's population as a share of the country's population, the average luminosity in the parish in 1998, the change in registered voters at each parish between 1998 and 2000, demographics, a dummy variable that takes a value of 1 when the number of null votes in a parish exceeds 5% and most of the polling stations in the parish were electronic, and state indicators. Additional controls detailed at the bottom of table. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 1.A.7: RNV and the Narrative of Democracy

	Chávez's Share (1998)	Chávez's Share (2000)			
	(1)	(2)	(3)	(4)	(5)
RNV	0.017	0.044**	0.032*	0.044**	0.032**
	(0.020)	(0.019)	(0.016)	(0.018)	(0.016)
Not Democracy Yet	0.071**	0.006	-0.046**	0.008	-0.044**
	(0.029)	(0.028)	(0.020)	(0.027)	(0.020)
RNV $\times$ Not Democracy Yet	-0.003	-0.011	-0.014	-0.025	-0.024
	(0.028)	(0.029)	(0.024)	(0.027)	(0.022)
Observations	869	872	869	872	869
Chávez's Share (1998)	-	No	Yes	No	Yes
Turnout Rate (1998)	Yes	No	No	Yes	Yes
R^2	0.372	0.322	0.640	0.347	0.652

Notes: Column 1 shows correlation between RNV's signal and above-median indicator for municipalities where pre-Chávez Latinobarometer respondents believed Venezuela was not yet a democracy on Chávez's vote share in 1998. Columns 2-4 show effect Chávez's vote share in 2000. RNV's coverage is the proportion of each parish's population covered with a signal with a field strength of at least 32 dBμV/m. All models include the distance to the nearest RNV tower, the parish's population as a share of the country's population, the average luminosity in the parish in 1998, the change in registered voters at each parish between 1998 and 2000, demographics, a dummy variable that takes a value of 1 when the number of null votes in a parish exceeds 5% and most of the polling stations in the parish were electronic, and state indicators. Additional controls detailed at the bottom of table. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 1.A.8: Potential Mechanisms: Dis-Intermediation of Political Parties

	Chávez's Share (1998)	Chávez's Share (2000)			
	(1)	(2)	(3)	(4)	(5)
RNV	0.021	0.061***	0.044***	0.044***	0.044***
	(0.021)	(0.019)	(0.015)	(0.015)	(0.015)
Distrust Political Parties	0.094***	0.025	-0.061***	-0.061***	-0.061***
	(0.023)	(0.024)	(0.017)	(0.017)	(0.017)
RNV $\times$ Distrust Political Parties	-0.014	-0.059*	-0.057**	-0.057**	-0.057**
	(0.029)	(0.032)	(0.023)	(0.023)	(0.023)
Observations	869	872	869	869	869
Chávez's Share (1998)	-	No	Yes	No	Yes
Turnout Rate (1998)	Yes	No	No	Yes	Yes
R^2	0.372	0.325	0.654	0.654	0.654

Notes: Column 1 shows correlation between RNV's signal and above-median indicator for municipalities where pre-Chávez Latinobarometer respondents distrusted political parties on Chávez's vote share in 1998. Columns 2-4 show effect Chávez's vote share in 2000. RNV's coverage is the proportion of each parish's population covered with a signal with a field strength of at least 32 dBμV/m. All models include the distance to the nearest RNV tower, the parish's population as a share of the country's population, the average luminosity in the parish in 1998, the change in registered voters at each parish between 1998 and 2000, demographics, a dummy variable that takes a value of 1 when the number of null votes in a parish exceeds 5% and most of the polling stations in the parish were electronic, and state indicators. Additional controls detailed at the bottom of table. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 1.A.9: Potential Mechanisms: By-Passing Traditional Means of Communication

	Chávez's Share (1998)	Chávez's Share (2000)			
	(1)	(2)	(3)	(4)	(5)
RNV	0.004	0.045**	0.040**	0.037**	0.034**
	(0.019)	(0.019)	(0.015)	(0.018)	(0.015)
Distrust Press	-0.078***	-0.080***	-0.026	-0.089***	-0.032*
	(0.018)	(0.021)	(0.017)	(0.020)	(0.017)
RNV $\times$ Distrust Press	0.051	-0.022	-0.056**	-0.008	-0.045*
	(0.033)	(0.033)	(0.025)	(0.031)	(0.023)
Observations	869	872	869	872	869
Chávez's Share (1998)	-	No	Yes	No	Yes
Turnout Rate (1998)	Yes	No	No	Yes	Yes
R^2	0.374	0.322	0.641	0.346	0.653

Notes: Column 1 shows correlation between RNV's signal and above-median indicator for municipalities where pre-Chávez Latinobarometer respondents distrusted the press on Chávez's vote share in 1998. Columns 2-4 show effect Chávez's vote share in 2000. RNV's coverage is the proportion of each parish's population covered with a signal with a field strength of at least 32 dBμV/m. All models include the distance to the nearest RNV tower, the parish's population as a share of the country's population, the average luminosity in the parish in 1998, the change in registered voters at each parish between 1998 and 2000, demographics, a dummy variable that takes a value of 1 when the number of null votes in a parish exceeds 5% and most of the polling stations in the parish were electronic, and state indicators. Additional controls detailed at the bottom of table. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

1.B Details on Data Sources and Empirical Strategy

1.B.1 Documents Used to Infer List of RNV Towers

This list was inferred from two sources of information accessed via the Wayback Machine.[42] The first source is an article in RNV's website from April 2004, covering the inauguration of 21 FM frequencies all over the country as part of the station's aggressive expansion process. The article's headline states that RNV was operating in 30 frequencies following this inauguration. The second source is a list of RNV's full list of frequencies from September 2004 (5 months after the inauguration). One can match the frequencies in the list to those inaugurated in April 2004 and see that the remaining frequencies are a handful of FM frequencies that broadcast music and 10 AM frequencies. Both articles are included below, along with translations from Google Translate. To the best of my knowledge, there is no academic discussion on the history of FM radio in Venezuela. However, in an interview, Polo Troconis, a Venezuelan radio host recalls that FM radio was only used for political talk shows (most notably Chávez's *Aló Presidente*) well into the 21st century – way after Chávez's presidency started (Velásquez Malavé, 2009). Since I cannot determine which of the 10 AM frequencies was inaugurated between April and September 2004, and I cannot guarantee that this additional frequency did not exist prior to April 2004 and that the article's headline reported 30 frequencies rather than 31 for stylistic reasons, I keep the 10 AM frequencies as the list of towers that were active in 1998. Section 1.5 explores robustness of the paper's main results to different ways of performing the simulation – including dropping one tower from the simulation at a time.

[42] The earliest captures of RNV's website available through the Wayback Machine are from August 2003.

First article:



The Wayback Machine - https://web.archive.org/web/20041205000121/http://www.rnv.gov.ve:80/noticias/index.php?act=ST&f=22&t=5123

La Radio de la Información Veraz

Venezuela AHORA ES DE TODOS

PORTADA
FRECUENCIAS
Nacionales
Internacionales
AUDIO EN VIVO
Informativo
Clásico
YVKE Mundial
NOTICIAS
Nacional y Política
Provincia
Economía
Educación
Salud
Internacionales
Deportes
Culturales
Tendencias
RNV
Novedades
OPINIÓN
¡Participe!
Artículos
Entrevistas
Documentos
OTROS IDIOMAS
English
GALERÍA
MULTIMEDIA
CONTÁCTENOS
CRÉDITOS

Presidencia
Vicepresidencia
Venpres
Venezolana de Televisión
vive

RNV > NOVEDADES

Se planifican estudios regionales

RNV cuenta con 30 frecuencias a nivel nacional

Con el propósito de fortalecer la política comunicacional del Gobierno Bolivariano, desde el 22 de abril inauguró 21 frecuencias FM en todo el país y reactivó su canal internacional

Véase también:
RNV ya cuenta con un canal internacional

PRENSA RNV
27 de Abril de 2004, 11:48 PM

Con el propósito de fortalecer la política comunicacional del Gobierno Bolivariano, desde el 22 de abril **Radio Nacional de Venezuela** inauguró 21 frecuencias FM en todo el país. De esta forma, RNV se convierte en el circuito radial más grande de Venezuela y pasa a ser la plataforma comunicacional con mayor capacidad de alcance en la población.

Durante este importante evento realizado en la sede de Radio Nacional de Venezuela, al cual asistieron productores de la radio y representantes del Gobierno Nacional, el ministro de Información y comunicación Jesse Chacón manifestó que a través de este tipo de iniciativas comunicacionales se democratiza el derecho a la información que tienen todos los venezolanos. El titular del MINCI informó que en pocos meses saldrá al aire una emisora de corte juvenil que pretende tratar al público nacional de este tipo como seres pensantes y capaces de forjarse un destino donde no se les trate como meros objetos susceptibles a ser manipulados. Jesse Chacón destacó también que se está trabajando para establecer un circuito de emisoras en amplitud modulada (AM), desplegadas en todo el país, las cuales tendrán como objetivo cautivar el público adulto inmerso en la población nacional

Por su parte, la directora de RNV, Helena Salcedo, expresó su satisfacción al concretar este proyecto resaltando que "ahora Radio Nacional de Venezuela es de todos los Venezolanos". Exaltó además la voluntad que tiene RNV de cumplir con el articulado del proyecto de la Ley de Responsabilidad Social de Radio y Televisión, acotó que la programación del Canal Informativo, el cual representa la matriz para todas las frecuencias del interior de Venezuela, será orientada cada vez más con el objetivo de cumplir con lo estipulado en este proyecto de ley.

Vale destacar que Radio Nacional de Venezuela tiene planificado la creación de estudios regionales donde están las frecuencias en el país, con la intención de promover la producción local y consolidar un cuerpo de productores independientes que se unan al Circuito RNV a través de programas destinados al público de cada zona.

Rompiendo fronteras
Otro logró que se destacó durante la inauguración oficial fue la puesta en funcionamiento del Canal Internacional, el cual difundirá por medio de una señal onda corta, a los Estados Unidos, Centro América, Rió de Janeiro, Buenos Aires, Chile y las Islas del Caribe; los logros del Gobierno Bolivariano y todas las potencialidades que tiene Venezuela.

Sobre este punto, el ministro Jesse Chacón aseveró que una de las estrategias que manejaron los gobiernos anteriores fue la de dejar deteriorar los equipos de Radio Nacional, Venezolana de Televisión, Venpres y la Imprenta: "durante los últimos 20 años no se realizó ningún tipo de inversión en los canales del Estado" informó. Esto trajo como consecuencia, explicó el ministro, el deterioro de transmisores que permitían difundir en vivo toda la programación de RNV hasta el exterior.

Lista de frecuencias
Región Sur
Barinas, Estado Barinas: 91.7 FM

MÁS NOTICIAS

Comienza la navidad en Radio Nacional de Venezuela

La Librería Mediática: Solicitud de adhesión a propuesta a la Ley Resorte

Comienza a funcionar emisora de radio en Granada con ayuda de RNV

"La Hojilla" todos los lunes y miércoles en RNV

Mantenimiento a la Torre de Transmisión en Puerto Ordaz

Descargue la "DIANA" del triunfo en Santa Inés

Descarga el VIDEO y la canción "¡DI QUE NO!" transmitida en Dossier

RNV ya cuenta con un canal internacional

ver más >

(Continues next page)

(Continued from previous page)

San Fernando, Estado Apure: 91.1 FM
Ciudad Bolívar, Estado Bolívar: 98.5 FM
Ciudad Guayana, Estado Bolívar: 99.1 FM
Puerto Ayacucho, Estado Amazonas: 96.1 FM

Región Occidental
Maracaibo, Estado Zulia: 91.7 FM
Barquisimeto, Estado Lara: 91.7 FM
Coro, Estado Falcón: 89.7 FM
San Carlos, Estado Cojedes: 88.9 FM

Región Los Andes
Mérida, Estado Mérida: 90.5 FM
San Cristóbal, Estado Táchira: 105.7 FM
Trujillo, Estado Trujillo: 106.5 FM

Región Capital y Central
Guarenas y Guatire, Estado Miranda: 99.5 FM
Río Chico y Barlovento, Estado Miranda: 91.9 FM
Valencia, Estado Carabobo: 90.5 FM

Región Oriental
El Tigre, Estado Anzoátegui: 95.7 FM
Maturín, Estado Monagas: 91.7 FM
Puerto La Cruz, Anzoátegui: 92.3 FM
Tucupita, Estado Delta Amacuro: 91.1 FM

LBR

Enviar por correo | Imprimir

<< Volver a Novedades

- Novedades ¡Ir!

© 2003 Radio Nacional de Venezuela
IPB v1.1.1 © 2003 IPS, Inc.

Translation of the article from Google Translate:

Radio Nacional de Venezuela : RNV has 30 frequencies nationwide.

With the purpose of strengthening the communication policy of the Bolivarian Government, since April 22 Radio Nacional de Venezuela inaugurated 21 FM frequencies throughout the country. In this way, RNV becomes the largest radio circuit in Venezuela and becomes the communication platform with the greatest capacity to reach the population.
27th of April 2004, 11:48 PM

During this important event held at the headquarters of Radio Nacional de Venezuela, which was attended by radio producers and representatives of the National Government, the Minister of Information and Communication Jesse Chacón stated that through this type of communication initiatives, the right to the information that all Venezuelans have. The head of the MINCI reported that in a few months a juvenile court station will go on the air that intends to treat the national public of this type as thinking beings and capable of forging a destiny where they are not treated as mere objects susceptible to being manipulated. Jesse Chacón also highlighted that work is being done to establish a circuit of amplitude modulated (AM) stations, deployed throughout the country, which will aim to captivate the adult public immersed in the national population.

For her part, the director of RNV, Helena Salcedo, expressed her satisfaction in concretizing this project, highlighting that "now Radio Nacional de Venezuela belongs to all Venezuelans." He also extolled the will of the RNV to comply with the articles of the project of the Social Responsibility Law of Radio and Television, he stated that the programming of the Information Channel, which represents the matrix for all the frequencies of the interior of Venezuela, will be oriented each again with the aim of complying with the provisions of this bill.

It is worth noting that Radio Nacional de Venezuela has planned the creation of regional studios where the frequencies are in the country, with the intention of promoting local production and consolidating a body of independent producers who join the RNV Circuit through programs aimed at the public. of each zone.

Breaking barriers
Another achievement that stood out during the official inauguration was the start-up of the International Channel, which will broadcast through a short wave signal to the United States, Central America, Rio de Janeiro, Buenos Aires, Chile and the Caribbean Islands. ; the achievements of the Bolivarian Government and all the potential that Venezuela has.

(Continues next page)

(Continued from previous page)

On this point, Minister Jesse Chacón asserted that one of the strategies used by previous governments was to allow the equipment of Radio Nacional, Venezolana de Televisión, Venpres and La Imprenta to deteriorate: "during the last 20 years no type of of investment in the channels of the State" he reported. This brought as a consequence, explained the minister, the deterioration of transmitters that allowed all RNV programming to be broadcast live abroad.

Frequency list
Southern region
Barinas, Barinas State: 91.7 FM
San Fernando, Apure State: 91.1 FM
Bolivar City, Bolivar State: 98.5 FM
Guayana City, Bolívar State: 99.1 FM
Puerto Ayacucho, Amazonas State: 96.1 FM

Western region
Maracaibo, Zulia State: 91.7 FM
Barquisimeto, Lara State: 91.7 FM
Choir, Falcon State: 89.7 FM
San Carlos, Cojedes State: 88.9 FM

Andes Region
Merida, Merida State: 90.5 FM
San Cristobal, Táchira State: 105.7 FM
Trujillo, Trujillo State: 106.5 FM

Capital and Central Region
Guarenas and Guatire, Miranda State: 99.5 FM
Río Chico and Barlovento, Miranda State: 91.9 FM
Valencia, Carabobo State: 90.5 FM

Eastern Region
El Tigre, Anzoategui State: 95.7 FM
Maturin, Monagas State: 91.7 FM
Puerto La Cruz, Anzoategui: 92.3 FM
Tucupita, Delta Amacuro State: 91.1 FM

Second article:

The Wayback Machine - https://web.archive.org/web/20041204173238/http://www.rnv.gov.ve:80/noticias/index.php?pg=frecuencias

RNV Radio Nacional de Venezuela

Venezuela AHORA ES DE TODOS

Presidencia
Vicepresidencia
Venpres
Venezolana de Televisión
vive

Radio Nacional de Venezuela

EMISORAS DE RADIO NACIONAL DE VENEZUELA EN TODO EL PAÍS

A menos que se indique lo contrario, las frecuencias corresponden al Canal Informativo.
(Actualizadas al 14 de septiembre de 2004)

Ciudad	Estado	Dial	Tipo
Barquisimeto	Lara	91.1	FM
Barquisimeto	Lara	1.050	AM
Caracas *(Canal Clásico)*	Distrito Capital	91.1	FM
Caracas	Distrito Capital	630	AM
Caracas *(Canal Musical)*	Distrito Capital	1.050	AM
Ciudad Bolívar	Bolivar	98.5	FM
Ciudad Bolívar	Bolivar	790	AM
Coro	Falcon	89.5	FM
Cumaná	Sucre	99.5	FM
El Tigre	Anzoátegui	95.7	FM
Guarenas y Guatire	Miranda	99.5	FM
Güiria	Sucre	91.5	FM
Guri	Bolivar	1310	AM
La Paragua	Bolivar	1330	AM
Maiquetía	Vargas	106.1	FM
Maracaibo	Zulia	91.7	FM
Maracay	Aragua	89.3	FM
Margarita	Nueva Esparta	102.3	FM
Maturín	Monagas	91.7	FM
Mérida	Merida	90.5	FM
Puerto Ayacucho	Amazonas	96.1	FM
Puerto Cabello	Carabobo	92.5	FM
Puerto La Cruz y Barcelona	Anzoátegui	92.3	FM
Puerto La Cruz y Barcelona	Anzoátegui	1.310	AM
Puerto Ordaz	Bolívar	560	AM
Rio Chico y Barlovento	Miranda	91.9	FM
San Carlos	Cojedes	88.9	FM
San Cristóbal	Tachira	105.7	FM
San Felipe	Yaracuy	99.7	FM
San Fernando de Apure	Apure	91.1	FM
Santa Elena de Uairén	Bolivar	1310	AM
Trujillo	Trujillo	106.5	FM
Tucupita	Delta Amacuro	91.1	FM
Valencia	Carabobo	90.5	FM
Valencia	Carabobo	770	AM
Valle La Pascua	Guárico	88.9	FM

En reparación
Las frecuencias en las ciudades señaladas a continuación están en reparación o en revisión y serán reactivadas en las próximas semanas.

Ciudad	Estado	Dial	Tipo
Barinas	Barinas	91.7	FM
Cabimas	Zulia	100.7	FM
Ciudad Guayana	Bolivar	99.1	FM
Ciudad Guayana	Bolivar	560	AM

(Continues next page)

(Continued from previous page)

El Vigia		Zulia	94.7 FM
Valles del Tuy		Miranda	95.9 FM

© 2003 Radio Nacional de Venezuela
IPB v1.1.1 © 2003 IPS, Inc.

Translation of the second article from Google Translate:

National Radio stations of Venezuela throughout the country
Unless otherwise indicated, the frequencies correspond to the Informative Channel.
(Updated to September 14, 2004)

City	State	Dial	Type
(Table omitted from translation)			

In repair
The frequencies in the cities indicated below are under repair or review and will be reactivated in the coming weeks.

City	State	Dial	Type
(Table omitted from translation)			

1.B.2 Simulating Radio Signals

How to replicate:

1. Input the location of each radio station according to the data gathered from FMRadio.org into Radio Mobile.
2. Create an additional unit 2m above the ground level to simulate reception into a standard radio or a vehicle. Call this unit Radio for convenience.
3. Create networks in Radio Mobile, choosing the following specification, as in Blouin and Mukand (2019):
    - Choose "broadcast" in Mode of variability.
    - Set surface refractivity to 310.
    - Set ground conductivity to 0.005.
    - Set relative ground permitivity to 15.
    - Choose "vertical" in the type of polarization.
    - Choose "equatorial" climate.
    - Set variability % of time to 60.
    - Set variability % of location to 60.
    - Set variability % of situations to 60.
    - Choose "visible and voice-net (command/subordinate/rebroadcast)" in topology.
4. Simulate each of the networks using Combined Cartesian Radio Coverage. In order to generate data as fine as possible I used a map with dimensions 2000 x 1500 pixels that covered the entire country.
    (a) On the left panel, select the fixed units that comprise each of the individual networks.
    (b) Set the mobile unit to Radio–the unit created above to represent a receiving radio or vehicle.
    (c) Select the desired network for the simulation.
    (d) Tick off the option "Save raster data to file (TXT)".
    (e) Set maximum range to 1000 km.
    (f) Set Link Direction to Mobile Rx.
    (g) In signal range to draw, choose dBμV/m.

(h) Set draw pixel size to 0.0001.

(i) Choose the displayed range of values from 0 to 200.

(j) Click "Draw" to begin simulation.

(k) After simulation finishes, save text file with results where desired.

5. The simulations from Radio Mobile are saved in a text file. We need to eliminate the preamble of those files so that the only information in the files are the variables names and their corresponding values.

6. In ArcMap, we clock on Add Data and look for the text file we wish to import.

7. ArcMap will save this file as a table. In the table of Contents, we have to right-click on the table and choose Display XY Data.

8. After choosing the correct set of coordinates and the right Projection and Coordinate Systems (the easiest thing to do is select the layer's we want to attach the data the data to, in order to avoid confusion).

9. ArcMap will now display the data as a separate layer. We need to right-click on this layer and export the data into the Data Frame we are using.

10. Now we can use ArcMap's geoprocessing tools. For instance:

    (a) To find the centroid of each parish, use the tool "Features to Point" ticking off the option "Include".

    (b) After importing the location of each radio tower, calculate the distance from the centroid of each parish to the nearest 3 towers using the tool "Generate Near Table" and choose "Geodesic" in method and set the maximum number of closest matches to 3.

    (c) Use "Spatial Join" to join the parishes into the layer of points representing the strength of the radio signal. This way we will have the name of the parish, municipality and state in which each point is located.

1.B.3 Field Strength Calculations

The electromagnetic spectrum is divided into bands that differ by the frequency of the waves that travel through it. FM broadcasts correspond to the Very High Frequency (VHF) band, which comprises frequencies of 30-300 MHz. The VHF band is usually divided into three other bands: Band I (47-68 MHz), Band II (87.5-108 MHz) and Band III (230-240 MHz). AM broadcasts fall into the Medium Frequency (MF) band, which includes frequencies in a

range 0.3-30 MHz. Table 7.3 from Hoeg and Lauterbach (2003) shows that the threshold of 45 dBμV/m corresponds to the value used for frequencies in the range of 1452–1480 MHz. These frequencies fall into the Ultra High Frequency (UHF) band, generally used by televisions, mobile phones, bluetooth or GPS devices. Table 2.2.3 from CEPT (2007) shows the minimum field strength needed to determine that a specific area receives a strong enough signal for the three VHF bands. These values range from 25 dBμV/m for Band I to 35 dBμV/m for Band III. These numbers are smaller in magnitude than the threshold I adopt in the paper. This is because these thresholds must yet face an adjustment of +13 dBμV/m if we want to increase the local reception from 50% to 99% (or +18.2 dBμV/m if there is potential interference from other signals), an additional +10 dBμV/m if we are considering the receiving antenna to be at 1.5m above ground level (rather than 10m), and a third adjustment to account for building penetration. This last value depends on the frequency of the signal being measured (see for example Table 4.3 from EBU (2013)). These adjustments should be adopted given that there are many factors that cannot be accounted for in the simulations.

Next, I provide details of the calculations used to determine the field strength threshold I adopt throughout the article. All formulae are from EBU (2013). The notation employed in this section does not extend to the rest of the article and vice versa.

The first step involves calculating the minimum receiver input power (P_s^{min}). This is determined by the following two equations:

$$\begin{aligned} P_n &= F + 10 * log_{10}(k * T_o * B) \\ P_s^{min} &= P_n + C/N, \end{aligned}$$

where P_n corresponds to the corresponding receiver noise input power, F is the receiver noise figure, k is Boltzmann's constant, T_o is the absolute temperature, B is the equivalent noise bandwidth, and C/N refers to the signal to noise ratio required by a particular system. The minimum receiver input power is added a correction for mobile reception: $P_s^{min,corr} = P_s^{min} + C/N_{mobile}$. We now calculate the effective antenna aperture (A_a):

$$A_a = G_D + 10 * log_{10}\left(1.64 * \frac{\lambda^2}{4\pi}\right),$$

where G_D is the antenna gain relative to a half wave dipole, and λ corresponds to the wavelength of the particular signal we are analysing. We can then calculate the minimum power flux density at the receiving place (Φ^{min}):

$$\Phi^{min} = P_n^{min,corr} - A_a.$$

This can then be used to calculate the minimum equivalent field strength at the receiving place (E^{min}), following:

$$\begin{aligned} E^{min} &= \Phi^{min} + P_{mmn} + 120 + 10 * log_{10}(120\pi) \\ &= \Phi^{min} + P_{mmn} + 145.8, \end{aligned}$$

where P_{mmn} is an allowance for man made noise. We can improve this calculation and incorporate other adjustments such as a location correction factor (or margin) (C_l), a correction for building penetration loss (L_b), and a correction for height loss (L_h) going from 10m to 1.5m above ground level. This yields a minimum median power flux density required at 10m above ground level for 50% of the time and 50% of locations, with corrections to achieve coverage in 99% of locations at 1.5m above ground level (Φ^{med}):

$$\Phi^{med} = \Phi^{min} + P_{mmn} + C_l + L_b + L_h.$$

Finally, the minimum median equivalent field strength is calculated as:

$$\begin{aligned} E^{med} &= \Phi^{med} + 120 + 10 * log_{10}(120\pi) \\ &= \Phi^{med} + 145.8. \end{aligned}$$

Thus, we have a collection of variables $(P_n, P_s^{min}, P_s^{min,corr}, A_a, \Phi^{min}, E^{min}, \Phi^{med}, E^{med})$ that can be calculated with the set of parameters (F, k, T_o, B, C/N, C/N_{mobile}, G_D, λ, P_{mmn}, C_l, L_b, L_h). Table 1.B.1 details the values I adopt for these parameters and their sources.

Table 1.B.1: Parameters from Field Strength Calculations

Parameter	Name (units)	Value	Source
F	Receiver noise figure (dB)	7	EBU (2013)
k	Bolzmann's constant (Ws/K)	1.38×10^{-23}	EBU (2013)
T_o	Absolute temperature (K)	290	EBU (2013)
B	Equivalent noise bandwidth (Hz)	1.536×10^6	EBU (2013)
C/N	Signal to noise ratio (Gaussian channel) (dB)	7.4	EBU (2013)
C/N_{mobile}	Signal to noise correction for mobile reception (dB)	5.6	EBU (2013)
G_D	Antenna gain relative to half wave dipole (dB)	-2.2	EBU (2013)
λ	Signal wavelength (m)	311.54	Calculated
P_{mmn}	Allowance for man made noise (dB)	1	EBU (2013)
C_l	Location correction margin (dB)	18.2	EBU (2013)
L_b	Building penetration loss (dB)	15.5	Fernandez et al. (2012)
L_h	Height loss adjustment from 10m to 1.5m (dB)	13	EBU (2013)

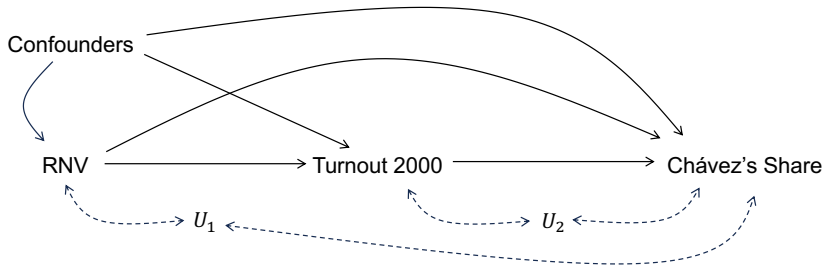
Notes: Parameter choices used to calculated median field strength threshold value. The table details the parameters, their names, the values I adopted, and the source of those values.

I calculated the wavelength (λ) with the following formula: $v = f * \lambda$, where v is the wave speed in metres per second, f is the frequency in hertz, and λ is the wavelength in

metres. In the formula, v takes the value of the speed of light (and any other wave in the electromagnetic spectrum), i.e., 2.997×10^8 m/s, and I assume a value for f of 0.962 MHz – the mean of the frequencies in which RNV operates in the sample (ranging from 0.56 to 1.33 MHz). I adopt a location correction margin rather than a smaller location correction figure to allow for possible interference with other signals. The value of 18.2 dB is the recommended value by EBU (2013) to guarantee mobile reception in 99% of locations. Finally, the building penetration loss parameter corresponds to the 98th percentile of the sample taken by Fernandez et al. (2012) from the engineering faculties in Bilbao and Vitoria-Gasteiz, and roughly to the 66th percentile of the penetration losses observed in their sample from Madrid. Their values are used instead of those reported by EBU (2013) because Fernandez et al. (2012) focus on frequencies in the MF range, while EBU (2013) focus on the VHF range. Inputting these parameters into the equations outlined above, yields a minimum median equivalent field strength value of 32.56 dBμV/m, which is rounded down to 32 dBμV/m given how conservative the building penetration loss values were.

1.B.4 Details on Mediation Analysis

To further explore whether effects on Chávez's share are driven by political mobilisation I propose a causal mediation relationship, which I represent with a Directed Acyclic Graph (DAG) in Figure 1.B.1. I follow Imai, Keele, and Yamamoto (2010) and Imai, Keele, and Tingley (2010) and estimate the average causal mediation effect (ACME) within a Baron and Kenny (1986) Linear Structural Equation Model, under the Sequential Ignorability assumption and Algorithm 1 from Imai, Keele, and Tingley (2010).



Notes: DAG representing the proposed mediated relationships between the variables in this setting. U_1 represents unobservables that would be akin to omitted variables. U_2 also represent unobservables related to the sequential ignorability assumption, which is non-refutable (Manski, 2007).

Figure 1.B.1: DAG for Mediation Analysis

Assumption 1: (Sequential Ignorability)

$$\{Y_i(t', m), M_i(t)\} \perp\!\!\!\perp T_i | \mathbf{X}_i = \mathbf{x}, \tag{1.2a}$$

$$Y_i(t', m) \perp\!\!\!\perp M_i(t) | T_i = t, \mathbf{X}_i = \mathbf{x}, \tag{1.2b}$$

for $t, t' \in \{0, 1\}$, where $M_i(t)$ denotes the potential value taken by the mediator for a unit i under treatment status $T_i = t$. Similarly, $Y_i(t', m)$ denotes the potential outcome for unit i under $T_i = t$ and $M_i = m$. $\mathbf{X}_i$ denotes a set of covariates. In this context, treatment status refers to reception of RNV, the mediator is turnout, and the relevant outcome is the change in Chávez's share between both elections, as illustrated in Figure 1.B.1.

The sequential ignorability assumption has two conditions: the first is that there is no omitted variable bias (Hicks & Tingley, 2011) represented in the DAG by U_1. In this context, this is a credible condition if we believe that topographic ruggedness is an exogenous determinant of RNV's reception given the set of covariates included in the regressions. A violation of this condition would be observed if parishes could self-select into reception of RNV. The second condition is much stronger in comparison. It implies that the potential mediator is ignorable given the treatment assignment and pretreatment confounders, which might not be the case even in randomised experiments. This is a nonrefutable assumption (Manski, 2007), and is represented in Figure 1.B.1 by the existence of U_2. Imai, Keele, and Yamamoto (2010) developed a sensitivity test to determine robustness of the results to violations of the second assumption.[43]

This sensitivity analysis hinges on a parameter ρ that quantifies the correlation of the residuals of two models: one in which the mediator is explained by the treatment and pretreatment covariates, and another model in which the outcome of interest is explained by the treatment, the mediator and the set of pretreatment covariates. Under sequential ignorability, ρ takes a value of zero. The test consists in re-estimating Algorithm 1 under different values of ρ. The objective is finding the values of ρ for which the ACME equals zero. This gives us an idea of how much of the variation between the outcome and mediator would need to be explained by a potential confounder (U_2) for the ACME to be unidentified.[44]

Algorithm 1: (Parametric Inference)

1. Fit models for the observed outcome and mediator variables,

2. Simulate model parameters from their sampling distribution,

[43] I implement the sensitivity test and identification algorithm with the Stata package mediate (Hicks & Tingley, 2011).

[44] See Maggio, Karra, and Canning (2021) for an example of an application.

3. Repeat the following three steps:

   (a) Simulate the potential values of the mediator,

   (b) Simulate the potential outcomes given the simulated values of the mediator,

   (c) Compute the causal mediation effects.

4. Compute summary statistics such as point estimates and confidence intervals.

Table 1.B.2 presents the results of these models. The coefficients associated with RNV correspond to the direct effect of being exposed to RNV on Chávez's share over and above the mediation that could have happened through turnout. Intuitively, the coefficient of turnout on Chávez's share is negative and statistically significant. However, little of this effect can be attributed to being exposed to RNV. For both models the ACME is very small in magnitude, and in the model including state indicators its 95% confidence intervals include zero. The fact that the absolute value of ρ is smaller than the R^2 of both models raises concerns about the ACME's identification: either because of a violation of the sequential ignorability assumption, or simply because the ACME is weak and close to zero anyway. Overall, from these results I cannot conclude that turnout mediated the effect of reception of RNV in a way that impacted Chávez's share in the 2000 election. The total effect, represented by τ, and the direct effect remain statistically significant and in the vicinity of the point estimates found in Table 1.1.

One potential pitfall of this analysis is the existence of intermediate confounders that could bias the estimates of the direct effect. This is an issue highlighted by Acharya, Blackwell, and Sen (2016a). This is represented in a DAG in Figure 1.B.2. If this were the case, the direct effect, represented with dotted lines in Figure 1.B.2 would be biased. Their methodology for addressing this concern and performing mediation analyses has been recently applied in Abebe et al. (2021), for instance. Further, the flexibility of their methodology allows me to nominate some of the covariates as potential intermediate confounders,[45] or assume that none are intermediate confounders and rerun the analyses under a different estimating algorithm. Table 1.B.3 shows the Average Causal Direct Effect under different specifications and DAGs. Columns 1 and 2 correspond to specifications without state indicators, while columns 3 and 4 include state indicators. Even-numbered columns assume that there are no intermediate confounders, they replicate the DAG from Figure 1.B.1. The odd-numbered columns correspond to the DAG in Figure 1.B.2, where

[45] Out of the covariates that I include in all models the ones that could be treated as intermediate confounders are the change in the number of registered voters in each parish and the indicator for corruption. The former could be determined by exposure to RNV if, for instance, it encourages people to register to vote if they would not have done it in a counterfactual world. The corruption indicator by definition is constructed with post-treatment data.

Table 1.B.2: Turnout as a Mediator of RNV

	(1)	(2)
RNV Signal (Direct Effect)	0.028***	0.018*
	(0.008)	(0.009)
Turnout Rate (2000)	-0.154***	-0.091**
	(0.042)	(0.040)
τ	0.032	0.020
τ 95% CI	[0.016,0.047]	[0.002,0.039]
ACME	0.004	0.002
ACME 95% CI	[0.002,0.007]	[0.000,0.005]
ρ ACME = 0	-0.111	-0.070
Covariates	Yes	Yes
State Indicators	No	Yes
Observations	1092	1092
R^2	0.519	0.620

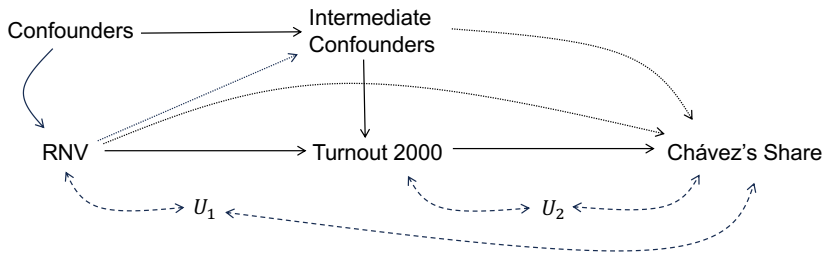
Notes: These results follow the methodology in Imai, Keele, and Tingley (2010); Imai, Keele, and Yamamoto (2010) and have been implemented in Stata using the mediate package (Hicks & Tingley, 2011). Column 1 and 2 differ in the inclusion of state indicators to the baseline model that includes the covariates detailed in section 1.4.1. RNV Signal is an indicator and quantifies the direct effect of receiving RNV's signal on Chávez's share over and above the mediated effect that could have happened through turnout. The Average Causal Mediated Effect (ACME) along with confidence intervals is reported at the bottom of the table. τ denotes the total effect combining the direct and ACME. The absolute value of ρ denotes the variation that would need to be explained by a confounder for the ACME to be unidentified either because of a violation of the sequential ignorability assumption or because the ACME is not statistically different from zero. This number should be compared to the regression's R^2 as a benchmark.

corruption and the change in registered voters per parish correspond to the proposed intermediate confounders. Overall, results are robust to using this alternative algorithm.

Table 1.B.3: Average Causal Direct Effect of RNV with Turnout as Mediator

	(1)	(2)	(3)	(4)
RNV Signal (Direct Effect)	0.032***	0.032***	0.021**	0.020**
	(0.007)	(0.007)	(0.009)	(0.009)
ACDE 95% CI	[0.018,0.046]	[0.017,0.046]	[0.003,0.038]	[0.003,0.038]
Intermediate Confounders	Yes	No	Yes	No
Select State FE	No	No	Yes	Yes
Observations	1092	1092	1092	1092

Notes: ACDE calculated using the methodology from Acharya et al. (2016a). RNV Signal is an indicator and quantifies the direct effect of receiving RNV's signal on Chávez's share over and above the mediated effect that could have happened through turnout. Columns 1 and 2 correspond to specifications without state indicators, while columns 3 and 4 include state indicators. Even-numbered columns assume that there are no intermediate confounders, they replicate the DAG from Figure 1.B.1. The odd-numbered columns correspond to the DAG in Figure 1.B.2, where corruption and the change in registered voters per parish correspond to the proposed intermediate confounders.



Notes: DAG representing the proposed mediated relationships between the variables in this setting with potential intermediate confounders. U_1 represents unobservables that would be akin to omitted variables. U_2 also represent unobservables related to the sequential ignorability assumption, which is non-refutable (Manski, 2007).

Figure 1.B.2: DAG with Intermediate Confounders

1.C Details of Heterogeneities Analysis

1.C.1 Demographic Summary Statistics and Variable Creation

For the variables that capture the materials of houses' outer walls, roofs and floors, I create a new variable that groups some of these materials under a classification that I call 'precarious'. For the case of outer walls, I group under 'precarious walls' all houses with walls made out of wood, trowelled and untrowelled adobe, and 'other' materials. For the case of 'precarious roofs' I group houses with roofs made out of asbestos, metal or 'other'. And for the case of 'precarious floor', I group houses with floors made out of dirt or 'other'.

To gain statistical power in the analyses, I aggregate these variables into indeces following the methodology from Anderson (2008). Namely, the indeces are the inverse correlation-weighted indeces of standardised versions of the variables chosen to comprise the index. Table 1.C.1 shows summary statistics of all these variables.

1.C.2 Tree-based Analysis

See Tables 1.C.2 and 1.C.3 for heterogeneous effects on all demographics.

Table 1.C.1: Summary Statistics of Demographics per Parish in Venezuela (1990)

	N	Mean	SD	Min	10th p.	25th p.	Median	75th p.	90th p.	Max
Proportion Men	1108	0.516	0.025	0.438	0.487	0.500	0.516	0.530	0.544	0.724
Unemployment Rate	1108	0.063	0.037	0.000	0.023	0.040	0.061	0.082	0.098	0.603
Unemployment Rate Men	1108	0.090	0.055	0.000	0.030	0.055	0.086	0.118	0.146	0.867
Unemployment Rate Women	1108	0.035	0.028	0.000	0.012	0.021	0.033	0.045	0.052	0.350
Illiteracy Rate	1108	0.219	0.101	0.028	0.100	0.139	0.212	0.278	0.342	0.686
Illiteracy Rate Men	1108	0.215	0.101	0.026	0.094	0.133	0.208	0.276	0.337	0.678
Illiteracy Rate Women	1108	0.223	0.102	0.030	0.103	0.145	0.214	0.278	0.352	0.695
Total Houses	1108	5805.610	12194.148	44.000	487.000	861.000	1784.500	4555.000	14503.000	95599.000
Prop. Occupied Houses	1108	0.814	0.104	0.263	0.688	0.767	0.836	0.887	0.918	0.979
Prop. Unoccupied Houses	1108	0.107	0.061	0.010	0.050	0.070	0.097	0.131	0.173	0.674
Prop. Occasional Houses	1108	0.063	0.073	0.000	0.012	0.021	0.040	0.075	0.137	0.587
Prop. Houses In Construction	1108	0.016	0.013	0.000	0.003	0.008	0.014	0.021	0.032	0.116
Occupants Per House	1108	5.218	0.491	3.600	4.700	4.950	5.200	5.460	5.760	8.280
Troweled Brick Outer Walls	1108	0.622	0.143	0.017	0.434	0.542	0.658	0.725	0.781	0.904
Untroweled Brick Outer Walls	1108	0.121	0.048	0.004	0.059	0.087	0.118	0.153	0.185	0.308
Concrete Outer Walls	1108	0.030	0.030	0.000	0.006	0.010	0.018	0.038	0.074	0.198
Wood Outer Walls	1108	0.005	0.014	0.000	0.001	0.001	0.002	0.005	0.012	0.163
Troweled Adobe Outer Walls	1108	0.086	0.086	0.000	0.010	0.024	0.063	0.122	0.183	0.578
Untroweled Adobe Outer Walls	1108	0.088	0.091	0.000	0.008	0.020	0.050	0.136	0.221	0.465
Other Material Outer Walls	1108	0.048	0.062	0.000	0.003	0.010	0.032	0.068	0.101	0.639
Precarious Outer Walls	1108	0.227	0.164	0.003	0.071	0.110	0.168	0.312	0.456	0.974
Precarious Wall Index	1108	-0.000	0.576	-0.855	-0.527	-0.331	-0.092	0.182	0.586	4.758
Concrete Roof	1108	0.165	0.179	0.000	0.021	0.038	0.084	0.254	0.408	0.933
Tiled Roof	1108	0.061	0.059	0.000	0.015	0.025	0.041	0.071	0.144	0.395
Asbestos Roof	1108	0.150	0.147	0.000	0.041	0.065	0.106	0.169	0.308	0.925
Metal Roof	1108	0.592	0.210	0.011	0.280	0.465	0.639	0.756	0.821	0.930
Other Roof	1108	0.033	0.066	0.000	0.007	0.010	0.015	0.026	0.066	0.730
Precarious Roof	1108	0.774	0.194	0.018	0.514	0.658	0.848	0.922	0.955	0.998
Precarious Roof Index	1108	-0.000	0.431	-1.561	-0.574	-0.230	0.122	0.260	0.418	1.295
Brick Floor	1108	0.149	0.170	0.000	0.012	0.028	0.072	0.221	0.386	0.957
Cement Floor	1108	0.735	0.153	0.041	0.559	0.674	0.769	0.838	0.885	0.954
Dirt Floor	1108	0.108	0.097	0.001	0.022	0.042	0.076	0.148	0.221	0.640
Other Floor	1108	0.008	0.063	0.000	0.001	0.001	0.002	0.002	0.007	0.855
Precarious Floor	1108	0.116	0.116	0.001	0.024	0.044	0.078	0.153	0.232	0.959
Precarious Floor Index	1108	-0.000	0.712	-0.615	-0.492	-0.385	-0.210	0.203	0.598	6.728
Poverty Index	1108	-0.000	0.354	-1.400	-0.365	-0.160	0.048	0.175	0.352	1.633

Notes: Summary statistics per parish using the 1990 census results. All data was digitised from OCEI (1993b, 1993c).

Table 1.C.2: Heterogeneous Effects: Quartiles Most and Least Affected by Exposure to RNV on Chavez's Share

	1Q		4Q		4Q - 1Q	
Variable	Mean	SD	Mean	SD	Difference	p-value
∆ Chávez's Share (Pred.)	-0.028	(0.007)	0.029	(0.014)	0.057***	(0.000)
RNV Signal (Indicator)	1.000	(0.000)	1.000	(0.000)	0.000	()
RNV Signal (Prop.)	0.927	(0.125)	0.920	(0.127)	-0.007	(0.677)
Chavez's Share 1998	0.565	(0.096)	0.576	(0.121)	0.011	(0.474)
Men (Prop.)	0.521	(0.028)	0.506	(0.023)	-0.015***	(0.000)
Parish Area (km^2)	629.767	(940.450)	342.139	(781.511)	-287.627**	(0.020)
Unemployment Rate	0.074	(0.044)	0.067	(0.045)	-0.007	(0.268)
Unemployment Rate Men	0.104	(0.061)	0.091	(0.059)	-0.012	(0.147)
Unemployment Rate Women	0.041	(0.038)	0.041	(0.027)	-0.000	(0.935)
Illiteracy Rate	0.202	(0.083)	0.167	(0.090)	-0.034***	(0.006)
Illiteracy Rate Men	0.199	(0.083)	0.163	(0.089)	-0.036***	(0.004)
Illiteracy Rate Women	0.205	(0.083)	0.172	(0.092)	-0.033***	(0.009)
Total Houses	7,643.305	(16,080.598)	13,358.680	(19,253.793)	5,715.375**	(0.024)
Occupied Houses (Prop.)	0.774	(0.122)	0.822	(0.104)	0.048***	(0.003)
Unoccupied Houses (Prop.)	0.112	(0.050)	0.097	(0.040)	-0.015**	(0.021)
Occasional Houses (Prop.)	0.094	(0.109)	0.067	(0.081)	-0.026*	(0.055)
Houses In Construction (Prop.)	0.020	(0.015)	0.013	(0.009)	-0.007***	(0.000)
Occupants Per House	5.117	(0.612)	4.997	(0.448)	-0.121	(0.115)
Troweled Brick Outer Walls (Prop.)	0.630	(0.109)	0.645	(0.119)	0.015	(0.351)
Untroweled Brick Outer Walls (Prop.)	0.130	(0.042)	0.120	(0.041)	-0.010	(0.104)
Concrete Outer Walls (Prop.)	0.035	(0.033)	0.047	(0.040)	0.012**	(0.026)
Wood Outer Walls (Prop.)	0.006	(0.009)	0.007	(0.008)	0.001	(0.454)
Troweled Adobe Outer Walls (Prop.)	0.072	(0.064)	0.058	(0.062)	-0.014	(0.109)
Untroweled Adobe Outer Walls (Prop.)	0.078	(0.083)	0.064	(0.085)	-0.014	(0.259)
Other Material Outer Walls (Prop.)	0.050	(0.044)	0.059	(0.039)	0.010*	(0.096)
Precarious Outer Walls (Prop.)	0.205	(0.126)	0.188	(0.128)	-0.017	(0.341)
Precarious Wall Index	-0.059	(0.375)	-0.072	(0.358)	-0.013	(0.805)
Concrete Roof (Prop.)	0.191	(0.180)	0.322	(0.236)	0.131***	(0.000)
Tiled Roof (Prop.)	0.051	(0.031)	0.060	(0.044)	0.009*	(0.082)
Asbestos Roof (Prop.)	0.152	(0.128)	0.118	(0.125)	-0.034*	(0.058)
Metal Roof (Prop.)	0.588	(0.195)	0.485	(0.213)	-0.103***	(0.000)
Other Roof (Prop.)	0.018	(0.025)	0.014	(0.013)	-0.004	(0.221)
Precarious Roof (Prop.)	0.758	(0.190)	0.618	(0.229)	-0.141***	(0.000)
Precarious Roof Index	-0.046	(0.404)	-0.340	(0.482)	-0.294***	(0.000)
Brick Floor (Prop.)	0.175	(0.179)	0.283	(0.211)	0.108***	(0.000)
Cement Floor (Prop.)	0.733	(0.152)	0.641	(0.170)	-0.091***	(0.000)
Dirt Floor (Prop.)	0.090	(0.065)	0.074	(0.074)	-0.017*	(0.091)
Other Floor (Prop.)	0.002	(0.004)	0.002	(0.003)	-0.000	(0.920)
Precarious Floor (Prop.)	0.093	(0.065)	0.076	(0.075)	-0.017*	(0.091)
Precarious Floor Index	-0.132	(0.341)	-0.220	(0.392)	-0.088*	(0.092)
Poverty Index	-0.002	(0.314)	-0.191	(0.403)	-0.188***	(0.000)
Observations	99		99		198	

Notes: Heterogeneous effects on Chávez's vote share along different along demographics calculated using random forests (Athey, Tibshirani, & Wager, 2019). Baseline model corresponds to the full specifications used for the results reported in Table 1.1.

Table 1.C.3: Heterogeneous Effects: Quartiles Most and Least Affected by Exposure to RNV on Turnout

	1Q		4Q		4Q - 1Q	
Variable	Mean	SD	Mean	SD	Difference	p-value
Δ Turnout (Pred.)	-0.016	(0.004)	0.005	(0.003)	0.021***	(0.000)
RNV Signal (Indicator)	1.000	(0.000)	1.000	(0.000)	0.000	()
RNV Signal (Prop.)	0.873	(0.147)	0.949	(0.112)	0.076***	(0.000)
Turnout Rate 1998	0.568	(0.064)	0.643	(0.039)	0.075***	(0.000)
Men (Prop.)	0.528	(0.020)	0.503	(0.018)	-0.025***	(0.000)
Parish Area (km^2)	1,251.688	(1,758.219)	81.818	(115.533)	-1,169.870***	(0.000)
Unemployment Rate	0.071	(0.079)	0.065	(0.025)	-0.006	(0.503)
Unemployment Rate Men	0.102	(0.110)	0.090	(0.039)	-0.012	(0.314)
Unemployment Rate Women	0.035	(0.043)	0.040	(0.015)	0.005	(0.294)
Illiteracy Rate	0.256	(0.088)	0.142	(0.065)	-0.114***	(0.000)
Illiteracy Rate Men	0.254	(0.088)	0.139	(0.067)	-0.115***	(0.000)
Illiteracy Rate Women	0.259	(0.090)	0.145	(0.063)	-0.114***	(0.000)
Total Houses	6,373.517	(14,757.073)	14,693.949	(23,047.236)	8,320.433***	(0.003)
Occupied Houses (Prop.)	0.783	(0.085)	0.819	(0.126)	0.036**	(0.020)
Unoccupied Houses (Prop.)	0.121	(0.046)	0.088	(0.035)	-0.033***	(0.000)
Occasional Houses (Prop.)	0.083	(0.071)	0.076	(0.111)	-0.008	(0.564)
Houses In Construction (Prop.)	0.013	(0.010)	0.018	(0.015)	0.005***	(0.010)
Occupants Per House	5.211	(0.452)	4.984	(0.453)	-0.227***	(0.001)
Troweled Brick Outer Walls (Prop.)	0.573	(0.115)	0.691	(0.090)	0.118***	(0.000)
Untroweled Brick Outer Walls (Prop.)	0.128	(0.046)	0.119	(0.039)	-0.010	(0.111)
Concrete Outer Walls (Prop.)	0.023	(0.022)	0.051	(0.039)	0.028***	(0.000)
Wood Outer Walls (Prop.)	0.005	(0.007)	0.007	(0.009)	0.001	(0.269)
Troweled Adobe Outer Walls (Prop.)	0.103	(0.070)	0.040	(0.038)	-0.063***	(0.000)
Untroweled Adobe Outer Walls (Prop.)	0.124	(0.098)	0.031	(0.038)	-0.093***	(0.000)
Other Material Outer Walls (Prop.)	0.044	(0.036)	0.062	(0.041)	0.018***	(0.001)
Precarious Outer Walls (Prop.)	0.276	(0.139)	0.139	(0.075)	-0.137***	(0.000)
Precarious Wall Index	0.105	(0.337)	-0.188	(0.322)	-0.293***	(0.000)
Concrete Roof (Prop.)	0.128	(0.127)	0.322	(0.228)	0.195***	(0.000)
Tiled Roof (Prop.)	0.052	(0.040)	0.062	(0.041)	0.010*	(0.084)
Asbestos Roof (Prop.)	0.135	(0.070)	0.176	(0.216)	0.041*	(0.075)
Metal Roof (Prop.)	0.662	(0.139)	0.428	(0.227)	-0.234***	(0.000)
Other Roof (Prop.)	0.024	(0.026)	0.012	(0.007)	-0.012***	(0.000)
Precarious Roof (Prop.)	0.821	(0.134)	0.616	(0.226)	-0.205***	(0.000)
Precarious Roof Index	0.059	(0.273)	-0.293	(0.540)	-0.353***	(0.000)
Brick Floor (Prop.)	0.110	(0.113)	0.296	(0.207)	0.186***	(0.000)
Cement Floor (Prop.)	0.760	(0.094)	0.651	(0.183)	-0.109***	(0.000)
Dirt Floor (Prop.)	0.127	(0.082)	0.052	(0.040)	-0.076***	(0.000)
Other Floor (Prop.)	0.003	(0.004)	0.002	(0.003)	-0.001	(0.204)
Precarious Floor (Prop.)	0.130	(0.082)	0.054	(0.041)	-0.076***	(0.000)
Precarious Floor Index	0.064	(0.428)	-0.336	(0.217)	-0.400***	(0.000)
Poverty Index	0.074	(0.178)	-0.153	(0.435)	-0.227***	(0.000)
Observations	99		99		198	

Notes: Heterogeneous effects on turnout along different along demographics calculated using random forests (Athey et al., 2019). Baseline model corresponds to the full specifications used for the results reported in Table 1.1.

1.D Text Analyses

1.D.1 Pre-processing Data

I scrape the transcripts of all *Aló Presidente* broadcasts and all speeches that Chávez gave as a president and as a candidate from the todochavez.gob.ve website. I process these documents in Python. First, I remove all punctuation marks, drop Spanish stopwords, drop words with fewer than 3 letters, drop numbers, lemmatise the text keeping nouns and adjectives, and manually remove words that are repeated unnaturally often.[46] I save the lemmatised version of each speech as a separate document and use them as inputs for the rest of the analyses.

1.D.2 Data Analyses

Term Frequency-Inverse Document Frequency (tf-idf) Scores. Intuitively, these numbers denote the relative importance of specific tokens along a collection of documents. Tokens that are more common in several documents are penalised. If I consider the collection of speeches as a time-series, tokens that are common across all cross-sections will be penalised, whereas tokens that are unique to each document (speech) will receive higher values. I compute the tf-idf scores using TfidfVectorizer from the scikit-learn library. I plot the tf-idf scores for the terms 'constitutional' and 'election' in Figure 1.3.

Personalisation. For this analysis, I adapt the methodology from Van Aelst et al. (2012) to measure the personalisation of each speech or broadcast. Personalisation is measured across three dimensions: 1) visibility, 2) mentions of Chávez's characteristics in personal rather than political contexts, and 3) mentions of Chávez's personal life. These three dimensions are split into 18 questions that I respond using OpenAI's gpt-4o.

As an input, I use the lemmatised version of each speech or broadcast, which I separate into 50,000-word chunks. I send these chunks to the model using the prompts outlined below. I save the output as independent txt files, which I analyse separately. I clean the output in Python and generate csv files that I then import into Stata. In Stata, I recode questions 3.1, 3.3, 3.5, 3.7, 3.9, and 3.11 multiplying the response by -1. I aggregate the responses under sections 3 and 4 by calculating their sum, and I recentre section 3 around zero by adding 6 to the sum. I repeat this exercise for all speeches and *Aló Presidente* broadcasts. Then I calculate the mean values of section 2, 3, and 4 for the speeches, and use these mean values to find the ratio of personalisation for each *Aló Presidente* broadcast.

[46] These words include 'presidente', 'chávez', 'venezuela', 'país', gobierno', 'venezolano', bueno', 'venezuelo', 'mérido', 'discurso', 'alocución', 'candidato'.

The aggregate measure shown in Figure 1.6 corresponds to the average of the three ratios.

Prompt 1:

"Here is a transcript of a broadcast of Alo Presidente. Respond to the following questions with numbers and nothing else: 1) how many times are political parties or politicans mentioned? 2) How many times is Hugo Chavez mentioned? 3.1 Is the characteristic of 'competence' mentioned within the text in a political context? For example: the leader does not understand the office he or she is responsible for. (0 = no; 1 = yes). 3.2 Is the characteristic of 'competence' mentioned within the text in a personal context? For example: the leader is a poor mother or father (0 = no; 1 = yes) 3.3 Is the characteristic of 'leadership' mentioned within the text in a political context? For example: the leader failed to rally his or her party behind him or her. (0 = no; 1 = yes) 3.4 Is the characteristic of 'leadership' mentioned within the text in a personal context? For example: was the leader seen as a natural person in command in his/her youth by classmates. (0 = no; 1 = yes) 3.5 Is the characteristic of 'credibility' mentioned within the text in a political context? For example: a broken promise by the candidate in the previous elections, say on lower taxes. (0 = no; 1 = yes) 3.6 Is the characteristic of 'credibility' mentioned within the text in a personal context? For example: the leader is criticized by a family member for not keeping his or her promises to spend more time with his/her family. (0 = no; 1 = yes)"

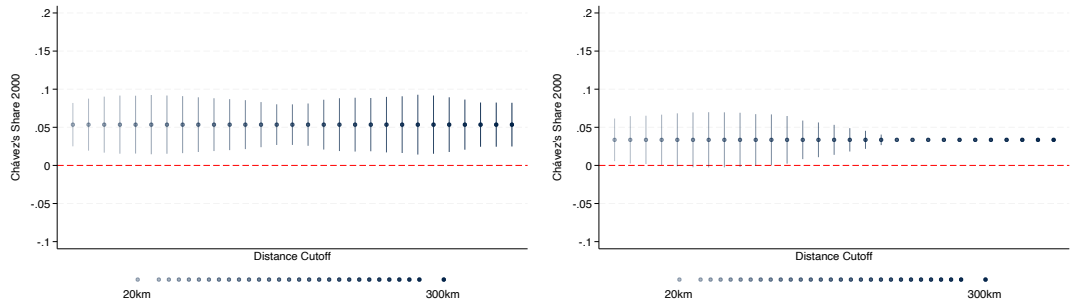
Prompt 2:

"Here is a transcript of a broadcast of Alo Presidente. Respond to the following questions with numbers and nothing else, printing the number of the question in each row e.g., '3.7)':
3.7) Is the characteristic of 'morality' mentioned within the text in a political context? For example: an investigation against the leader for accepting bribes or undermining the career of a rival. (0 = no; 1 = yes)
3.8) Is the characteristic of 'morality' mentioned within the text in a personal context? For example: the leader was caught cheating on his or her spouse. (0 = no; 1 = yes)
3.9) Is the characteristic of 'rhetorical skills' mentioned within the text in a political context? For example: a reference to a great speech by the leader in parliament. (0 = no; 1 = yes)
3.10) Is the characteristic of 'rhetorical skills' mentioned within the text in a personal context? For example: a reference to a great speech by the leader in a private ceremony or to one made before he or she entered politics. (0 = no; 1 = yes)
3.11) Is the characteristic of 'appearance' mentioned within the text in a political context? For example: a reference to the 'presidential appearance' of the candidate. (0 = no; 1 = yes)
3.12) Is the characteristic of 'appearance' mentioned within the text in a personal context? For example: a reference to the past of the leader as a winner of a beauty pageant. (0 = no; 1 = yes).
Does the text contain references to one of these indicators with respect to Hugo Chavez:
4.1) Family life. This includes family relationships and all aspects of domestic life. (0 = no; 1 = yes)
4.2) Past life or upbringing. This includes all biographical information. (0 = no; 1 = yes)
4.3) Leisure time. This includes all information on hobbies, vacations, and recreational activities. (0 = no; 1 = yes)
4.4 Love life. This includes all information on sexual relationships, marriage and divorce. (0 = no; 1 = yes)"

1.E Robustness

1.E.1 Conley Standard Errors

I test the robustness of my results to calculating Conley (1999) standard errors. I implement this in Stata via the ols_spatial_HAC command (Hsiang, 2010). I probe my results by setting a range of distance cutoffs within which errors are allowed to correlate in 10km increments from 20km to 300km. Results are presented in Figure 1.E.1. Panel (a) corresponds to the model without state indicators, and Panel (b) corresponds to the model that includes state indicators. Results from the former are fully robust, and results from the latter become weaker in the range of 80-100 km remaining significant at 10%.



(a) Without State Indicators (b) With State Indicators

Notes: Robustness of main results to spatial HAC adjusted errors along a range of distances. Computations were carried out in Stata using the ols_spatial_HAC command (Hsiang, 2010).

Figure 1.E.1: Robustness to Spatial HAC Adjusted Errors

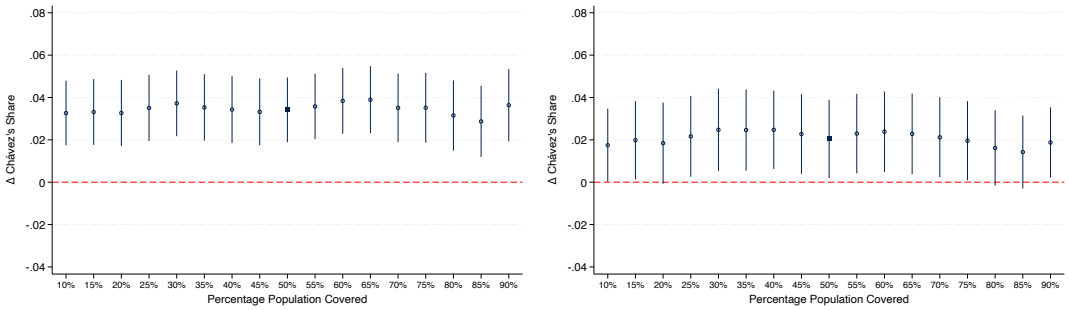
1.E.2 Family-wise Error Rates

In order to correct for Family-Wise Error Rates (FWER), I use the Westfall and Young (1993) algorithm. This algorithm predates the commonly used Romano and Wolf (2005a, 2005b) resampling method and is more flexible, albeit at the cost of computational power and the subset pivotality assumption. This assumption is a property of the distribution of p-values of our hypotheses (P). The assumption holds when 'the joint distribution of the subvector $\{P_i; i \in K\}$ is identical under the restrictions $\bigcap_{i \in K} H_{0i}$ and H_0^C, for all subsets $K = \{i_1, ..., i_j\}$ of true null hypotheses', where H_0^C denotes the complete null hypothesis (i.e., where all null hypotheses H_i are true). In this setting, all my null hypotheses are of the shape $H_0 : \beta = 0$, where β refers to any estimated coefficient. Each hypothesis is tested with a univariate t-test. If we assume that any arbitrary subset of the hypotheses are true, the joint distribution of the t-statistics of those true hypotheses will not depend on the values of the remaining estimates. The joint distribution of those t-statistics will

not change regardless of whether all hypotheses are true, or whether only the hypotheses associated with those t-statistics are true. In particular, the joint distribution of those t-statistics will depend on my sample size (i.e., degrees of freedom) and the correlation between the estimates for which the hypotheses are true (Siddiqui, 1967). Thus, subset pivotality holds in this setting.[47] Table 1.A.3 includes the Westfall-Young corrected p-values in curly brackets. All results are robust to the implementation of this stricter algorithm.

1.E.3 Changing Percentage of Population Covered

Here I check for robustness in my main results to changes in the percentage of the population being covered while keeping the field strength threshold fixed. Figure 1.E.2 shows the size of the coefficient for different definitions of the indicator using 5 percentage point increases on the population being covered. Panel (a) corresponds to the specification without state indicators and Panel (b) to the stricter specification that includes full state indicators. From these figures we can conclude that the results seem robust to modifying the definition of the indicator. The reason why the point estimates look rather flat while the population threshold is varying considerably is related to the coarseness of the indicator (see Figure 1.A.1). Regardless of the approach, the most evident feature of the figure is that movements along the population axis do not impact the coverage mean as much as movements along the field strength axis.



(a) Without State Indicators (b) With State Indicators

Notes: Robustness of main results to changes along the population thresholds of the main independent variable.

Figure 1.E.2: Robustness to Different Percentages of Population Being Covered by RNV's Signal

[47] This is similar to example 2.1 from Westfall and Young (1993).

1.E.4 Changing Signal Strength Threshold

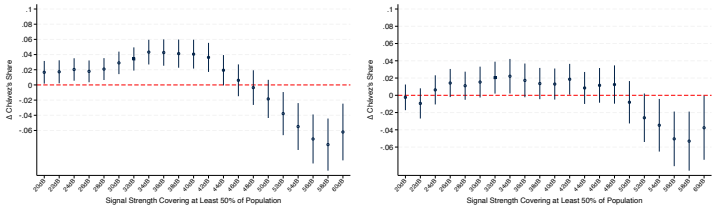
This section turns to checking for robustness in the results from section 1.4.1 by changing the field strength definition in the RNV signal indicator. Changes along this dimension should be interpreted carefully given two important pitfalls. Namely, that even though increasing this threshold could allow us to 'zoom into' parishes where RNV's signal is the strongest (coding with 0 parishes where the signal is not as strong or simply not there), empirically we cannot distinguish these results from two scenarios: (i) a scenario where we could be imposing unnecessarily conservative constraints on the field strength measure, making the analysis unrealistic; or, (ii) a scenario where we could end up measuring exposure to a different signal, given that signals with a higher frequency need a higher field strength to be detected. Moreover, increasing the field strength greatly reduces the number of parishes that are coded as receiving RNV's signal, which translates into reduced statistical power.

Figure 1.E.3 shows the coefficients associated with reception of RNV when the definition of the indicator is changing in increases of 2 dBμV/m. Panel (a) and (b) show the results for specifications without and with state indicators. In both cases we can conclude that a positive and statistically significant effect is found for field strengths in the vicinity of 32 dBμV/m. In fact, if anything, 32 dBμV/m seems to be the upper limit at which this positive effect is found. This is aligned with how conservative the calculations for determining the 32 dBμV/m threshold were. For instance, with a correction for building penetration 1 SD lower than what was adopted, the threshold would have been 28 dBμV/m. For higher measures of field strength, there is a null effect, as expected.

Panels (c) and (d), and (e) and (f) in Figure 1.E.3 show the coefficients associated with exposure to RNV defined over the same range of field strength values when 30% and 70% of the population is covered, respectively. These additional figures confirm that results are not that sensitive to changes along the population dimension.

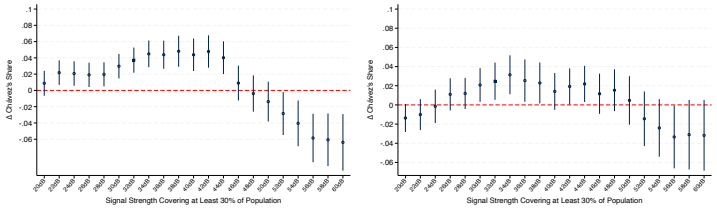
Even though some of these models find statistically significant point estimates for extreme field strength values, only around 5% of the sample is being coded as receiving RNV's signal in these extreme cases. This should raise concerns in terms of statistical power. The minimum detectable effect[48] (MDE) for the main definition of the signal indicator with 80% power at 5% significance is 0.021 SD. In contrast, for 50% of the population receiving

[48] The MDE is calculated with the following formula: $MDE = (t_{1-\kappa} + t_{\alpha/2})\sqrt{\frac{1}{P(1-P)} \times \frac{\sigma^2}{N}}$, where the values t are the corresponding critical values of the Student's t distribution at values $(1-\kappa)$ i.e., power (0.80), and $\alpha/2$ i.e., significance (0.05) divided by two; P is the proportion of units in the sample receiving a 'treatment', σ^2 corresponds to the variance of the outcome (in this case the Chávez's vote share in 2000, i.e., 0.015), and N, the number of observations (1,104). This formula follows the notation and explanation in Gupta and Kopper (2021).



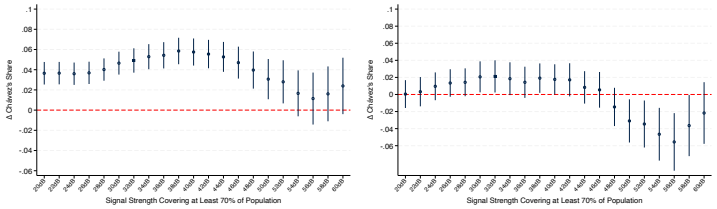
(a) Without State Indicators (50% Population)

(b) With State Indicators (50% Population)



(c) Without State Indicators (30% Population)

(d) With State Indicators (30% Population)



(e) Without State Indicators (70% Population)

(f) With State Indicators (70% Population)

Notes: Robustness of main results to changes along the median field strength threshold associated with the main independent variable. Panels (a) and (b) use the 50% population threshold used throughout the paper. Panels (c) and (d) reduce this threshold to 30%. Panels (e) and (f) increase it to 70%.

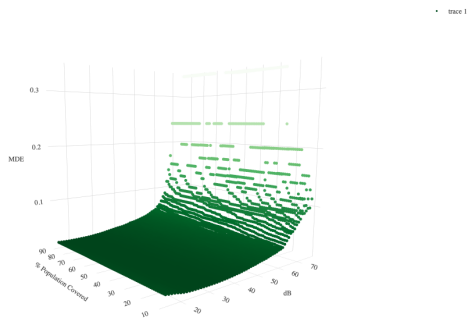
Figure 1.E.3: Robustness to Different Measures of Signal Strength

a signal with a field strength of at least 60 dBμV/m, the MDE is 0.048 SD – an effect more than 4 times as large. Figure 1.E.4 is a 3D plot showing the MDE for all the combinations of field strength and population coverage in the dataset. For argument's sake, the coefficient associated with exposure to RNV using the definition of 50% of the population receiving 32 dBμV/m has an effect size of 0.204 SD (95% C.I.: 0.080, 0.328), while the coefficient associated to the indicator defined over 60 dBμV/m has an effect size of -0.147 SD (95% C.I.: -0.043, 0.140).[49] In the case of the latter, the 95% confidence intervals include both zero and the MDE.

1.E.5 Percentage of Covered Area, Rather than Population

A commonly used approach to measure exposure in studies of mass media consumption is defining coverage over the area of a geographic unit, rather than its population. While

[49] Calculated with Stata's package ESIZEREG (Linden, 2019).



Notes: Minimum-Detectable Effect under different definitions of the main independent variable. The vertical axis represents the MDE, and the horizontal axes the two dimensions that could be modified in the definition: the population threshold, and the signal strength threshold. The MDE is calculated with the following formula: $MDE = (t_{1-\kappa} + t_{\alpha/2})\sqrt{\frac{1}{P(1-P)} \times \frac{\sigma^2}{N}}$, where the values t are the corresponding critical values of the Student's t distribution at values $(1-\kappa)$ i.e., power (0.80), and $\alpha/2$ i.e., significance (0.05) divided by two; P is the proportion of units in the sample receiving a 'treatment', σ^2 corresponds to the variance of the outcome (in this case the Chávez's vote share in 2000, i.e., 0.015), and N, the number of observations (1,104). This formula follows the notation and explanation in Gupta and Kopper (2021).

Figure 1.E.4: MDE for All RNV Signal Indicators

both approaches might have similar means, the area approach is not appropriate for this setting since the units of analysis i.e., parishes, are much larger than in settings that adopt the area approach (e.g., Blouin & Mukand, 2019). The area approach weighs each squared kilometre equally when determining whether a larger area is receiving a radio signal. On the contrary, the population approach weighs a square kilometre by the population that lives in it, which might not be perfect in case there is measurement error.

However, in the worst case scenario, following the area approach we could end up attributing a very strong reception to an area where effectively no one lives, and this area would weigh the same in the definition of our independent variable as the squared kilometre with the highest population density in, say, Caracas. Following the area approach simplifies calculations, and could be used in setting where the unit of analysis is small. However, in settings such as this one, or in cross-country analyses such as in Guriev et al. (2021) or Mensah et al. (2021), it makes sense to follow the population approach.

I check for robustness of this article's findings to changing the indicator to take a value of 1 when a certain area of Venezuela's parishes receive a strong enough signal. I adopt a threshold of 50% as the minimum area that a parish should have covered with a signal of at least 32 dBμV/m in order to be counted as receiving RNV's signal. The results replicate those presented in Section 1.4.1.

Table 1.E.1: Effect of RNV's Signal on Chávez's Share 1998-2000 Using Area Approach.

	Δ Turnout		Δ Chávez's Share	
	(1)	(2)	(3)	(4)
RNV Signal (Proportion Area)	-0.020***	-0.017**	0.021***	0.019**
	(0.006)	(0.007)	(0.008)	(0.009)
Observations	1096	1096	1092	1092
Control Mean	0.594	0.594	0.593	0.593
Covariates	Yes	Yes	Yes	Yes
Demographics	Yes	Yes	Yes	Yes
State Ind.	No	Yes	No	Yes
R^2	0.338	0.447	0.514	0.682

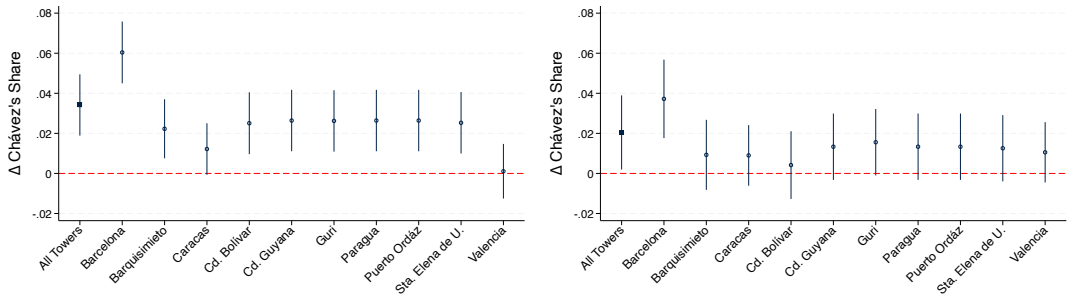
Notes: Effect of RNV's Signal on Chávez's share between 1998 and 2000. RNV's Signal is a dummy variable that takes the value of 1 when at least 50% of a parish's area is covered with a signal with a field strength of at least 32 dBμ/m. Luminosity is the average luminosity in the parish in 1998. Δ Voters is the change in registered voters at each parish between 1998 and 2000. Possible corruption is a dummy variable that takes a value of 1 when the number of null votes in a parish exceeds 5% and most of the polling stations in the parish were electronic. Population is the parish's population as a share of the country's population. No Signal (State) is an indicator that takes a value of 1 when a parish is located in a state where no parish receives RNV's signal. No Signal (Mun) is an indicator that takes a value of 1 when a parish is located in a municipality where no parish receives RNV's signal. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

1.E.6 Changing Tower Network Composition

Figure 1.E.5 shows the estimates associated with reception of RNV's signal when the field strength simulation is carried out dropping at a time each of the 10 towers considered part of the network. It is clear that dropping one tower in a region where there are several towers and where there is not as much topographic variation as in other places (as exemplified in Panel (b) Figure 1.4) will not impact the estimates considerably. This is the case for the towers located in Ciudad Bolívar, Ciudad Guyana, Puerto Ordáz, Guri and Paragua (see Panel (a) Figure 1.4). As for the tower located in Santa Elena de Ushuaia, being located in the southern part of the country, its radio signal reaches mainly 3 municipalities located in the state of Bolívar: the southern parts of the Municipalities of Sucre and Angostura, and most of the Municipality of Gran Sabana. These municipalities as a whole do not seem to influence the outcome of interest.

This is not the case, however, for the towers located in Barcelona, Barquisimieto, Caracas and Valencia, as removing each of these towers influences considerably the point estimates of the regressions. In the case of Barcelona, it increases the point estimate, while for the case of the other three towers it reduces the magnitude of the point estimate until it is statistically indistinguishable from zero. For the case of the municipalities affected by the tower in Barcelona, the reason the coefficient is increasing in magnitude is because the parishes that are no longer being coded as receiving RNV's signal had a lower than

average change in support for Chávez. The opposite is the case for the parishes reached by the towers located in Barquisimieto, Caracas and Valencia. As is evident in panel (b), the stricter specification runs into trouble given a lack of power even if point estimates are fairly consistent across the board.



(a) Without State Indicators (b) With State Indicators

Notes: Robustness of main results under different networks constructed leaving one of the towers out at a time. Horizontal axes denote the name of the tower that is being dropped from the network.

Figure 1.E.5: Robustness to Dropping One Tower From Simulation

1.E.7 Exposure to a Placebo Network

I perform this test to assess whether the results observed so far are but an artefact of the simulation process itself. This test is performed by creating a placebo network and running the models from section 1.4.1. For the placebo network, I follow Blouin and Mukand (2019) and Björkegren (2019), and create a placebo network by 'placing' 10 radio towers in the 10 tallest peaks of Venezuela. I then simulate the field strength of this placebo network under the same set of assumptions used for RNV's simulations. Figure 1.A.4 shows the result of this simulation along with the location of these 10 fictitious towers. Table 1.E.2 provides more details on the tallest peaks of Venezuela, but it is worth remarking that all of them are located in the Andes, specifically in the state of Mérida.

As in the case for RNV's signal, a parish is coded as receiving the placebo signal when more than half of the population living in the parish is receiving a signal with a field strength of at least 32 dBμV/m. All controls included in the regressions are the same as those in Table 1.1, with the exception of the variables capturing the distance from the centroid of each parish to the nearest RNV tower. Naturally, in this case I use the distance to the nearest placebo tower. Results are shown in Table 1.E.3, and they confirm that electoral outcomes are not statistically related to reception of this placebo signal. This suggests that the results found so far are not spuriously generated by the simulation process itself. Rather, it adds credibility to the claim that the results found using data generated when

simulating RNV's network indeed capture a relationship between reception of RNV's signal and the electoral outcomes being analysed.

Table 1.E.2: Ten Highest Peaks in Venezuela – Location of Placebo Towers

Mountain	Latitude	Longitude	Altitude (m)	State	Mountain Range
Pico Bolívar	8.541112	-71.04639	4978	Mérida	Sierra Nevada de Mérida
Pico Humboldt	8.549167	-70.99583	4942	Mérida	Sierra Nevada de Mérida
Pico La Concha	8.554723	-71.02556	4922	Mérida	Sierra Nevada de Mérida
Pico Bonpland	8.540833	-71.00222	4883	Mérida	Sierra Nevada de Mérida
Pico Piedras Blancas	8.85889	-70.95166	4748	Mérida	Sierra La Culata
Pico El León	8.511389	-71.10056	4740	Mérida	Sierra Nevada de Mérida
Pico El Toro	8.529444	-71.09139	4727	Mérida	Sierra Nevada de Mérida
Pico Los Nevados	8.864445	-70.92222	4700	Mérida	Sierra La Culata
Pico Pan de Azúcar	8.810278	-71.00083	4643	Mérida	Sierra La Culata
Pico Mucuñuque	8.756667	-70.80251	4609	Mérida	Sierra de Santo Domingo

Notes: Name, coordinates, altitude, and location of the 10 highest peaks of Venezuela, used in the construction of a placebo radio network. The simulations to generate the network used the same assumptions as those used to simulate RNV's network. Latitude, longitude and altitude values were retrieved from Google Earth.

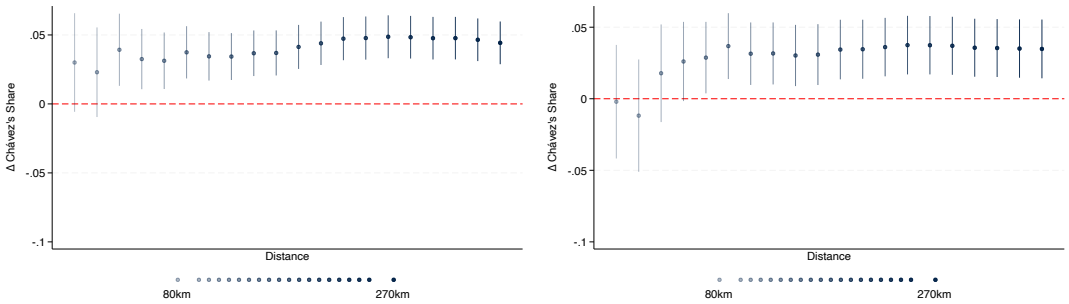
Table 1.E.3: Test Against Reception of Placebo Network

	Δ Turnout	Δ Chávez's Share
	(1)	(2)
Receives Placebo Signal	-0.011	0.008
	(0.009)	(0.011)
Observations	1096	1092
Control Mean	0.599	0.626
Covariates	Yes	Yes
Demographics	Yes	Yes
State Ind.	Yes	Yes
R^2	0.445	0.685

Notes: Placebo Signal is an indicator that takes the value of 1 when at least 50% of a parish's population receives a signal with a field strength of at least 32 dBμV/m from a placebo network comprised of 10 fictitious radio towers placed in the 10 highest peaks of Venezuela. The covariates in all specifications follow the same structure as those in Table 1.1. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

1.E.8 Subsampling

I condition the sample to parishes located within a specific distance to their nearest RNV tower. The objective of this exercise is pushing the identification of the effect to be within the same states. Figure 1.E.6 below shows the magnitude of the coefficient associated with reception of RNV for a series of subsamples defined over a wide range of distances to the nearest RNV tower for the two main specifications of the paper. The values of the distances range from 80km to 270km. These values correspond to the 30th and 80th percentiles of the distribution of distances.



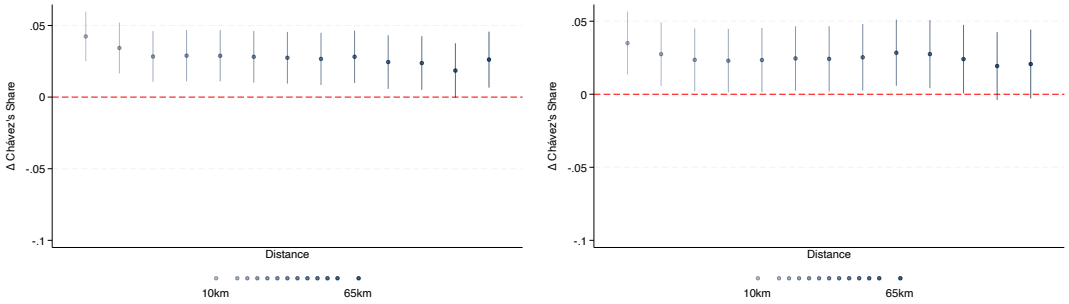
(a) Without State Indicators (b) With State Indicators

Notes: Robustness of main results to limiting the distance to the nearest towers. Results shown in figure correspond to the values ranging from the 30th to the 80th percentiles of the distance distributions.

Figure 1.E.6: Subsamples by Distance to Nearest RNV Tower

The coefficients from both the specification with and without state indicators remain consistent and positive throughout. An obvious and warranted concern with these calculations relates to statistical power. However, given the effect sizes found in previous sections, the most conservative rules explored here should still be able to detect effects of these sizes. For instance, for the subsample located within 270km of the nearest RNV tower, the MDE is 0.019SD.

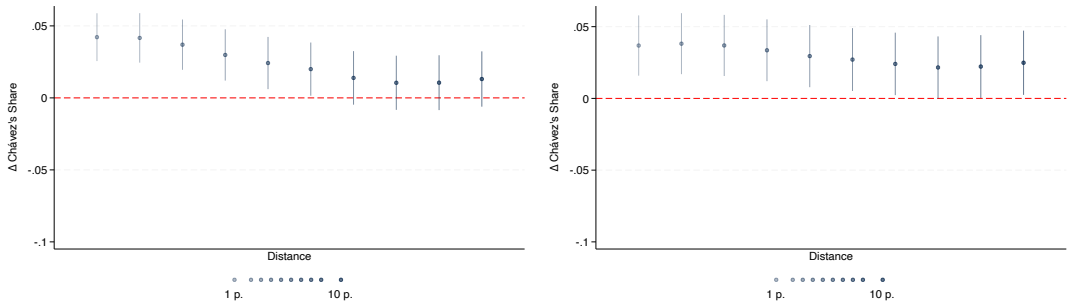
As a complement to this analysis, I create 'doughnuts' around the radio towers using ranges of distances (Figure 1.E.7) or distance percentiles (Figure 1.E.8). Results are enchanged to these alternative specifications.



(a) Without State Indicators (b) With State Indicators

Notes: Robustness of main results to creating doughnuts around the nearest towers. Results shown in figure correspond to the values obtained after dropping observations outside the range of distances capped at the top and bottom by 10km to 65km.

Figure 1.E.7: Subsamples by Distance to Nearest RNV Tower: Doughnuts on Distances



(a) Without State Indicators

(b) With State Indicators

Notes: Robustness of main results to creating doughnuts around the nearest towers. Results shown in figure correspond to the values obtained after dropping observations from the top and bottom 1p. to the 10p. of the distance distributions.

Figure 1.E.8: Subsamples by Distance to Nearest RNV Tower: Doughnuts on Distance Percentiles

1.E.9 Sharing a Border with Colombia

There is a concern that results could be driven by populations living in parishes that are close to Colombia. I include an indicator for being in a state that shares a border with Colombia in Table 1.E.4. Results remain unchanged.

Table 1.E.4: Main Results with Indicator for Sharing a Border with Colombia

	Δ Turnout 2000	Δ Chávez's Share 2000
	(1)	(2)
RNV Signal (Proportion)	-0.030***	0.034***
	(0.009)	(0.012)
State Borders Colombia	-0.037	0.018
	(0.026)	(0.034)
Observations	1096	1092
Control Mean	0.599	0.583
Covariates	Yes	Yes
Demographics	Yes	Yes
State Indicators	Yes	Yes
F-Statistic	30.515	70.614
R^2	0.449	0.683

Notes: Robustness of the main results accounting for states that share a border with Colombia. Robust standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

1.E.10 Correct for Non-Random Response to Random Shocks

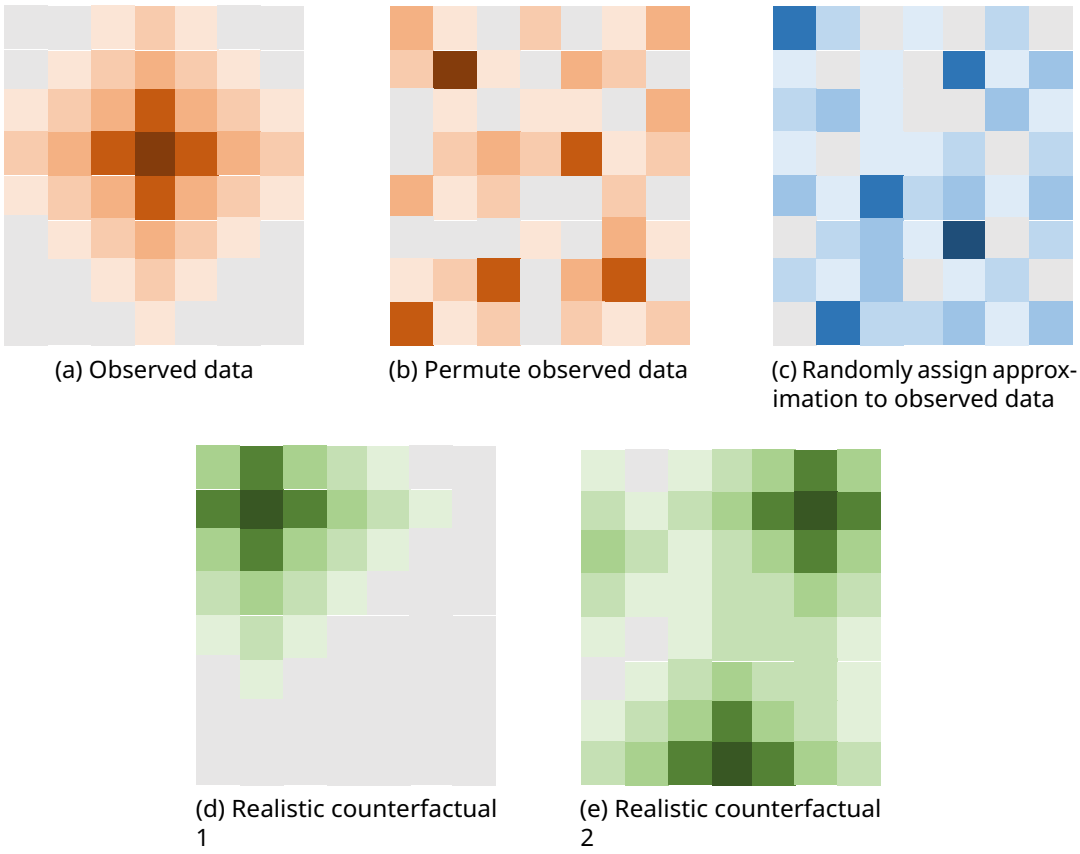
Borrowing notation from Borusyak and Hull (2021), we are interested in estimating the effect that a variable x_i has on outcome y_i for unit i.[50] To estimate this effect we may use a composite instrument $z_i = f_i(g; w)$.[51] $f_i(\cdot)$ is a non-stochastic function defined over a vector of exogenous shocks (g), and a set of predetermined variables (w). Borusyak and Hull (2021) define an expected instrument as $\mu_i = \mathbb{E}[f_i(g; w)|w]$, i.e., the average of z_i over different realisations of the shocks g, conditional on w. In my setting, we assume the shock driving the identification strategy is topographic ruggedness. The set of predetermined variables, on the other hand, corresponds to idiosyncratic aspects of the setting such as the location of the radio towers relative to the different parishes in the sample, or the population of each parish.

In order to compute the expected instrument, we need to define a shock assignment process as the conditional distribution of $g|w$, i.e., $G(g|w)$. If $G(\cdot)$ is known, then we can calculate $\mu_i = \int f_i(\gamma; w) dG(\gamma|w)$. In non-experimental settings, specifying $G(\cdot)$ is difficult. Thus, I explore three alternative methods for specifying $G(\cdot)$ and calculating candidate instruments $m_i(w)$ in order to correct for a potential omitted variable bias. First, following the advice set in Appendix D.7 of Borusyak and Hull (2021) I perform a permutation of the realised values of topographic ruggedness. The second approach involves approximating the empirical distribution of my measures of ruggedness (the exogenous process) with independent draws from several candidate theoretical distributions. The last approach relies on simulating realistic counterfactual topographies using Perlin (1985) noise – a type of noise used to create procedurally generated textures. I motivate using each of these approaches in Figure 1.E.9.

Each panel of Figure 1.E.9 shows an 8x8 grid that represents a specific geographic extent. Panel (a) shows the observed data where darker colours represent higher values of an observed metric (e.g., higher topographic ruggedness). This panel emphasises a natural geo-spatial relationship in the observed data. The first approach to offering a counterfactual realisation of g is represented in Panel (b), where the observed values are randomly permuted across the geographic extent. Note this keeps the overall distribution of g constant – represented by maintaining the number of squares of each colour consistent from Panel (a) to (b). Panel (c) represents approximating the observed data via a theoretical distribution, and randomly assigning draws from this distribution across the geographic extent. Note that in this case the number of squares of a certain gradient will not match

[50] In my setting $i = p$.

[51] It may well be the case that $x_i = z_i$, and the relationship between x_i and y_i reflects the relationship between a composite variable and an outcome of interest.



(a) Observed data

(b) Permute observed data

(c) Randomly assign approximation to observed data

(d) Realistic counterfactual 1

(e) Realistic counterfactual 2

Notes: Figure illustrates the three approaches used to define the shock assignment process when correcting for a potential OVB akin to a non-random response to exogenous shocks. Figure (a) represents the observed data, which with a natural geo-spatial relation. Panel (b) illustrates permuting the observed values of ruggedness across the geographic extent. Panel (c) represents randomly assigning independent draws from a candidate theoretical distribution that approximates the distribution of values from (a). Panels (d) and (e) represent counterfactual topographies. For this exercise, I attempt to implement the correction following the three approaches.

Figure 1.E.9: Approaches to Correct for Non-Random Response to Random Shocks

perfectly those in Panel (a), but the distributions will be close. Panels (d) and (e) represent realistic counterfactual realisations of g. These representations are realistic in a sense that the natural geo-spatial relationship in the data is preserved, even if the distribution of the counterfactual data is not necessarily the same as the one from (a). My preferred approach is the third one.

For these analyses I mainly rely on two topographic ruggedness variables from Nunn and Puga (2012): the Topographic Ruggedness Index (TRI) and a measure of local slope variability. Both of these measures are aggregated at the level of the parish with weights that correspond to each pixel's area following Nunn and Puga's advice. Taking the mean of a measure of dispersion as a way of aggregating data is committing an ecological fallacy. Thus, I incorporate a third (non-parametric) measure of ruggedness: the median absolute deviation from the median (MAD) of the TRI for each parish. I calculate the MAD TRI for

each parish p using the following formula:

$$\text{MAD TRI}_p = \sqrt{\sum_j \left(\tilde{\text{TRI}}_p - \text{TRI}_j\right)^2},$$

where $\tilde{\text{TRI}}_p$ is the median of the TRI for each parish, and j corresponds to each point feature within each parish's polygon. Since all my analyses are carried at the level of the parish, I consider this measure superior to the TRI and Slope given the strong parametric assumption that went into the computation of the latter two.

Permuting observed topographic ruggedness. This is an approach suggested by Borusyak and Hull (2021) in their Appendix D.7. I implement it by reshuffling my three (observed) topographic ruggedness variables 1,000 times.[52] I then find the best fitting third degree polynomial on the proportion of people in the parish receiving a signal of at least 32 dBμV/m via LASSO for each of my three measures of ruggedness.[53] I then take these linear models and use the estimated coefficients to predict counterfactual values of the proportion of people receiving RNV by using the reshuffled values of ruggedness. This yields 1,000 values of $f_p(g;w)|w$ which I then average and use as instruments. Results from this approach are reported in Table 1.E.5.

Table 1.E.5: IV Regressions Correcting for Non-Random Response to Random Shocks: Permuted ruggedness

	Chávez's Share 2000					
	(1)	(2)	(3)	(4)	(5)	(6)
RNV Coverage	0.025*	0.025*	0.025*	0.025*	0.025*	0.025*
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Expected Exposure (TRI MAD)		0.063*				
		(0.033)				
Expected Exposure (TRI)				0.065*		
				(0.034)		
Expected Exposure (Slope)						0.065*
						(0.034)
Observations	988	988	988	988	988	988
Estimation	2SLS	OLS	2SLS	OLS	2SLS	OLS
R^2	0.652	0.654	0.652	0.654	0.652	0.654

Notes: Regressions showing correction for a potential omitted variable bias akin to a non-random response to random shocks. I define the shock-assignment process as permutations of the observed values of ruggedness. Odd-numbered columns show the 2SLS correction, and even-numbered columns show their OLS counterpart. For 2SLS specifications, the instrument is the re-centred measure of exposure to RNV. Sample is restricted to match the sample of Table 1.E.8.

[52] I implement this in Stata by generating random draws from a uniform distribution, sorting the data according to these values, and assigning observations a new (counterfactual) ID according to their ranking. I repeat this process 1,000 times, and reassign the observed values of ruggedness by these counterfactual IDs.

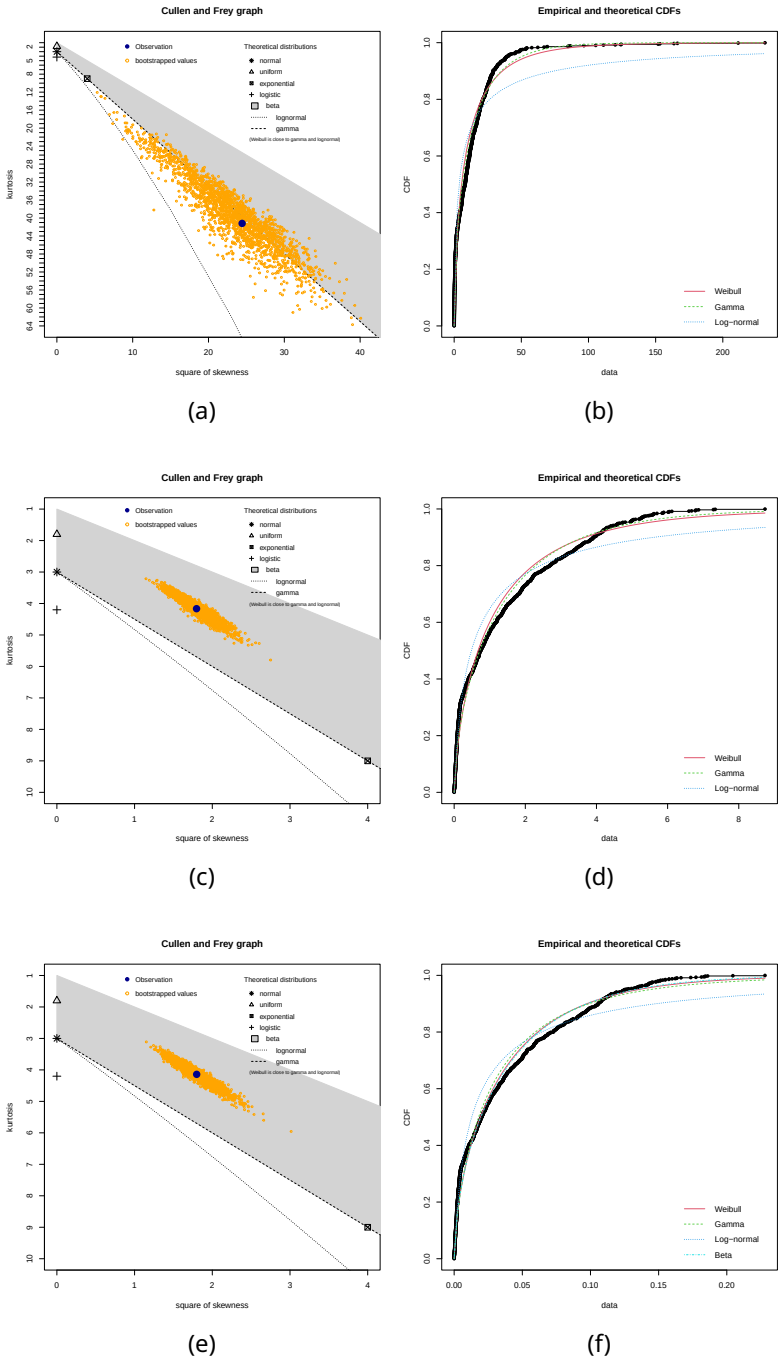
[53] I use the package lasso2 in Stata. The covariates I include are the distances to the three nearest AM towers, the longitude and latitude of each parish, my measure of topographic ruggedness, and the population of each parish. Naturally, I also include all of these variables raised to the second and third powers. To keep the model sparse I target the extended bayesian information criterion.

Theoretical approximation of empirical values of ruggedness. For this method I take a data-driven approach and find the theoretical distributions that best fit the empirical distribution of each measure of ruggedness in my sample.[54] The Cullen and Frey graphs for my three measures are shown in Panels (a), (c) and (e) of Figure 1.E.10.[55] From these figures I conclude that the distributions are close to the beta family and to a gamma distribution. For the TRI MAD and TRI, I fit three conventional right-skewed distributions: the gamma, Weibull and log-normal. For the case of the slope measure, since it is defined as a percentage, I can also fit a beta distribution.[56] Panels (b), (d) and (f) of Figure 1.E.10 show the empirical and theoretical CDFs of the distributions for the three corresponding measures.

Next, I compute a goodness of fit test,[57] which yields the results shown in Table 1.E.6. The AIC and BIC favour the gamma for the case of the TRI MAD and TRI, and beta for the slope measure. For robustness, I also work with the second best distributions: the Weibull and gamma, respectively. The next step involves taking a series of random draws of these distributions, parameterised with the values found when fitting each distribution into the data.[58] Finally, I take the same approach as above to find 1,000 counterfactual values of the proportion of individuals receiving RNV. Averaging these values yields six new expected instrument candidates for each parish in the sample. Results with these instruments are in Table 1.E.7.

Perlin noise as pseudo-randomly generated topographies. I start by generating 1,000 draws of Perlin (1985) noise,[59] which I then rescale to match the average altitude in Venezuela.[60] Next, I perturb these values with scaled draws from a beta distribution to approximate mountains.[61] Having 1,000 counterfactual topographies allows me to calculate counterfactual values of the TRI,[62] which I then aggregate as outlined above into the MAD for each parish. Figure 1.E.11 shows the observed distribution of the TRI MAD in my data, along with 100 distributions of MAD calculated from 100 of my 1,000 counterfactual topographies. Finally, I take these 1,000 counterfactual values of TRI MAD

[54] I do this with the package fitdistrplus (Delignette-Muller & Dutang, 2015) in R (R Core Team, 2020).
[55] In all cases, I bootstrapped 2,000 values to give a clearer sense of the dispersion of the third empirical moment.
[56] This is implemented with the function fitdist().
[57] I use the function gofstat().
[58] I set the number of observations to 1,105 to match the full sample, and the number of iterations to 1,000.
[59] I use the package ambient and the function gen_perlin() in R (R Core Team, 2020), setting a frequency of 0.8.
[60] The noise is generated with values centred around 0 ranging approximately from -1 to 1. I add and multiply these values to 450, as 450m is the average altitude in Venezuela.
[61] For the draws of the beta distribution I pick shape parameters 0.5 and 5 to ensure a right-tailed distribution. I multiply these values by 1,000. Adding perturbations to Perlin (1985) noise is not uncommon. In fact several draws of the noise are aggregated in practical settings to add realism to the textures created.
[62] I implement the function tri() from the spatialEco library (Evans & Murphy, 2021).



Notes: Cullen and Frey Graphs and CDFs of measures of topographic ruggedness. Panels (a) and (b) correspond to the MAD TRI, panels (c) and (d) to the TRI, and panels (e) and (f) to the slope variability. I generate these figures with the fitdistrplus library (Delignette-Muller & Dutang, 2015) in R (R Core Team, 2020) bootstrapping 2,000 values.

Figure 1.E.10: Cullen and Frey Graphs and CDFs of Measures of Topographic Ruggedness

Table 1.E.6: Information Criteria for Different Right-Skewed Distributions Fit to Topographic Ruggedness Measures of Venezuela

	Panel A: TRI MAD			
	Weibull	Gamma	Log-Normal	—
AIC	7453.767	7412.958	7902.645	–
BIC	7463.782	7422.973	7912.661	–

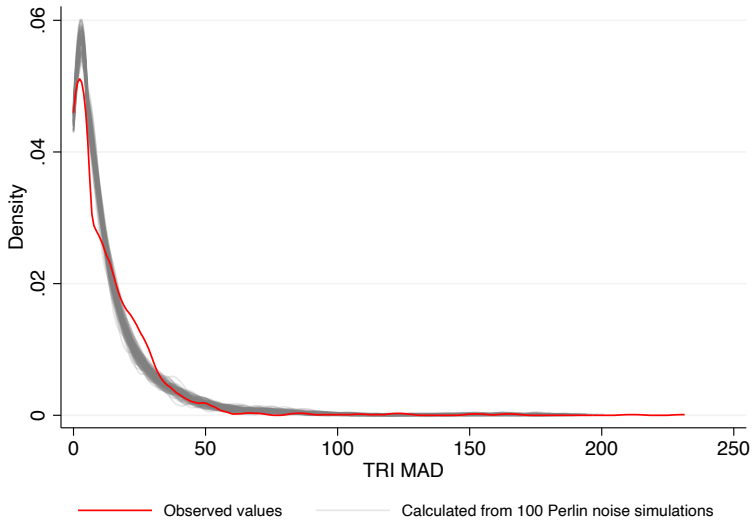
	Panel B: TRI			
	Weibull	Gamma	Log-Normal	—
AIC	2700.615	2677.503	2965.817	–
BIC	2710.630	2687.519	2975.832	–

	Panel C: Slope			
	Weibull	Gamma	Log-Normal	Beta
AIC	-5299.112	-5323.371	-5052.855	-5333.305
BIC	-5289.097	-5313.356	-5042.840	-5323.290

Notes: Goodness of fit tests for candidate theoretical distributions for the TRI MAD (panel A), TRI (panel B), and Slope (panel C). I use the AIC and BIC as criteria for selecting two distributions per variable.

and calculate 1,000 counterfactual values of the proportion of individuals receiving RNV following the same approach outlined above. Averaging these values gives me a final instrument, which I use in the results presented in Table 1.E.8.

The recommendation from Borusyak and Hull (2021) is that several potential candidate instruments $m_p(w)$ are included at the same time given that $G(\cdot)$ is unknown. I run the same regressions including the four instruments at the same time and present the results in Table 1.E.9. Results remain the same.



Notes: Observed distribution of TRI MAD shown in red overlayed against 100 counterfactual distributions generated with Perlin noise.

Figure 1.E.11: Observed and Counterfactual Distributions of TRI MAD

Table 1.E.7: IV Regressions Correcting for Non-Random Response to Random Shocks: Theoretical approximation

	Panel A: TRI MAD Chávez's Share 2000			
	(1)	(2)	(3)	(4)
RNV Coverage	0.025*	0.025*	0.025*	0.025*
	(0.014)	(0.014)	(0.014)	(0.014)
Expected Exposure TRI MAD (Weibull)		0.063*		
		(0.033)		
Expected Exposure TRI MAD (Gamma)				0.063*
				(0.033)
Observations	988	988	988	988
Estimation	2SLS	OLS	2SLS	OLS
R^2	0.652	0.654	0.652	0.654

	Panel B: TRI Chávez's Share 2000			
	(1)	(2)	(3)	(4)
RNV Coverage	0.025*	0.025*	0.025*	0.025*
	(0.014)	(0.014)	(0.014)	(0.014)
Expected Exposure TRI (Weibull)		0.066*		
		(0.034)		
Expected Exposure TRI (Gamma)				0.065*
				(0.034)
Observations	988	988	988	988
Estimation	2SLS	OLS	2SLS	OLS
R^2	0.652	0.654	0.652	0.654

	Panel C: Slope Chávez's Share 2000			
	(1)	(2)	(3)	(4)
RNV Coverage	0.025*	0.025*	0.025*	0.025*
	(0.014)	(0.014)	(0.014)	(0.014)
Expected Exposure Slope (Beta)		0.065*		
		(0.034)		
Expected Exposure Slope (Gamma)				0.066*
				(0.034)
Observations	988	988	988	988
Estimation	2SLS	OLS	2SLS	OLS
R^2	0.652	0.654	0.652	0.654

Notes: Regressions showing correction for a potential omitted variable bias akin to a non-random response to random shocks. I define the shock-assignment process as independent draws from theoretical distributions that approximate the observed distribution of different measures of ruggedness. Odd-numbered columns show the 2SLS correction, and even-numbered columns show their OLS counterpart. For 2SLS specifications, the instrument is the re-centred measure of exposure to RNV. Sample is restricted to match the sample of Table 1.E.8.

Table 1.E.8: IV Regressions Correcting for Non-Random Response to Random Shocks: Perlin noise

	Chávez's Share 2000	
	(1)	(2)
RNV Coverage	0.024*	0.024*
	(0.014)	(0.014)
Expected Exposure		0.068**
		(0.033)
Observations	988	988
Estimation	2SLS	OLS
R^2	0.652	0.654

Notes: Regressions showing correction for a potential omitted variable bias akin to a non-random response to random shocks. I define the shock-assignment process as independent draws of Perlin noise to generate counterfactual topographies. I then calculate the TRI MAD on these counterfactual topographies. Odd-numbered columns show the 2SLS correction, and even-numbered columns show their OLS counterpart. For 2SLS specifications, the instrument is the re-centred measure of exposure to RNV. Missing observations were not covered by the Perlin noise simulations. I decided to keep them as missing values rather than inputting data.

Table 1.E.9: IV Regressions Correcting for Non-Random Response to Random Shocks: Multiple instruments

	Chávez's Share 2000	
	(1)	(2)
RNV Coverage	0.023*	0.023*
	(0.014)	(0.014)
Expected Exposure TRI MAD (Perlin Noise)		1.681***
		(0.528)
Expected Exposure TRI MAD (Reshuffling)		-3.290
		(4.659)
Expected Exposure TRI MAD (Random Draws)		1.683
		(4.623)
Observations	988	988
Estimation	2SLS	OLS
R^2	0.652	0.657

Notes: Regressions showing correction for a potential omitted variable bias akin to a non-random response to random shocks. I use as instruments three re-centred measures of exposure generated through Perlin noise, permuting observed values of ruggedness, and taking independent draws from theoretical distributions. Odd-numbered columns show the 2SLS correction, and even-numbered columns show their OLS counterpart. For 2SLS specifications, the instruments are the re-centred measure of exposure to RNV. Sample is restricted to match the sample of Table 1.E.8.

1.F Alternative Determinants of Electoral Turnout

I explore other candidate determinants of electoral turnout; namely, the closeness of the previous election via what is known as the Downsian Cost Hypothesis, and Chávez's incumbent advantage. I find correlational evidence in both cases, though results cannot be interpreted as causal. In a horse-race regression, Chávez's margin is comparable in magnitude to the impact of reception of RNV. However, overall, reception of RNV is a better predictor of turnout than closeness.

Closeness and Electoral Turnout. This is a test of the Downsian Cost Hypothesis (DCH) (Downs, 1957; Riker & Ordeshook, 1968; Tullock, 1968). The DCH posits there is a negative relationship between the closeness of an election and electoral turnout: the closer an election (i.e., the smaller the winning margin), the higher the turnout.[63] Methodologically, this is a simple hypothesis to test. I follow the recommendations of Cox (1988) and Cox and Munger (1989), and I use raw margins, rather than relative margins to avoid finding spurious relationships, and I weight each observation with the parish's population share. Also, rather than using margins to predict contemporary turnout, as several papers in the literature do, I use Chávez's margin in 1998 to predict turnout in 2000.

Table 1.F.1 presents the results of several models that test the hypothesis. All models include the square of the raw margin and they differ by the set of additional covariates they include. Columns (4) through (6) shows the results of LASSO models that pick the variables from the full set of covariates that best explained the variation in the outcome, penalising model complexity by minimising the BIC. Overall, I find correlational evidence in support for the DCH wherein parishes that saw a closer result in 1998 had a higher turnout in 2000 with the caveat that these models might be prone to the same aggregation bias highlighted by Matsusaka (1993) and Matsusaka and Palda (1993).

Table 1.F.1: Testing the Downsian Cost Hypothesis in Venezuela's Presidential Elections 1998-2000

	(1)	(2)	(3)	(4)	(5)	(6)
Closeness (1998) / 100,000	-0.608***	-0.383***	-0.184**	-0.056*	-0.098***	-0.089***
	(0.093)	(0.080)	(0.083)	(0.031)	(0.037)	(0.033)
Observations	1095	1095	1095	1095	1095	1095
Avg. Luminosity (1998)	No	Yes	Yes	Yes	Yes	Yes
All States FE	No	No	Yes	Yes	No	Yes
All Muns. FE	No	No	No	No	Yes	Yes
LASSO	No	No	No	Yes	Yes	Yes
R^2	0.104	0.157	0.355	0.309	0.424	0.450

Notes: Controls include luminosity, state indicators, and municipality indicators. Robust standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

[63] There is a large literature in comparative political science testing this hypothesis and formalising it into rational voter models.

Incumbency Effect on Turnout. Another factor that could affect turnout is incumbency. Incumbency effects have long been studied. For instance, Cox and Katz (1996) find that incumbents can 'scare off' competition, Erikson and Titiunik (2015) find no question of an incumbent advantage in the United States, but the evidence form other countries is mixed and nuanced (e.g., Desposato & Petrocik, 2003; Grimmer, Hersh, Feinstein, & Carpenter, 2011; Hyytinen, Meriläinen, Saarimaa, Toivanen, & Tukiainen, 2018; Klašnja & Titiunik, 2017; Uppal, 2009). Following an empirical strategy that closely resembles that used to test the DCH, Table 1.F.2 presents the results of several models that predict the change in turnout between both elections with Chávez's raw margin from 1998.

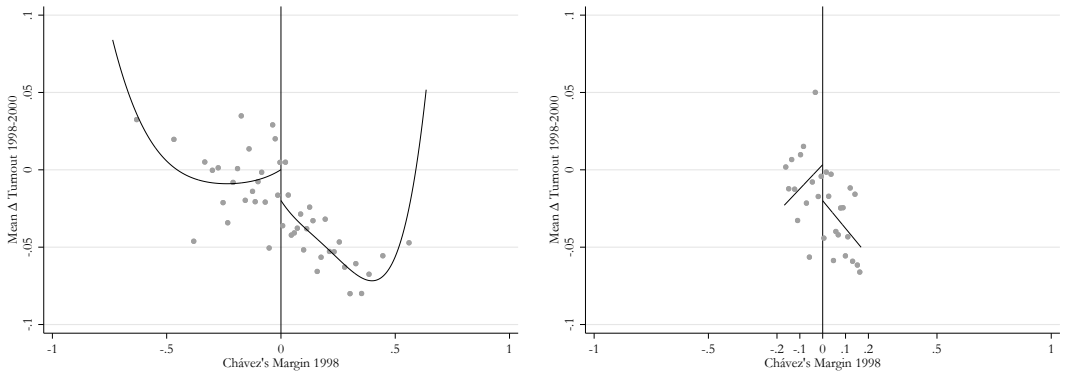
Similar to the closeness measure used above, I find evidence suggesting turnout decreasing in parishes where Chávez's share in 1998 was greater. However, to test this more formally, I employ a regression discontinuity design (see Cattaneo & Titiunik, 2022; Hahn, Todd, & der Klaauw, 2001; Lee, 2008; Lee & Lemieux, 2010; Thistlethwaite & Campbell, 1960) following the advice from Cattaneo et al. (2019). Panel (a) Figure 1.F.1 displays the mimicking variance regression discontinuity plot with quantile-spaced bins. The purpose of this figure is to visually inspect the distribution of the outcome on both sides of the proposed discontinuity.

Table 1.F.2: Chávez's Margin 1998 on Changes in Electoral Turnout 1998-2000

	(1)	(2)	(3)	(4)	(5)
Chávez's Margin (1998) / 100,000	-0.125**	-0.125***	-0.079***	-0.079***	-0.079***
	(0.058)	(0.044)	(0.025)	(0.025)	(0.025)
Observations	1095	1095	1095	1095	1095
Avg. Luminosity (1998)	No	Yes	Yes	Yes	Yes
All States FE	No	No	Yes	No	Yes
All Muns. FE	No	No	No	Yes	Yes
LASSO	No	No	Yes	Yes	Yes
R^2	0.033	0.144	0.450	0.450	0.450

Notes: Controls include luminosity, state indicators, and municipality indicators. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

In order to calculate the point estimate at the proposed discontinuity and to avoid Runger's phenomenon, I use a locally linear estimator, a triangular kernel and choose to set an MSE-optimal bandwidth on either side of the cutoff. Panel (b) of Figure 1.F.1 illustrates this. Even though these bandwidths are appropriate for point estimation, they are not for inference (see Calonico, Cattaneo, Farrell, & Titiunik, 2019; Cattaneo et al., 2019). Thus, Table 1.F.3 reports the point estimates, along with robust bias-corrected 95% confidence intervals, which have been calculated using coverage-error optimal bandwidths. Even though the visual inspection of the discontinuity plot and the point estimates seemed to suggest an effect in the vicinity of 2 percentage points, the robust confidence intervals include zero. Hence, I find no statistical evidence that turnout decreased between 1998



(a) VM Plot with Q-S Bins (b) Locally Linear Estimation

Notes: RDD plot in Panel A is a Variance Mimicking Quantile-Space bins plot (QSMV). Panel B is a regression discontinuity plot, using the same bandwidth on both sides of the cut-off, fitting a first-order polynomial. Both plots were generated using the rdplot command in Stata (Calonico, Cattaneo, & Titiunik, 2014).

Figure 1.F.1: Regression Discontinuity Plots

and 2000 in parishes where Chávez barely won in 1998, even if by and large there seems to be a statistically significant relationship at the mean.

Table 1.F.3: Regression Discontinuity Inference

	(1)	(2)	(3)	(4)	(5)	(6)
Margin Cutoff	-0.023	-0.022	-0.017	-0.021**	-0.012*	-0.008
	(0.015)	(0.015)	(0.014)	(0.009)	(0.006)	(0.006)
Observations	1093	1092	1093	1093	1093	1091
Robust 95% CI	[-.057 ; .008]	[-.054 ; .01]	[-.047 ; .011]	[-.041 ; -.003]	[-.023 ; .005]	[-.021 ; .007]
Robust p-value	0.140	0.170	0.216	0.026	0.189	0.326
Avg. Luminosity (1998)	No	Yes	No	No	No	Yes
All States FE	No	No	Yes	No	Yes	Yes
All Muns. FE	No	No	No	Yes	Yes	Yes

Notes: Additional controls include luminosity, state indicators, and municipality indicators. Calculations carried out using the rdrobust command in Stata (Calonico et al., 2014). Robust standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

Does State Media Trump Incumbency? If I include both dummy variables in a linear regression on turnout in 2000, while the point estimate associated with exposure to RNV is greater in magnitude than that of incumbency, it is not in a statistical sense. Furthermore, the null of a test of equality between both coefficients fails to be rejected. I present these horse-race results in Table 1.F.4. If I assume incumbency and RNV's reception are exogenous, I can test whether the incumbency effect explains more of the effect associated with the reception of RNV and *vice-versa*. I carry this out via two complementary twofold Oaxaca-Blinder decompositions, the first of which relies on the RNV reception indicator, and the second on an indicator for parishes where Chávez won in 1998. These indicators separate the data into two groups with statistically significant differences in turnout. While I rely on the indicators to separate the data into groups, in the regressions I include the continuous versions of the variables to take advantage of their variation.

Results are in Table 1.F.5, where columns (1) and (2) correspond to separating the data with the reception of RNV indicator, and columns (3) and (4) separate the data with an indicator for parishes where Chávez won in 1998. Even-numbered columns include state indicators. While the reception of RNV can account for about 20% of the difference in turnout observed in parishes where Chávez won in 1998, Chávez's incumbent advantage can account for only 10% of the difference observed in parishes coded as receiving RNV.[64] This is not different from what other variables such as luminosity can explain. From this simple exercise, I conclude that reception of RNV is better at explaining the difference in turnout that could be attributed to Chávez's incumbency than *vice-versa*.

Table 1.F.4: RNV Reception vs Incumbency Effect on Electoral Turnout in 2000

	(1)	(2)
(β_1) RNV Signal	-0.022***	-0.021***
	(0.006)	(0.007)
(β_2) Chávez Won 1998	-0.025***	-0.024***
	(0.004)	(0.005)
Observations	1096	1096
P-Value $H_0 : \beta_1 < \beta_2$	0.363	0.374
P-Value $H_0 : \beta_1 = \beta_2$	0.727	0.747
State Indicators	No	Yes
R^2	0.362	0.465

Notes: Linear regression of turnout in 2000 following the base specification from section 1.4.1 adding an indicator for parishes where Chávez won in 1998. The purpose of these regressions is performing a horse-race between both the incumbency and RNV indicators. I am not able to reject the null that both coefficients have the same magnitude.

To complement this decomposition I also estimate the Shapley values associated with each of these variables in a simple linear regression of turnout in 2000 such as those reported in Table 1.F.4. Figure 1.F.2 reports the relative importance of the Shapley values associated with each of these regressors. Again, reception of RNV explains nearly twice as much of the variation in turnout that closeness can explain.

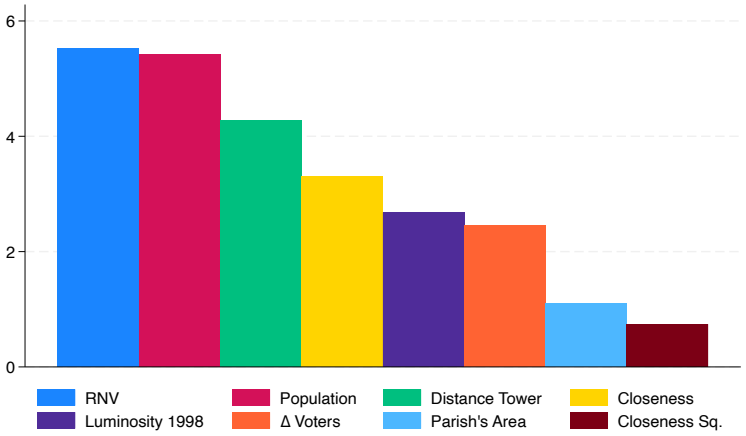
To summarise, I find suggestive correlational evidence in support for the DCH and an incumbency advantage. To the extent that I can test it, I find no causal evidence of the incumbency effect at the margin. In a linear regression, the coefficient associated with Chávez's incumbency advantage is comparable in magnitude to the effect associated with exposure to RNV. However, under a Oaxaca-Blinder decomposition, I find that over twice as much as the difference in turnout rates between parishes where Chávez won and lost in 1998 can be explained by exposure to RNV than what can be explained by Chávez's margin between parishes receiving RNV. This finding is replicated by looking at the contribution

[64] This interpretation relies on the assumption that there are no omitted variables.

Table 1.F.5: Oaxaca-Blinder Decompositions

	Receives RNV		Chávez Won 1998	
	(1)	(2)	(3)	(4)
Panel A: Group Differences				
Group 1	0.583***	0.583***	0.590***	0.590***
	(0.003)	(0.003)	(0.004)	(0.004)
Group 2	0.553***	0.553***	0.562***	0.562***
	(0.004)	(0.004)	(0.003)	(0.003)
Difference	0.030***	0.030***	0.028***	0.028***
	(0.005)	(0.005)	(0.005)	(0.005)
Panel B: Explained Differences				
Chávez's Margin (1998) / 100,000	0.003***	0.003***		
	(0.001)	(0.001)		
Proportion RNV Signal			0.005***	0.006***
			(0.002)	(0.002)
Observations	1096	1096	1096	1096
Select FEs	No	Yes	No	Yes

Notes: Oaxaca-Blinder decomposition of main results under two different dimensions. Columns (1) and (2) separate the sample along reception of RNV. Columns (3) and (4) separate the data along an indicator for whether Chávez won in 1998. In both cases, these are assumed exogenous. Panel B reports the explained differences by the included covariates in the regressions. The interpretation of these numbers is relative to the unexplained difference reported in Panel A. Receiving RNV's signal is a better predictor than Chávez's Margin. Robust standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.



Notes: The figure shows the relative contribution of each variable in explaining the variation in turnout in 2000 captured by a linear regression. Covariates include those in the specifications described in section 1.4.1 except demographics and state indicators to make the estimating algorithm more efficient. The R^2 of the regression with all these covariates is 0.3349. The numbers on the vertical axis should be interpreted as percentage points of 33.49%. The covariate with the highest explanatory power is the lag of turnout with a value of 73.83%. I exclude this covariate from the figure to focus on values of similar magnitudes.

Figure 1.F.2: Relative Importance of Shapley Values

of each of these variables to explaining turnout by calculating Shapley values. This leads me to conclude that reception of RNV is a better predictor of turnout in 2000 than other classical determinants such as the closeness of the previous election or the margin of the incumbent candidate.

1.G Details of Survey Analysis

I test whether exposure to RNV had any medium-term effects on political preferences by aggregating all available waves of Latinobarometer pre-Chávez and until 2004 – the year when RNV underwent an expansion process which included the creation of several FM frequencies as documented in Appendix 1.B.1. The waves I work with are from 1996, 1997, 2000, 2001, 2002, 2003, and 2004. Latinobarometer does not interview the same individuals in each of its waves. In order to make the comparison across time as clean as possible, I restrict my sample to municipalities that were part of the survey in at least two points in time: one pre-Chávez and one after Chávez became president. Further, not all questions were asked in every wave of the survey. Again, to make the comparison as clean as possible, I restrict the sample to include the same question in at least one period pre-Chávez and one period after he became president in 2000.

I recode the questions to either have a binary outcome, or an ordinal outcome whenever I can take advantage of the variation in the responses. My main independent variables are the share of people receiving RNV in 1998 in the relevant municipality, a post-1998 indicator – the year Chávez became president –, and the interaction of these two variables. All regressions include a battery of covariates which include the mean of the luminosity in the municipality in 1998, the age, gender and educational attainment of the respondent; indicators for the number of luxury items the individuals owns or has access to;[65] state indicators, and year indicators when the question is asked in more than two waves. The coefficient of interaction terms in non-linear regressions cannot be interpreted as causal (Ai & Norton, 2003). Thus, I complement my logit and ordered logit regressions with linear probability models. Figure 1.G.2 shows the results for all binary outcomes, and Figure 1.G.3 for all ordinal outcomes. In both figures, Panel (a) presents the linear models, while Panel (b) their non-linear counterparts. Several trends emerge, which I summarise below, as well as in Figure 1.G.1, which only includes the OLS results of these regressions.

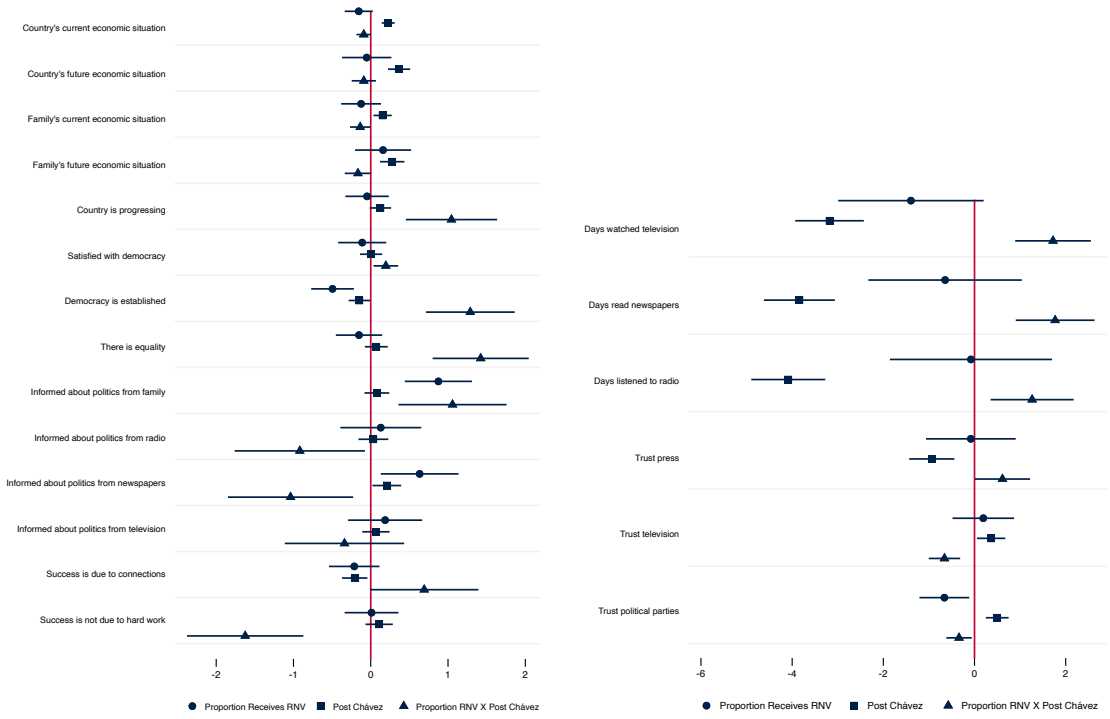
Heightened sensitivity to inequality; economic pessimism; institutional optimism. When exposed to RNV, individuals tend to be more sensitive towards the prevailing inequalities in the country: respondents strongly believe the current income distribution is not fair, they also tend to think that women do not currently have access to the same jobs and salaries as men. Although I find no difference when it comes to access to education. With time, respondents became slightly more optimistic about their personal and country's economic conditions. However, when exposed to RNV this optimism was slightly offset. Despite the economic pessimism, respondents overall believed there is equality in the

[65] These goods include a television, a refrigerator, a computer, a washer, a telephone, a car, a second house, drinking water, hot water, and sewage.

country, and that success can be achieved through hard work, even if they feel success is also due to personal connections, although the latter result is weak.

Respondents believed the country was progressing overall. Crucially, with a higher exposure to RNV, I find that people seem to feel more satisfied with democracy. Respondents also believe democracy has been fully established in Venezuela. These findings are particularly relevant as they highlight a tension between liberal and direct democracy exemplified by the adoption of the Constitution of 1999.

Substitution of other means of communication; higher interest in politics. Overall, post 1998 people spend less days watching television, listening to radio or reading newspapers. This trend away from traditional means of communication might be due to an increased availability of the internet, though this was not part of the Latinobarometer survey. This decrease is smaller though when respondents are exposed to RNV. Being exposed to RNV is also associated with an increased interest in politics and discussing politics with family members. Although, individuals seem to rely less on newspapers and radio for informing themselves about politics as exposure to RNV increases. Exposure to RNV is also associated with a decrease in trust in the press and television. This finding can be easily rationalised in light of Chávez's constant attacks to other means of communication as discussed in section 1.2. Finally, I also observe an overall increase in trust in political parties in the years after 1998, which is slightly offset by exposure to RNV. This rise in trust in parties naturally follows the strong fallback from the early- to mid-1990s. The fact that the recovery is not as strong in areas with stronger exposure to RNV is probably related to the cult around Chávez as an individual, the *outsider* quality that made him an attractive candidate in the first place.

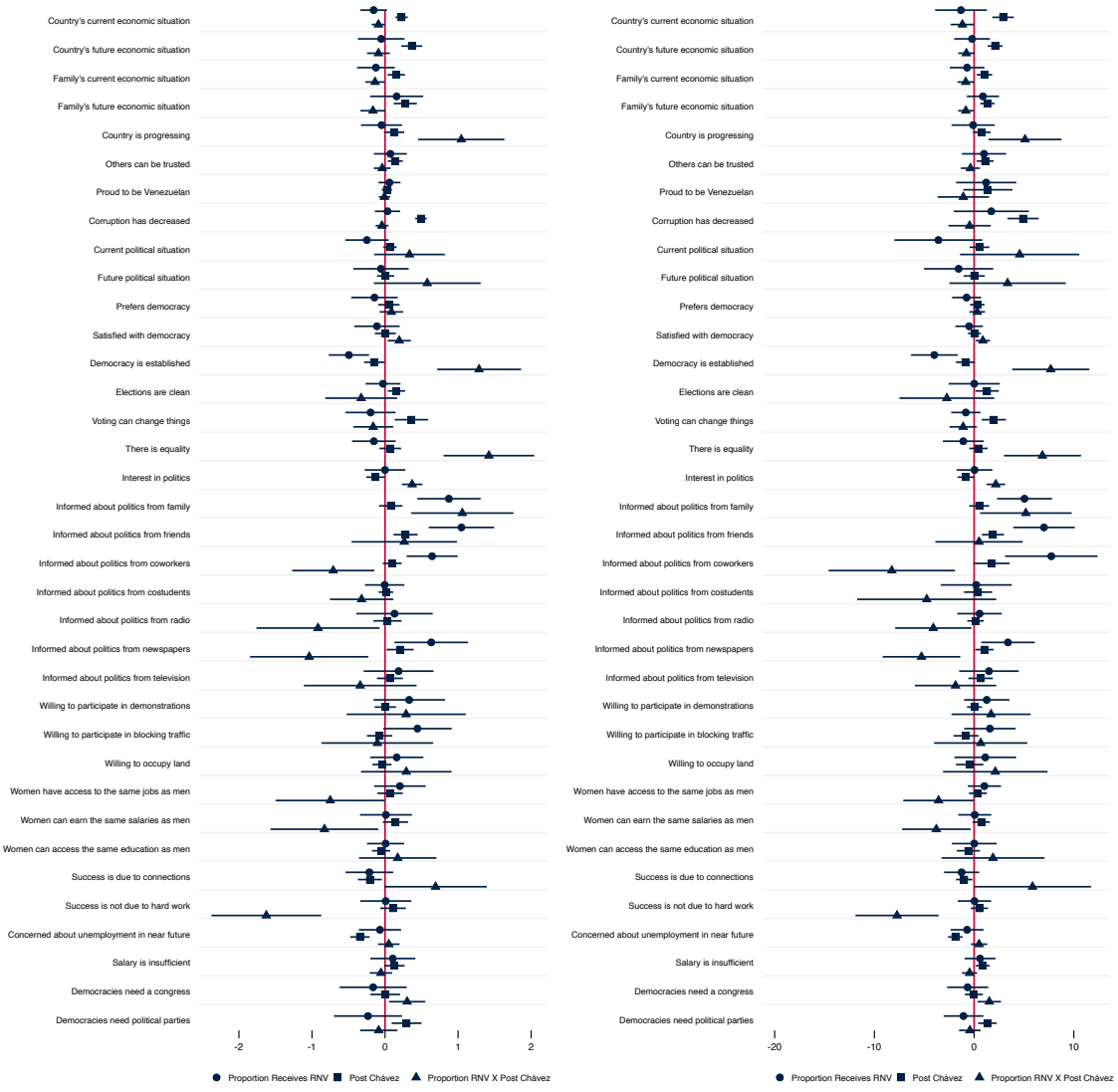


(a) Binary Outcomes

(b) Ordered Outcomes

Notes: Specifications are OLS regressions. All regressions include a battery of covariates which include the mean of the luminosity in the municipality in 1998, the age, gender and educational attainment of the respondent; state indicators, year indicators when the question is asked in more than two waves; and indicators for the number of luxury items the individuals owns or has access to. These goods include a television, a refrigerator, a computer, a washer, a telephone, a car, a second house, drinking water, hot water, and sewage.

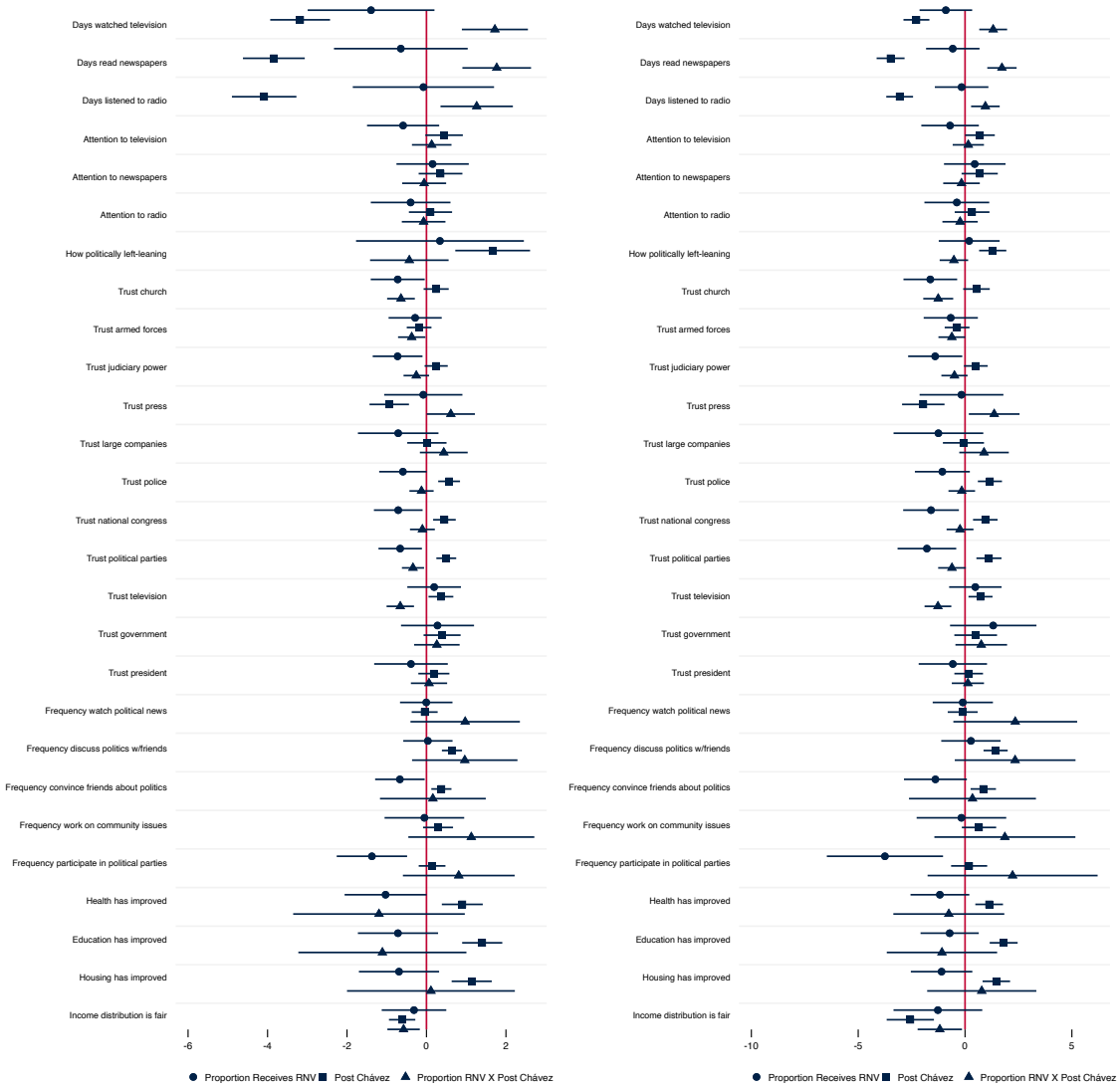
Figure 1.G.1: RNV Reception on Survey Outcomes



(a) LPMs (b) Logits

Notes: Specifications are OLS regressions in Panel A, and Logit regressions in Panel B. All regressions include a battery of covariates which include the mean of the luminosity in the municipality in 1998, the age, gender and educational attainment of the respondent; state indicators, year indicators when the question is asked in more than two waves; and indicators for the number of luxury items the individuals owns or has access to. These goods include a television, a refrigerator, a computer, a washer, a telephone, a car, a second house, drinking water, hot water, and sewage.

Figure 1.G.2: Analysis of Latinobarometer: Binary outcomes



(a) LPMs

(b) Ordered Logits

Notes: Specifications are OLS regressions in Panel A, and ordered logit regressions in Panel B. All regressions include a battery of covariates which include the mean of the luminosity in the municipality in 1998, the age, gender and educational attainment of the respondent; state indicators, year indicators when the question is asked in more than two waves; and indicators for the number of luxury items the individuals owns or has access to. These goods include a television, a refrigerator, a computer, a washer, a telephone, a car, a second house, drinking water, hot water, and sewage.

Figure 1.G.3: Analysis of Latinobarometer: Ordered outcomes

1.H Historical Context Details

The Constitution of 1999. Some of the changes implemented in the 1999 Constitution in Venezuela include changing the country's name to República Bolivariana de Venezuela; using a non-androcentric language in the Constitution's text; guaranteeing human rights and also social rights (e.g., healthcare, housing and employment); giving international treaties an equal standing as the Constitution; formalising an obligation from the state to prevent any discrimination against anyone (e.g., women and minorities); creating a transfer scheme for female homeowners wherein they receive social benefits proportional to household labour; instituting a right to timely, true, and impartial information; eliminating state funding of political parties; establishing four different mechanisms for public referenda. The new Constitution also demarcated indigenous lands; it guaranteed the protection of indigenous languages, and bilingual education; it prohibited patents based on indigenous knowledge; it guaranteed political representation for indigenous communities. It also established environmental rights.

The Constitution separated government into 5 branches, rather than 3: the executive, judiciary, legislative, electoral and citizen powers; the latter of which could call for referenda to change the constitution, abrogate laws and even recall the president's mandate. Citizens could participate directly in politics through referenda, local and communal councils, social audits, citizen assemblies, direct involvement in the government, and cooperatives. If the government betrayed the constitution, citizens had the obligation to commit civil disobedience. The legislative power changed from a bicameral to a unicameral system. Presidential terms were extended from 5 to 6 years. The participation of the state in the direction of the economy was also increased. The constitution also granted great power to the military. It transformed it into a civil corporation, where only the president could approve of military promotions (thus aligning incentives between the executive branch of government and the military). It gave the military the rights to vote and to hold public office; tasked them with the implementation of government programmes; promoted civil-military unity; and, as in Cuba, encouraged the creation of popular defence units in case of an invasion from the United States. It also united the air force, navy, national guard and army into a single National Armed Force.

Initial political programme of Chávez's government. According to Wilpert (2006), their economic policies for the next 6 years were centred around 5 axes: 1) generating sustainable economic growth by diversifying the economy, promoting the vertical integration of the oil sector, preventing capital flight and stabilising the exchange rate; 2) reducing economic volatility through the creation of a macroeconomic stabilisation fund financed

with oil surpluses; 3) creating a social economy around micro-enterprises and cooperatives, redistributing land, organising economic activity on the basis of solidarity and the common good, and the participation of government-owned micro-credit banks; 4) fiscal sustainability through increased revenues and expenditure efficiency; and 5) ensuring state ownership of the Venezuelan oil company, and a strong leadership within the Organization of Petroleum Exporting Countries. Chávez's government also implemented a considerable set of social policies, which focused on land reform, education, housing, and social security. However, these were mostly implemented after 2003.

Chapter 2

Poverty and Low Human Capital: The Legacy of African Slavery in Venezuela and the Deep Historical Roots of *Chavismo*

Abstract

This paper studies the economic and political effects of the legacy of the use of African enslaved labour in colonial Venezuela. I take advantage of a unique colonial onomastic norm for assigning surnames to enslaved people – when enslaved people were taken to Venezuela they were registered and received the name of their ethnic groups or place of origin as surnames. The persistence of these surnames through time allows me to identify individuals with enslaved ancestry in contemporary Venezuela. I find a negative effect of enslaved ancestry on development outcomes and a positive effect in support for Hugo Chávez. The mechanism of historical persistence is low human capital and low-productivity occupational specialisation. By digitising new historical sources, I am also able to identify locations where enslaved labour was used during the colonial period. I find that individuals with enslaved ancestry do not live in locations where enslaved labour was utilized during the colony. The latter locations are not economically worse-off today and were not more likely to support Chávez. Taken together, these findings emphasise differences between mobile and place-based factors of historical persistence, suggest a historical explanation for factors contributing to the rise of *Chavismo*, and highlight the long shadow of enslavement even in peripheral areas of a colonial empire.

Acknowledgements: I want to thank my DPhil supervisors Sanjay Jain and Paola Conconi for encouraging me to write this article, for the continuous support and many helpful suggestions. I also want to thank James Robinson, Elias Papaioannou, Toke Aidt, Gabriel León Ablán, Dean Yang, David Doyle, Nelson Ruíz, Gianni de Fraja, Sandra Sequeira, Julien Labonne, Andrea Bernini, Doug Gollin, Christopher Woodruff, Nicolas Cerkez, Kai Gehring, Andreas Fuchs, Krisztina Kis-Katos, Martin Gassebner, Panu Poutvaara, Oliver Harman, Rocco Zizzamia, Niccolò Meriggi, Liyu Zheng, Daniela Horta, Jasmin Baier, Antonia Delius, Krittika Ray, Brynde Kreft, Jordan Imahori, and John Walker for helpful comments at various stages of this project. Thank you to participants and discussants at the CSAE Worshop 2024, the CSAE Conference 2025, the EPCS Conference 2025, the NICEP Conference 2025, the RES Conference 2025 for their helpful comments, questions, and suggestions. I am also grateful to Francesco Feriozzi for helping me find one of the maps used in this article. I acknowledge financial support from the University of Oxford's Department of Economics Oxford Economics Papers (OEP) fund in the realisation of this project during my DPhil. All mistakes and omissions are my own responsibility.

2.1 Introduction

> *Hate against me has a lot to do with racism. Because of my big mouth, because of my curly hair. And I'm so proud to have this mouth and this hair – because it's African.*
>
> *- Hugo Chávez*

> *[Hugo Chávez's] image bothers the posh ladies of Caurimare and other parts of the country, while for the lowly he symbolizes 'one of them' who has attained the Presidency of the Republic understanding their problems and fulfilling promises not fulfilled by the parties of the past. (…) Colour differences do matter (…). Skin color is a mark which, depending on the shade of café con leche [coffee with cream], hinders or facilitates social mobility, just as it opens or closes windows of opportunity.*
>
> *- Patricia Márquez, Vacas flacas y odios gordos: la polarización en Venezuela*

The long-term effects of colonialism and slavery have been widely studied (e.g., Acemoglu, Johnson, & Robinson, 2001; Dell, 2010), but their influence on contemporary political preferences remains relatively unexplored. This paper studies the economic legacy of the use of African enslaved labour in Venezuela, how it persisted through time, and how Hugo Chávez was able to appeal to voters whose ancestors were enslaved.

Explanations for historical persistence can be conceptually separated into two distinct groups: place-based factors, such as legal frameworks or formal norms (e.g., Acemoglu et al., 2001); and mobile factors, like culture and social norms (Becker & Voth, 2023; Nunn & Wantchekon, 2011), or human capital (Rocha, Ferraz, & Soares, 2017). Venezuela's unique history allows us to separate locations where enslaved labour was utilized during the colonial period, and locations where individuals with enslaved ancestry currently live. This unique setting allows us to test for persistence of place-based and mobile factors on contemporary outcomes.

To identify enslaved ancestry in contemporary Venezuela, I take advantage of a unique feature of colonial history in the Caribbean and Central Latin America. Enslaved African people were brought to these locations mainly to work in plantations. Upon arrival, they were often registered and given the name of their ethnic group or region of origin as a

surname.[1] This norm is documented in Acosta Saignes (1984) and Ortega (1992). This arrangement created a unique set of surnames that has persisted (and evolved) for over 200 years and is still relevant today. An example of this phenomenon is the surname Carabalí. Both Acosta Saignes (1984) and Ortega (1992) highlight that the origin of this surname is Kalabar in Nigeria. The transposition of 'r' and 'l' is a typical example of a linguistic phenomenon called metathesis, entailing the transposition of sounds or syllables in a word.[2,3]

Acosta Saignes (1984) provides a list of surnames that indicate African enslaved ancestry and that exist in contemporary Venezuela.[4] I match these surnames to the surnames of all adults in Venezuela through the Tascón-Maisanta list.[5] While this database has been used in research previously (Hsieh et al., 2011),[6] this is the first time the list of names itself is used as a source of data. Out of the 12 million people in the database, I classify nearly 2 million as having enslaved ancestry through their surnames.[7]

Interestingly, these individuals are not located near places to which enslaved people were brought from Africa during the colonial period. I define two proxies for exposure to the use of enslaved labour during the colony: proximity to colonial cocoa plantations – retrieved from a newly digitised[8] historic data source – and proximity to the ports of arrival of enslaved individuals, as documented in the Trans-Atlantic and Intra-American Slave Trade Databases (Eltis, 1999). The fact that these measures and the enslaved ancestry metric go in opposite directions is not altogether surprising. After slavery was banned in 1854, former slaves relocated far from cocoa plantations and lived as subsistence farmers. This migration was partly in response to rumours that they would be enslaved again if they remained on plantation land, but also to break away from extractive quasi-feudal labour institutions that survived from the colonial period such as the *encomienda* system

[1] This is different from, for instance, the prevailing norm in other places like the United States and even Colombia, where enslaved individuals were given as a surname a version of the name of whoever bought them.

[2] In this case, the former word is easier to pronounce for a Spanish speaker so the surname evolved to Carabalí.

[3] Thank you to Francesco Feriozzi and Sydelle de Souza for pointing this out.

[4] This list is reproduced in Appendix 2.B.4.

[5] The Tascón-Maisanta list is a database created and distributed by Chávez's regime in 2004 to intimidate the opposition following the collection of signatures to have a recall referendum against Chávez. The database includes the personal information of all adults that were registered to vote in the country: their full name, address, age, and date of birth, whether they were recipients of government programmes, and whether they signed the petition to have a presidential recall referendum. See Section 2.3 and Appendix 2.B.4 for more details.

[6] Hsieh et al. (2011) found that individuals who signed in favour of the referendum were systematically discriminated against when applying for jobs and were more likely to remain unemployed for a longer period of time.

[7] To the best of my knowledge, there has not been another major migration shock from Africa to Venezuela since the colonial period.

[8] I digitise the location of these plantations from a map drawn by de Santos Cabrera (n.d.) in the mid-XVIIIth century, and define one of the measures of exposure as the proximity to these plantations.

(Arcila Farías, 1966; Brito Figueroa, 1973).

The fact that the majority of these individuals became subsistence farmers informs the empirical strategy I follow to push the causality between enslaved ancestry and recent socio-economic outcomes. To address concerns around potentially omitted variables or simultaneity, I rely on two-state least squares (2SLS) specifications. As an instrument, I use the soil's suitability for growing wheat. Wheat was a staple of subsistence farmers' diets in Venezuela during the XVIIIth century, and one of the most widely-produced cereals in the country (Codazzi, 1960; Codazzi et al., 1940). Crucially though, it was not produced for exports – making it a historically relevant predictor for the location of subsistence farmers after conditioning for the location of cocoa *haciendas*.[9] I argue that the exclusion restriction for the validity of the instrument is satisfied, by including a comprehensive list of covariates, several of which are new variables that I digitise from historical sources.

As a first set of results, I identify a negative effect of the share of individuals with enslaved ancestry on three measures of local economic development. First, I find a negative effect on night-light density, proxying for overall development[10] (Henderson et al., 2011, 2012). Second, I find that higher illiteracy rates are associated with higher shares of individuals with enslaved ancestry. Interestingly, there is no effect on measures of schooling – these include enrolment rates at primary, secondary and pre-school levels, or the children per school ratio of a given municipality. This is in line with current research in education whereby children might be attending school but learning very little (N. Angrist, Djankov, Goldberg, & Patrinos, 2021). Access to schools is not the problem – the issue lies in the quality of what is being taught. And third, I find a negative effect on child mortality. This effect remains constant if I disaggregate the data into pre-natal, infant ($<$ 1 year) or older-child ($>$ 1 year) mortality. It is also not the case that people are simply more likely to die in these locations, as I see no effect on the overall mortality rate. The effect is entirely driven by children – the most vulnerable members of communities.

Second, I provide correlational evidence that suggests that these effects are due to the persistence of low human capital. For this analysis I pool all available Venezuelan censuses and find that the measure of enslaved ancestry is persistently correlated with illiteracy and lower educational attainment. Enslaved ancestry is also positively correlated with the share of people engaging in primary activities, specifically subsistence agriculture and other low-skill occupations. It is also negatively correlated with the proportion of individuals engaging in secondary and tertiary activities, as well as high-skill occupations.

[9] Using a crop's suitability as an instrument is akin to the strategy used by Acharya et al. (2015, 2016b) and similar in spirit to that used by Giuliano and Nunn (2021).

[10] Other desirable proxies for economic development such as Meta's Relative Wealth Index (Chi, Fang, Chatterjee, & Blumenstock, 2022) are not available for the case of Venezuela at the time of writing.

I also find a persistent negative correlation with an asset-based measure of welfare.

Third, I identify a strong effect of the proportion of individuals with enslaved ancestry on support for Hugo Chávez during his first presidential election in 1998.[11] The effect on his vote share becomes stronger after I control for the aforementioned development outcomes. I account for this increase in the magnitude and precision of the estimates via a Gelbach (2011) decomposition. Informed by the previous set of results, I also perform a mediation analysis following Acharya et al. (2016a). The average causal direct effect of African enslaved ancestry on Chávez's share of votes remains strong and statistically significant, suggesting that underdevelopment is not enough to explain the appeal of Chávez to these individuals.

Chávez's anti-system rhetoric was able to catalyse the frustration of hundreds of thousands of Venezuelans – having been excluded from politics by a consolidated elite for nearly 50 years, being dissatisfied by the series of major economic reforms adopted at the end of the 1980s, and feeling discontent due to prevailing inequalities and persistent poverty (Cannon, 2013). I find correlational evidence suggesting that locations with higher shares of individuals with enslaved ancestry are also locations where individuals tended to express more dissatisfaction with political life before Chávez was president. Survey respondents also expressed distrust towards the president and political parties; they expressed an overall lack of political engagement, and exhibited zero-sum thinking. Other factors that might have contributed to Chávez's appeal which I cannot fully capture empirically include the persuasiveness of Chávez's rhetoric on top of the rhetoric itself (De Anda Casas, 2025a; Frajman, 2014), or factors related to race and representation (Cannon, 2013; Herrera Salas, 2005; Ishibashi, 2003).

Fourth, following a similar identification strategy, I rely on cocoa suitability to instrument for proximity to colonial cocoa *haciendas* and estimate the effect between exposure to the use of enslaved labour and all the aforementioned outcomes. I find no statistical evidence of an effect on any of the development outcomes. However, the signs of the point estimates suggest that, if anything, locations exposed to the use of enslaved labour are better-off today by those metrics. I also find no statistically significant effect on the share of votes that Chávez received. These results strengthen the previous findings and nuance the implications of this paper: the legacy of African slavery in Venezuela lives with the descendants of those slaves.

This paper contributes to the literature in three ways. First, to the best of my knowledge, it is the first work in the persistence literature[12] that explores a historical explanation of

[11] These specifications are informed by those in De Anda Casas (2025a).

[12] This paper builds on a large literature which includes, for instance, Acemoglu et al. (2012, 2001); Acemoglu,

factors that contributed to the rise of a left-wing populist candidate. The closest work would be Becker and Voth (2023), which offers an explanation for the appeal of a strongman leader in the historical persistence of culture and religion. In the case of this paper, however, the mechanisms for persistence are low human capital and occupational specialisation that help explain contemporary measures of poverty. Another closely related paper is Nunn and Wantchekon (2011), which explores the persistence of being exposed to the slave trade on levels of mistrust in Africa. While Nunn and Wantchekon's analysis does centre on individuals and the experiences of their ethnic groups during the colonial period, it lacks the political dimension that this paper contributes.

Second, being able to distinguish locations where enslaved labour was used during the colonial era, from areas where the people negatively affected by it live today, makes this a unique setting yielding a new set of findings. Conceptually, this pushes the idea that both human capital and formal norms such as laws can persist, but the former is mobile – it is carried by people – while the latter are immobile (see for instance Fisman and Miguel (2007), or the example of Nogales in Acemoglu and Robinson (2012)). This distinction frees us from the 'place-based curse' of institutions.

The third contribution is from a historical point of view. Venezuela was considered part of the periphery of the Spanish empire for nearly 200 years (Mahoney, 2010; Wright, 1990). Mahoney (2010) contrasts areas with extractive institutions such as Peru, Mexico, Colombia, Bolivia and Guatemala with peripheral areas such as Argentina, Chile, Costa Rica, Uruguay and Venezuela, arguing that the former received their mercantilist institutions under Habsburg rule, as opposed to the latter that developed under Bourbon rule and thus were endowed with more liberal institutions. There are two reasons why Venezuela remained on the periphery of the Spanish Empire. First, there were no easily extractable commodities, apart from pearls along the coast and some gold deposits along the Andes. Second, there were no organised population masses given that the territory was mostly inhabited by a few nomadic indigenous groups. In line with arguments in Acemoglu et al. (2002) and Engerman and Sokoloff (2012b), there were thus no economic incentives for the Spanish to build institutions in this part of the world, until the XVIIIth century's shift in demand for cocoa in Europe. The fact that we observe historical persistence even in fringe areas that received comparatively liberal institutions speaks to the importance of human capital as a persistence mechanism.

This result ties back to the evidence also put forth by Valencia Caicedo (2019), Rocha et

Johnson, and Robinson (2002, 2005); Acemoglu and Robinson (2012); Acharya et al. (2015, 2016b); Dell (2010); Dell, Lane, and Querubin (2018); Engerman and Sokoloff (2005a, 2005b, 2012a, 2012b); Fisman and Miguel (2007); Giuliano and Nunn (2021); Michalopoulos and Papaioannou (2013, 2014, 2016); Miguel and Roland (2011); Nunn (2008); Nunn and Puga (2012); Nunn and Wantchekon (2011); Rocha et al. (2017); Valencia Caicedo (2019); Voigtländer and Voth (2012), etc.

al. (2017) and Dell (2010). The work of Dell (2010), Acemoglu et al. (2012) and Acharya et al. (2015, 2016b) focuses on areas where heavily mercantilist institutions existed. Valencia Caicedo (2019) studies institutional persistence at the limit of the extent of the Spanish empire by focusing on Paraguay. However, both Rocha et al. (2017) and Valencia Caicedo (2019) focus on historical features that allowed certain populations to thrive in the long-run through the preservation of human capital. In contrast, the evidence in this paper speaks to the persistence of low human capital even in a fringe area that developed under comparatively liberal institutions.

The paper is organised as follows. Section 2.2 describes the historical context, providing an overview of the colonial period in Venezuela, the post-independence years, and the legacy of African slavery in today's Venezuela. Section 2.3 details all the data sources. Section 2.4 outlines the empirical strategy discussing the validity and relevance of the candidate instrument used in the 2SLS strategy. Section 2.5 presents the results in four subsections: 2.5.1 discusses the effect of enslaved ancestry on the set of development outcomes, 2.5.2 presents the evidence on historical persistence, 2.5.3 turns to the effect of enslaved ancestry on support for Hugo Chávez, and 2.5.4 complements the previous subsections by highlighting the null relationship between these outcomes and the locations where African slavery existed in colonial Venezuela. Section 2.6 provides additional robustness checks, and section 2.7 concludes.

2.2 Historical Context

The Spanish Colonial Period. Throughout most of the colonial period, the territory that is now Venezuela was considered the periphery of the Spanish Empire (Mahoney, 2010; Wright, 1990). The territory lacked developed indigenous populations, or easily-exploitable mineral commodities, and remained an administratively fragmented society (Wright, 1990). The main economic activities in this territory during the XVIth century included slave trafficking, fishing pearls, and an incipient mining industry that had 'unflattering results, and in many cases barely any or no utility' (Brito Figueroa, 1966).

African slaves began to be imported to the territory around 1525 – almost a century before they began to be imported into the United States (Brito Figueroa, 1973). They were mostly used for extracting the few commodities that were easily available i.e., pearls and gold. It is estimated that during the XVIth century 6,595 slaves were brought into the territory and 10,147 in the XVIIth (Brito Figueroa, 1966). The boom in slave imports occurred during the XVIIIth century when 34,099 slaves were brought in (Brito Figueroa, 1966). Other estimates place the number at 70,513 (Ortega, 1992). Slave imports followed

a boom in European demand for cocoa. African slave labour only substituted indigenous labour to an extent. During this period of time, two forms of coerced labour i.e., slavery and serfdom, coexisted in places such as cocoa *haciendas*.

In 1728 a charter called the *Compañía Guipuzcoana* was created. It had the monopoly of cocoa production until 1784. The company represented the biggest demand for African slaves in Venezuela. Some examples of locations which these individuals came from include Kalabar in Nigeria or the Mina group from Mali-Benin (Ortega, 1992). Cocoa production was concentrated around the coasts and near the valleys such as the Orinoco basin. Slaves were also forced to work in the production of sugar and tobacco, which were mainly produced to satisfy the territory's internal demand, and not to the extent that cocoa was produced. In 1775 cocoa represented 75% of the exports of La Guaira – Venezuela's largest port. In comparison, leather represented 17%, tobacco 6%, and indigo 0.8%. At the turn of the century, coffee became an important crop as well, but never as important as cocoa (Vila, 1965).

By 1800 Venezuela had 898,043 inhabitants. 61% of them had African ancestry: 407,000 *pardos*, descendants of slaves and whites, who worked in service trades, crafts, urban labour, or as serfs in *encomiendas*; 33,362 free or *manumiso* individuals; 87,000 slaves; and 24,000 *cimarrones*, slaves that had escaped and lived in quasi-nomad groups. There were 184,722 whites, who comprised a landowning elite and controlled all political institutions (Brito Figueroa, 1966, 1973).

The Post-Independence Transition. The beginning of the XIXth century was characterised by political and economic instability. The War of Independence (1811-1823) provided an opportunity for slaves to emancipate. Some joined the pro-independence army, others escaped and became *cimarrones*. Ultimately, this created a crisis for the landowning elite, as the price of slave-labour increased. Slowly, African slavery lost importance. Lombardi (1971) argues that 'no one after 1830 seriously attempted to justify slavery as either a positive or even a relative good.' Other historians such as Brito Figueroa (1966) argue it simply was no longer profitable to continue with that system.

These conditions led to the adoption of the Manumission Laws of 1821 and 1830. These laws established that descendants from enslaved people would gain freedom after they turned 18. This age was later pushed to 21. In the meantime, these individuals were known as *manumisos*. Manumission was not a new system – it existed in Venezuela since the XVIIIth century (Brito Figueroa, 1966). A *manumiso*'s labour could be sold, but the individual per se. The price of a *manumiso*'s labour was half the price of a slave of the same age. When the first generation of *manumisos* under the Manumission Laws was about to reach freedom in 1840 a new system was introduced – the Apprenticeship system. *Manumisos*

were then forced to remain 'under tutelage' until the age of 25.

The adoption of the manumission system created a period of transition where the number of slaves rapidly declined between 1820 and 1850. African slavery simply transformed itself into other types of forced labour. Agustín Codazzi – Venezuela's first geographer – estimated that the slave population in 1838 was only 5% of the country's total population (Codazzi et al., 1940). However, Watts (1987) and Engerman and Sokoloff (2012b) document that nearly 60% of the population still had African enslaved ancestry even in 1840. Eventually, slavery was fully abolished in 1854: slaves were now free-men and slave-owners received pecuniary reparations in return.

These free-men were immediately integrated into the quasi-feudal labour system that already existed. They were allowed to live and work on the fringe of land controlled by white, *criollo* or former slave-owner elites in exchange for tributes. Other times they became landowners' pawns and received a symbolic payment in exchange for working their land. Sometimes they were paid with tokens that could only be exchanged for food and other goods at shops [*tiendas de raya*] owned by the landowners themselves. These tokens were often not enough to cover the (monopolistic) prices of these goods, and the workers became indebted to the landowners for the rest of their lives (Brito Figueroa, 1966, 1973).

Former slaves who decided not to stay in the same land where they used to work migrated to unoccupied land and became *pisatarios* (Brito Figueroa, 1966). They were 'poor peasants who practised a subsistence agriculture of little profitability, destined to produce fruits and basic grains in the popular diet' (Brito Figueroa, 1966). Other people migrated to cities where they became masons, barbers, tanners, blacksmiths, artisans or domestic servants (Brito Figueroa, 1966). Migration also took place as there was a fear that 'people who continued to reside in the places where they had been slaves would again be reduced to slavery' (Brito Figueroa, 1973). Between 1858 and 1863 a civil war ensued. The outcome of the civil war was the consolidation of political power of pre-existing elites. In the aftermath, the country opened to foreign capital investment, which kick-started Venezuela's economic development during the late XIXth and early XXth centuries.

The 'Myth of Racial Democracy' and the Legacy of African Slavery. By the end of the colonial period, *mestizaje* was so prevalent that it is estimated that up to 75% of the population were either of African origin or their descendants (Wright, 1990). For most of the XIXth and XXth centuries, race remained an understated yet present cleavage in Venezuelan society. For instance, up until the late 1940s, it was impossible for non-whites to immigrate into the country. Even if the notion of outright discrimination was shunned by most of Venezuelan society, it was a fact that the poorest individuals were

descendants of African slaves. Acosta Saignes (1984) highlighted that these individuals remain stuck in poverty primarily due to a lack of access to education and economic opportunities. Racism also remained entrenched in society. In fact, it has been documented that racial tensions scapegoating people of African descent, and members of indigenous communities, increased following shocks to the price of oil in the 1980s (Herrera Salas, 2005).

The 'myth of racial democracy' refers to the homogenising ideology of *mestizaje*, which according to Herrera Salas (2005) and Ishibashi (2003) was but an attempt to 'whiten' the population and mask the prevailing inequalities in the country. Even though race was not an explicit component of Hugo Chávez's original presidential campaign, the Constitution of 1999 did explicitly grant much more autonomy to minority groups such as indigenous communities (see De Anda Casas, 2025a; Garcia-Serra, 2001; Herrera Salas, 2005; Wilpert, 2006). Following Chávez's election in 1998, race soon became a controversial topic in Venezuela with many instances of outright racism displayed in media and condoned by members of Chávez's political opposition (see e.g., Cannon, 2013; Herrera Salas, 2005; Ishibashi, 2003).

2.3 Data Sources

Enslaved ancestry. I combine two data sources to measure African enslaved ancestry in Venezuela. The first input is a list of African slave surnames from Acosta Saignes (1984). I provide the full list and additional context in Appendix 2.B.3. The second input is the Tascón-Maisanta database, which includes the full names of over 12 million citizens in Venezuela that were registered to vote in 2004. I provide more context on this data source and details on how I processed it in Appendix 2.B.4. I split individuals' names into paternal and maternal surnames, and apply a hybrid algorithm[13] to match them to the list of names from Acosta Saignes (1984).[14] This algorithm yields a similarity score between names. I classify those with a score above 85% as names denoting enslaved ancestry.[15,16] Finally, I calculate the share of individuals with names denoting enslaved ancestry in each civil parish. The map in Figure 2.A.1 plots this variable. I probe the robustness of how I defined enslaved ancestry in Section 2.6.

[13] A hybrid algorithm combines a vectorial and phonetic decomposition of tokens i.e., strings. In general, these algorithms perform better than their uniquely vectorial or phonetic counterparts (see e.g., Hodge & Austin, 2003; Pfeifer, Poersch, & Fuhr, 1996; Zobel & Dart, 1995).

[14] I apply the soundex_ext algorithm though the matchit function in Stata.

[15] The threshold of 15% is akin to a false-positive rate. In this case, I opt for a number above the usual 20% assumed as standard in social sciences.

[16] I show robustness to changes in this threshold, and the choice of algorithm, in Section 2.6.

Crop suitability. I rely on data from the Food and Agriculture Organization's Global Agro-Ecological Zones project. Specifically, I download their suitability indexes calculated over all land in each grid cell (rather than just cropland) for the range of years from 1961 to 1990. These rasters are processed in ArcGIS Pro, where I calculate the weighted average suitability at the level of the civil parish and then apply the inverse hyperbolic sine (i.h.s.) transformation to those values. The raw data for cocoa and wheat are mapped in Figures 2.A.4 and 2.A.5, respectively.

Development outcomes. Following Henderson et al. (2011, 2012), one of the main development outcomes is night-time luminosity averaged at the level of the civil parish in 1998. The second outcome corresponds to illiteracy rates at the level of the parish from the 1990 Venezuelan census. The third outcome is the mortality rate of children relative to the number of children living in each municipality of the country. For robustness, I disaggregate the mortality rate into pre-natal mortality, infant ($<$ 1 year) mortality, and older ($>$ 1 year) children mortality. I complement this data with the number of deaths relative to the population in a given municipality, disaggregated into preventable deaths, highly-preventable, moderately preventable, and non-preventable deaths.

Political variables. The main political variables correspond to the Venezuelan presidential election of 1998 – Hugo Chávez's first election. I retrieved the data from ESDATA.info through the Wayback Machine, and aggregate all information at the level of the civil parish. The main outcome is the share of votes that Chávez received. I also calculate the turnout rate relative to the number of registered voters at the level of the parish. To explore why Chávez appealed to voters living in locations with higher enslaved ancestry, I rely on the 1996 and 1997 waves of Latinobarometer.[17]

Historical sources. Throughout the paper, I use three key historical sources. First, I retrieve the location of cocoa *haciendas* by digitising a map of the Province of Caracas in the mid-XVIIIth century. The map was originally drawn by de Santos Cabrera (n.d.), and includes the location of cocoa *haciendas*, cities and villas, monasteries and missions, military strongholds, farms, roads, and lakes. I calculate the distance from the centroid of each parish in the country to the location of each *hacienda* and transform it using the i.h.s.. I also calculate the distance to the nearest XVIIIth century city, villas, monasteries, farms and missions, apply the same transformation, and include them as controls in all regressions. This map is reproduced in Figure 2.B.1. Figure 2.B.2 shows the map overlaid on a contemporary map of Venezuela, and Figure 2.B.3 shows the data points extracted from the map. Figure 2.A.2 maps the distance from the centroid of each civil parish to the

[17] I decide to focus on pre-Chávez waves, and I use those that report which state/municipality respondents live in.

closest *hacienda*.

The second source proxies for locations used in large-scale agricultural production during the colonial period. This source is the map from Codazzi (1840) that depicts Venezuela's grasslands, jungles, and areas devoted to agriculture. The map is part of the first atlas created for a South American country and is reproduced in Figure 2.B.4. Finally, I use the Trans-Atlantic and Inter-American Slave Trade Databases (Eltis, 1999) to identify enslaved people's ports of arrival in Venezuela. Figure 2.B.6 shows the location of these ports and Figure 2.A.3 maps the location from the centroid of each civil parish to the closest port.

Additional controls. These variables include the distance to the nearest contemporary port, the distance to the nearest shore, land cover in 2000, the population of each civil parish, and whether there is indigenous population living in a given parish, as well as additional demographics. The location of contemporary ports comes from VesselFinder. I calculate the distance from the centroid of each parish to the nearest shore using ArcGIS Pro. For consistency, I apply the i.h.s. transformation to all distance variables. For contemporary land cover, I use the Global Land Cover 2000 map, which is available through the European Commission (2003).[18] The two key controls that I retrieve from the land cover data is the proportion of land in a parish used in agriculture, and the proportion of land that is classified as urban. These variables are important as they contribute to the validity of the instrumental variables strategy.

The population of each parish is calculated using the Global Population Density Grid Time Series data from NASA (CIESIN, 2011, 2017). The indigenous population of each parish in Venezuela is captured by digitising the data in OCEI (1993a).[19] In some specifications I include additional demographics – for instance, an index created following the methodology in Anderson (2008) which aggregates precarious living conditions at the level of the municipality, with data from the 1990 census. These demographics were digitised from the 25 books published by OCEI (1993b, 1993c) and have also been used in De Anda Casas (2025a). For the persistence analysis I rely on the harmonised censuses from IPUMS (2020). The shapefile of Venezuela is from Porto Tapiquén (2020).

[18] I opt for this database over more detailed maps such as those from Sentinel-2 as the latter are only available from 2015 onwards.

[19] Just as in De Anda Casas (2025a), I digitise this data using KofaxPowerPDF.

2.4 Empirical Strategy

2.4.1 Wheat Suitability as a Valid Instrument

The main threat to identifying a causal relationship between enslaved ancestry and the outcomes arises from potentially omitted variables. Thus, the main specifications follow a two-stage least-squares (2SLS) strategy using crop suitability as an instrument.[20] This strategy is akin to that used by Acharya et al. (2015, 2016b) to study the causal impact of enslavement on current political outcomes in the United States. The validity of the instrument is achieved after conditioning on enough confounders that allow us to claim that crop suitability is only affecting our outcomes through the instrumented variable.

In all specifications, I include a crucial set of covariates comprising a set of variables that capture local development and economic activity during the Spanish colonial period and today.[21] Specifically, I control for the distance to XVIIIth century cities, villas, monasteries, missions, and farms;[22] as well as the proportion of each contemporary civil parish that was used for large-scale agricultural production in the XIXth century.[23] Additionally, I control for the proportion of the land in each parish that is currently used in agriculture, and that is classified as urban.[24] Additional controls include the distance to the closest shore, the distance to the closest contemporary port, a parish's latitude and longitude, and their squares;[25] the population of each parish relative to the country's population,[26] as well as state indicators.[27] The collection of these variables is what I refer to as my baseline specification.

The distances to settlements during the XVIIIth century are included as they can capture patterns of labour mobility after Venezuela's independence from Spain, as well as areas that have a long history of higher economic activity. The distance to the location of missions and monasteries is important given that most missionary groups established in this area were Jesuits – it has been argued that areas where Jesuits were located during the colony, and extractive institutions were not in place, have persistently higher rates of educational attainment and higher incomes to this day (Valencia Caicedo, 2019).

[20] I present simple correlational evidence and explore heterogeneities using OLS in Section 2.6.
[21] I show robustness of results to excluding potential intermediate outcomes i.e., 'bad controls' in Section 2.6.
[22] These variables were created in ArcGIS Pro after digitising the map by de Santos Cabrera (n.d.). For more details, see Appendix 2.B.1.
[23] This variable comes from digitizing the bottom panel of the map from Codazzi (1840)
[24] These variables come from the Land Cover 2000 map from the European Commission (2003).
[25] These variables were calculated using ArcGIS Pro. The location of all contemporary ports is from VesselFinder.
[26] This variable comes from NASA's Global Population Density Grid Time Series (CIESIN, 2011, 2017).
[27] I do not control for municipality indicators as multiple parishes are singletons.

Controlling for variables that capture areas devoted to agricultural production during the XIXth century and today is relevant as it allows us to interpret any relationship between the endogenous variable and the proposed instrument as resulting from the suitability of the specific crop used as instrument, rather than agricultural suitability broadly defined. Including these controls highlights the potential suitability for growing a specific crop over and above the (observed) production in any given area today. Controlling for areas devoted to large-scale agricultural production during the XIXth century closes the causal path between colonial land-estates [*latifundios*] and *encomiendas* that produced agricultural goods for exports[28], and all outcomes. I probe the robustness of this strategy in Section 2.6.

Thus, for all civil parishes p I estimate the first stage equation represented by equation 2.1. As a second stage, I use the predicted enslaved ancestry from each parish to estimate the causal effect β_2 on outcome y_p, as equation 2.2 denotes. In both equations, $\mathbf{X_p}$ denotes a vector of controls, with γ and δ being the vectors of coefficients associated with these controls for equations 2.1 and 2.2, respectively. Finally, υ_p and ϵ_p denote each equation's residuals.

$$\text{EnslavedAncestry}_p = \beta_1 \text{WheatSuitability} + \mathbf{X_p}\gamma' + \upsilon_p \tag{2.1}$$

$$y_p = \beta_2 \widehat{\text{EnslavedAncestry}}_p + \mathbf{X_p}\delta' + \epsilon_p \tag{2.2}$$

2.4.2 Instrument's Relevance and First Stage Results

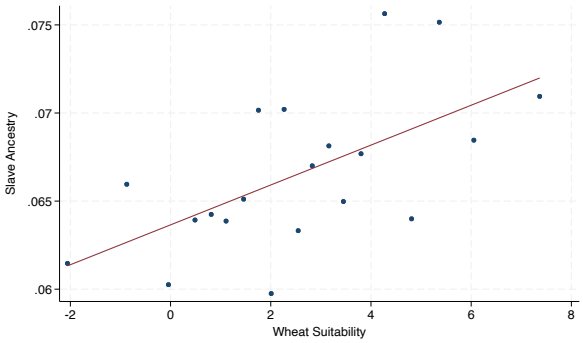
The relevance of the instrument can be argued from a historical perspective, and quantified through the first stage of the empirical strategy (equation 2.1). Following the abolition of slavery in 1854, most former slaves relocated to unoccupied land including the fringes of the large estates [*latifundios*] owned by white and *criollo* elites (Brito Figueroa, 1966, 1973). Most of these individuals became subsistence farmers. Wheat was one of the most widely produced cereals in colonial Venezuela[29] and a staple in subsistence farmers' diets[30] (Codazzi, 1960; Codazzi et al., 1940). Crucially, wheat was not produced for export.[31] Thus, wheat suitability leverages three historically relevant factors: first, it was a staple in farmers' diets; second, it was widely harvested, but not for exports; and third, it was

[28] As described in Section 2.2, the War of Independence did not involve any land redistribution. Landowning elites continued controlling the same land even after the war (see Brito Figueroa, 1966, 1973)

[29] The four most important cereals in Venezuela's agricultural production were wheat, rice, maize and millet.

[30] Farmers also ate pulses including beans, kidney beans, *alverjas*, *tapiramos*, *quinchonchos*, chickpeas, lentils, peas, millet, rice, and aniseed (Codazzi, 1960; Codazzi et al., 1940).

[31] The main crops produced for export included cotton, indigo, cocoa, coffee, sugarcane, coconut and tobacco (Codazzi, 1960; Codazzi et al., 1940).



Notes: Binscatter plot representing the first stage results of intrumenting enslaved ancestry with wheat suitability. Controls include state indicators, a dummy for parishes with indigenous population, population in 1990; latitude, longitude, and their squares; distance to XVIIIth century villas, monasteries, missions, farms; the area of a parish devoted to large scale agriculture in the XVIIIth century; distance to contemporary ports, distance to shoreline, proportion of land classified as urban, and proportion of land devoted to agriculture.

Figure 2.1: Binscatter Plot Showing Strength of First Stage Results

ubiquitous and easy to grow in land that was not already used for growing other crops.[32]

Table 2.A.1 reports the first stage results, as well as the results of conventional identification tests. Column 1 reports coefficient β_1 of the baseline specification (equation 2.1). There is a strong, positive and statistically significant relationship between both variables, which I illustrate in Figure 2.1. The null of underidentification is comfortably rejected given the Kleibergen-Paap statistic that Table 2.A.1 reports. Further, the Cragg-Donald F-statistic used in the weak identification test rests between the provided Stock-Yogo critical values of 8.96 and 16.38 for 15% and 10% maximal IV size, respectively. While the instrument is not the strongest, I am comfortable with its level of potential bias as the statistic is closer to the stricter critical value. Naturally, I do not include an overidentification test as equation 2.1 is just-identified.

2.5 Main Results

2.5.1 Enslaved Ancestry and Development Outcomes

I analyse the effect of enslaved ancestry on a group of outcomes that are good indicators of local economic development broadly defined. The first of these outcomes is luminosity following Henderson et al. (2011, 2012). I calculate the average luminosity in each parish in Venezuela in 1998, and then apply the i.h.s. transformation to interpret results as a

[32] In the first treatise of Venezuelan geography published in 1840, Agustín Codazzi states that wheat is a cereal 'so necessary in Europe [that] finds in Venezuela a multitude of plains, plateaus and valleys of great extension susceptible of its cultivation', but the fact that these 'lands [are] suitable for coffee, sugar cane, indigo, cotton, banana, corn, and a portion of plants that substitute bread, has been the cause that wheat has not taken a considerable increase; however, there are entire cantons whose only bread is wheat' (Codazzi, 1960).

semi-elasticity.[33] The results are reported in column 1 of Table 2.1. In the median parish in Venezuela, 6.3% of the population has African enslaved ancestry according to this measure. This implies that the median parish in terms of enslaved ancestry will be 2% less luminous than the average parish in the country.

The next of these outcomes is the illiteracy rate measured at the level of the parish in 1990. This data was digitised from the census results published in 25 volumes by OCEI (1993b, 1993c). I focus on literacy as it captures a fundamental component of human capital that goes beyond schooling. I find that in the median parish in Venezuela in terms of enslaved ancestry, the illiteracy rate is 24 percentage points higher.

Importantly, these results are not due to poor schooling. I replicate this exercise with enrolment rates at different educational levels i.e., primary, secondary and pre-school; as well as with the number of children per school in a given municipality, and find no effects on any of these outcomes (see Table 2.A.3). While it may well be that the results on illiteracy are driven by the adults in a parish, this finding resonates with recent work in the economics of education that highlights the difference between school access and school quality (N. Angrist et al., 2021), whereby children might indeed be attending school, but learning very little.

The third outcome I explore in this section is child mortality relative to the number of children in a given municipality. Here, I find that the median parish in terms of enslaved ancestry, will be in a municipality with a child mortality rate 3 percentage points higher than the average. I check whether this effect is driven by children of a particular age by splitting the mortality rate into pre-natal, infant ($<$ 1 year), and older children ($>$ 1 year), and find that the effect remains statistically significant across specifications (Table 2.A.4).

Interestingly, it is not the case that people in general are more likely to die in these places. Table 2.A.5 reports the same exercise on the number of preventable, highly-preventable, moderately preventable, and non-preventable deaths relative to the population of a given municipality. Here, I find no effects of enslaved ancestry, which emphasises the precarious picture painted by the finding on children's mortality rates.

2.5.2 Persistence Mechanism

In this section I present suggestive evidence that the mechanism of persistence behind the development outcomes analysed in section 2.5.1 is low human capital and skill formation. For these analyses I rely on the harmonised censuses from IPUMS (2020), which for the case of Venezuela comprise the four waves from 1970 to 2000. The downside of

[33] Results are robust to using the raw weighted averages instead.

Table 2.1: Effect of Enslaved Ancestry on Development Outcomes

	(1) Luminosity (i.h.s.)	(2) Illiteracy (1990)	(3) Child Mortality (2000)
Enslaved Ancestry	-33.713***	4.152***	0.553**
	(12.980)	(1.401)	(0.239)
Observations	1098	1098	1098
Population	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes
Specification	2SLS	2SLS	2SLS

Notes: Main results on development outcomes, which include luminosity in 1998, illiteracy rates in 1990, and child mortality ratio in 2000. Enslaved ancestry is instrumented with wheat suitability. Baseline controls include state indicators, a dummy for parishes with indigenous population, population in 1990; latitude, longitude, and their squares; distance to XVIIIth century villas, monasteries, missions, farms; the area of a parish devoted to large scale agriculture in the XVIIIth century; distance to contemporary ports, distance to shoreline, proportion of land classified as urban, and proportion of land devoted to agriculture. Robust standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

using this harmonised data is a loss of statistical power, as the data is aggregated using boundaries consistent with Venezuela's political division in 1970, and only up to the second administrative level i.e., municipalities. With this level of aggregation, the instrumental variables approach used in the other sections of the results is not a viable identification strategy. Hence, this section relies solely on correlational evidence, and I consider the findings as suggestive.

First, I pool all the available waves of the Venezuelan census. In these regressions, all specifications include the baseline set of controls, as well as state and year indicators. Table 2.2 reports these conditional correlations. To analyse human capital, I define an indicator for literacy, educational attainment up to primary school, and a third indicator for attainment up to secondary school; finally, I also use years of schooling. The outcomes are the collapsed weighted averages for each municipality-year. Across the board, I observe a negative and strongly significant correlation between enslaved ancestry and these measures of human capital (Panel A, columns 1-4). The findings on primary and secondary school, albeit suggestive in nature, are not necessarily at odds with the causal evidence provided in section 2.5.1 wherein I discussed null effects for enrolment, rather than attainment as is the case here.

The second set of outcomes that I analyse involves the structure of the economy. Here, I separate industrial activities into broad primary, secondary and tertiary sectors; and I calculate the weighted rate of participation in each of these sectors for the economically

active population at the level of the municipality. This information is only available from the last three censuses, which is why we have fewer observations in columns 1-3 of Panel B. Here, I observe a strong and positive correlation between enslaved ancestry and participation in primary industries; in contrast to a strong and negative correlation with secondary and tertiary industries. I check the robustness of these results to alternative ways of aggregating industrial activities and occupations. Table 2.A.7 separates the proportions of participation into individual industries, rather than groups of industries. The positive correlation with primary industries reported in Table 2.2 is entirely driven by participation in agriculture, while the other negative correlations are not driven by any specific industries.

Third, I find a negative correlation between enslaved ancestry and a measure of welfare, which I define as an Anderson (2008) index on the proportion of people in each municipality-year that have access to electricity, piped water, and sewage disposal; and that own a telephone, refrigerator, television, and a flushable toilet. I find that enslaved ancestry is negatively correlated, albeit weakly, with this index, as reported in column 4 of Table 2.2's Panel B.

Finally, as a complement to this analysis, I repeat the exercise with each individual wave of the census. Across the board, the correlations maintain the signs reported in the pooled analysis, albeit often at the expense of statistical significance due to the much smaller samples. These correlations are presented as a collection of binscatter plots in Figure 2.2.

2.5.3 Enslaved Ancestry and Votes for Chávez

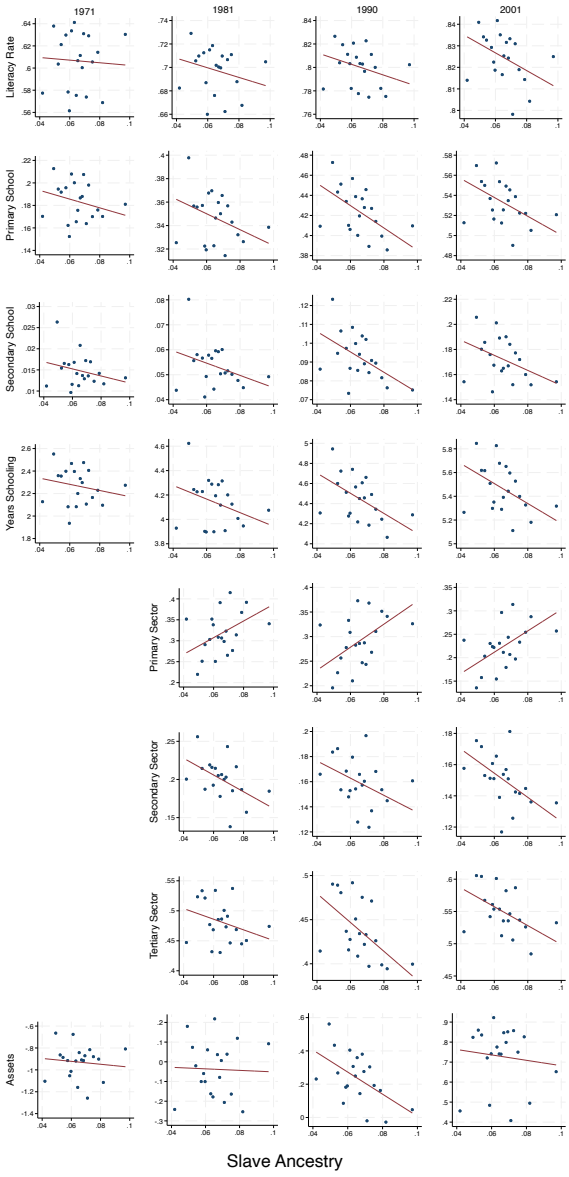
Electoral Results

Next, I explore whether enslaved ancestry is related to support for Hugo Chávez in his first presidential run. I follow the same 2SLS strategy to identify this effect with the addition of new controls that allow me to provide a cleaner estimate. The new controls include the unemployment rate measured at the level of the parish in 1990, an Anderson (2008) index that aggregates precarious living conditions measured at the level of the municipality in the 1990 census with data digitised from the books published by OCEI (1993b, 1993c).[34,35] These specifications are informed by those of De Anda Casas (2025a). The results are presented in column 1 of Table 2.3.

On top of these controls, I also add the outcomes from section 2.5.1 with the exception

[34] The index is an inverse-correlation Anderson (2008) index created with standardized versions of the proportion of houses with walls made out of wood, trowelled and untrowelled adobe, and 'other' materials; the proportion of houses with roofs made out of asbestos, metal or 'other' materials; and the proportion of houses with floors made out of dirt or 'other' materials.

[35] Column 2 of Table 2.A.1 reports the corresponding first stage results.



Notes: Binscatter plots of enslaved ancestry measured at the level of the municipality with census variables capturing human capital, occupational specialisation, and an asset-based measure of welfare by census wave. Outcomes include literacy, individuals with primary and secondary schooling completed, average years of schooling, participation in primary, secondary, and tertiary activities and an asset-based index created with the methodology from Anderson (2008). Controls include state indicators, a dummy for parishes with indigenous population, population in 1990; municipalities' latitude, longitude, and their squares; distance to XVIIIth century villas, monasteries, missions, farms; the area of a parish devoted to large scale agriculture in the XVIIIth century; distance to contemporary ports, distance to shoreline, proportion of land classified as urban, and proportion of land devoted to agriculture.

Figure 2.2: Binscatter Plots of the Historical Persistence of Enslaved Ancestry

Table 2.2: Low Human Capital and Skill Formation as a Mechanism of Persistence

Panel A	Human Capital			
	(1) Literacy	(2) Primary School	(3) Secondary School	(4) Years Schooling
Enslaved Ancestry	-0.347**	-0.779***	-0.367***	-6.686***
	(0.154)	(0.188)	(0.113)	(1.749)
Observations	628	628	628	628
Panel B	Industrial Structure			Welfare
	(1) Primary Sector	(2) Secondary Sector	(3) Tertiary Sector	(4) Assets Index
Enslaved Ancestry	2.203***	-0.852***	-1.329***	-2.423*
	(0.465)	(0.183)	(0.352)	(1.351)
Observations	471	471	471	628
Population	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes
Year Indicators	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes
Contemporary Land Coverage	Yes	Yes	Yes	Yes
Dist. Shore and Ports	Yes	Yes	Yes	Yes
Specification	OLS	OLS	OLS	OLS

Notes: Correlation of enslaved ancestry measured at the level of the municipality with census variables capturing human capital, occupational specialisation, and an asset-based measure of welfare. Panel A focuses on human capital. Outcomes include literacy, individuals with primary and secondary schooling completed, and average years of schooling. Panel B focuses on occupational specialisation disaggregated into primary, secondary, and tertiary activities; and welfare, proxied by an asset-based index created with the methodology from Anderson (2008). Controls include census wave indicators, state indicators, a dummy for parishes with indigenous population, population in 1990; municipalities' latitude, longitude, and their squares; distance to XVIIIth century villas, monasteries, missions, farms; the area of a parish devoted to large scale agriculture in the XVIIIth century; distance to contemporary ports, distance to shoreline, proportion of land classified as urban, and proportion of land devoted to agriculture. Robust standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

of the children's mortality rate, as this was measured in 2000.[36] The sequential addition of these variables is presented in columns 2 and 3 of Table 2.3, with column 4 representing my preferred (unrestricted) specification. The magnitude of the effect attributed to enslaved ancestry increases with the addition of these variables.

In order to 'account' for this increase, I perform a Gelbach (2016) decomposition. These results are presented in Table 2.A.8, which indicate the change in the magnitude of the coefficient is overwhelmingly (ca. 80%) due to the inclusion of luminosity, which has a statistically significant and negative relation with the outcome. Naturally, if we take the sum of the impact of these variables and subtract it from the coefficient of enslaved ancestry, we recover the magnitude from the restricted specification in column 1.

[36] Table 2.A.6 in the Appendix explores adding the children's mortality rate. The main results remain unchanged.

The natural subsequent question is whether the outcomes analysed in section 2.5.1 mediated the effect of enslaved ancestry on Chávez's share. Column 5 of Table 2.3 presents the Average Causal Direct Effect (ACDE) of enslaved ancestry, estimated with the methodology from Acharya et al. (2016a). In this specification, this coefficient is net of any effect that could have been mediated by luminosity and the illiteracy rate. The magnitude of the coefficient is similar to the restricted specification from column 1, but the precision of the estimate has increased. Reassuringly, the coefficient from column 5 is also similar to the net effect from column 4, after having subtracted the sum of impacts from the Gelbach decomposition in column 1 of Table 2.A.8.

Table 2.3: Effect of Enslaved Ancestry on Votes for Hugo Chávez

	2SLS Linear Regressions				ACDE
	(1)	(2)	(3)	(4)	(5)
Enslaved Ancestry	2.799*	3.513**	3.736**	3.906**	2.469**
	(1.613)	(1.700)	(1.690)	(1.709)	(1.188)
Observations	1087	1087	1087	1087	1087
Luminosity	No	No	Yes	Yes	Yes
Illiteracy Rate	No	Yes	No	Yes	Yes
Population	Yes	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes	Yes
Additional Controls	Yes	Yes	Yes	Yes	Yes
Specification	2SLS	2SLS	2SLS	2SLS	2SLS

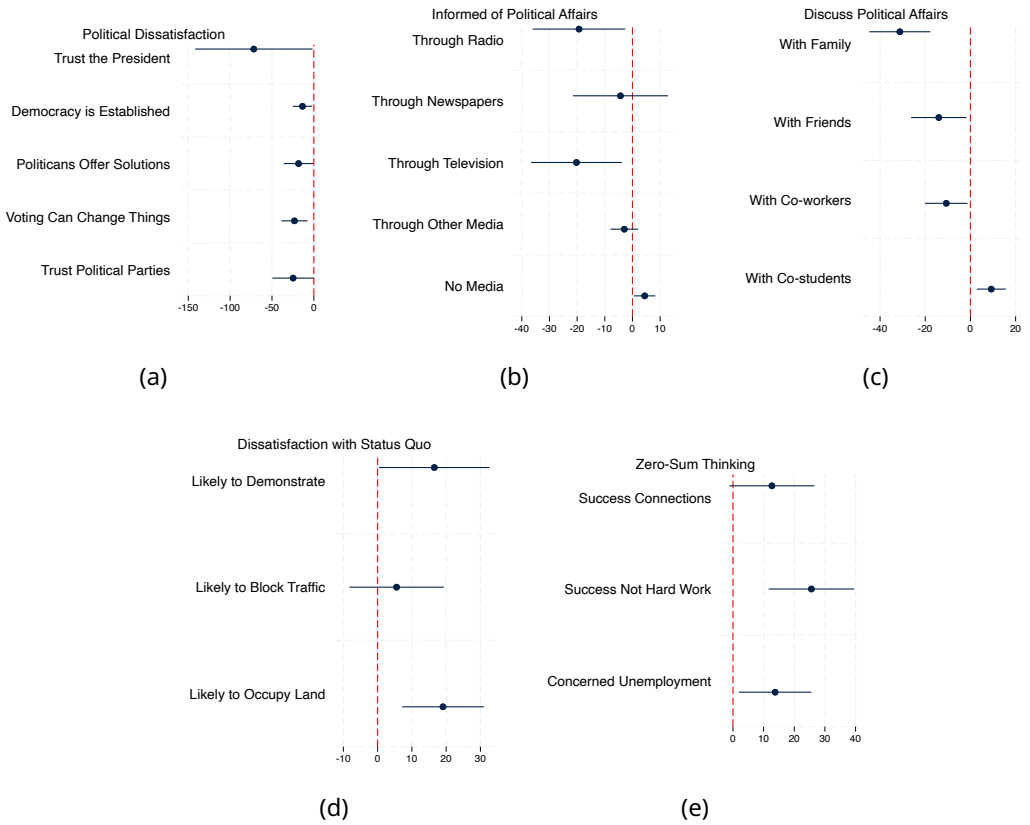
Notes: Main results on Chávez's vote share in 1998. Enslaved ancestry is instrumented with wheat suitability. Baseline controls include state indicators, a dummy for parishes with indigenous population, population in 1990; latitude, longitude, and their squares; distance to XVIIIth century villas, monasteries, missions, farms; the area of a parish devoted to large scale agriculture in the XVIIIth century; distance to contemporary ports, distance to shoreline, proportion of land classified as urban, and proportion of land devoted to agriculture. Additional controls include the unemployment rate, and an Anderson (2008) poverty index computed with physical house characteristics. Robust standard errors in parentheses. ACDE is estimated using the methodology from Acharya et al. (2016a) and assuming that luminosity and illiteracy are potential mediators. Robust standard errors in parentheses. For Column 5, standard errors are estimated via bootstrap with 1,000 repetitions. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Why Chávez?

I explore why Chávez was able to appeal to these voters by using the 1996 and 1997 waves of Latinobarometer. These waves include respondents' locations at the level of the municipality. I correlate the measure of enslaved ancestry at the level of the municipality with respondents' answers to several questions asked in the survey. These correlations are

reported in Tables 2.A.9-2.A.28 with increasingly strict specifications. The coefficient for enslaved ancestry is remarkably stable across each set of regressions. Figure 2.3 shows a summary of these findings under the strictest specifications.

I observe that respondents in locations with higher enslaved ancestry express greater dissatisfaction with political life in their country. This includes distrusting the president and political parties (panel (a)), low engagement with current political affairs reported in public media (panel (b)) and a low interest in discussing politics with family and peers (panel (c)). Individuals report a dissatisfaction with the status quo by expressing a likelihood to participate in demonstrations or to occupy land as part of a demonstration (panel (d)). Interestingly, respondents also exhibit zero-sum thinking (panel (e)). They are more likely to express that success is due to personal connections (though this is barely statistically significant), that success is not due to hard work, and to express concerns about becoming unemployed. It can be argued that, as an anti-system politician (see e.g., De Anda Casas, 2025a; Frajman, 2014), Chávez could plausibly appeal to these types of voters.



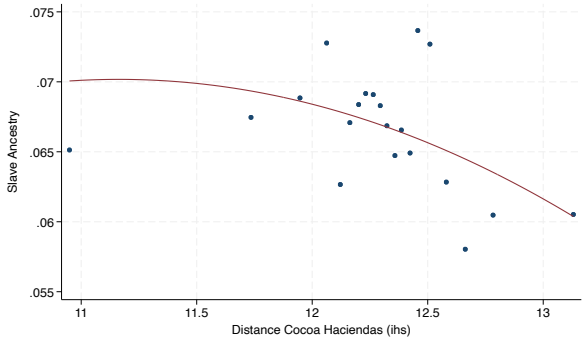
Notes: Coefficient plots for the correlation between enslaved ancestry and several variables from Latinobarometer with 95% confidence intervals. Data includes the 1996 and 1997 waves of Latinobarometer. All regressions include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television.

Figure 2.3: Enslaved Ancestry and Latinobarometer Respondents' Values

2.5.4 Locations Exposed to the Practice of Slavery

In this section I provide evidence to back the claim that the results are driven by individuals who have enslaved ancestry, and who are not necessarily living in locations where the institution of slavery prevailed during the colonial period – essentially, checking the long-term development and political consequences on locations exposed to the practice of slavery.

For this analysis I rely on a digitised version of the map of XVIIIth century Venezuela drawn by de Santos Cabrera (n.d.) as it details the location of cocoa-producing *haciendas* during the colonial period. This map is reproduced in Figure 2.B.1. I redefine the main independent variable as the distance from the centroid of each contemporary civil parish to the nearest XVIIIth century cocoa-producing *hacienda*. First, I confirm that individuals with enslaved ancestry are indeed not located in areas close to cocoa-producing *haciendas*. I check this via a battery of conditional correlations in Table 2.A.34. The strictest specification is represented via a binscatter plot in Figure 2.4. Regardless of the specification, the positive and statistically significant correlation suggests that parishes farther from cocoa-producing *haciendas* have higher shares of enslaved ancestry.



Notes: Binscatter plot of residualised enslaved ancestry correlated against distance to cocoa *haciendas*. The plot includes a quadratic fit line illustrating the relationship between the two variables. Enslaved ancestry is residualised with state indicators, a dummy for parishes with indigenous population, population in 1990; latitude, longitude, and their squares; distance to XVIIIth century villas, monasteries, missions, farms; the area of a parish devoted to large scale agriculture in the XVIIIth century; distance to contemporary ports, distance to shoreline, proportion of land classified as urban, and proportion of land devoted to agriculture.

Figure 2.4: Residualised Enslaved Ancestry and Distance to Cocoa *Haciendas*

Next, I test whether there is a causal relationship between exposure to slavery and the outcomes following the same strategy as in the previous sections, though here the instrument is the soil's cocoa suitability. Equations 2.3 and 2.4 represent the empirical strategy I follow. As above, $\mathbf{X_p}$ represents a vector of controls with γ and δ being their corresponding coefficients. The validity of the instrument follows from the discussion provided in section 2.4.1. The instrument's relevance is quantified in columns 3 and 4 of

Table 2.A.1. We observe a strong and statistically significant coefficient with the expected (negative) sign for β_3, a comfortable rejection of the underidentification null, and a strong enough performance of the instrument.

$$\text{DistCocoaHaciendas}_p = \beta_3 \text{CocoaSuitability} + \mathbf{X_p}\gamma' + \upsilon_p \tag{2.3}$$

$$y_p = \beta_4 \widehat{\text{DistCocoaHaciendas}}_p + \mathbf{X_p}\delta' + \epsilon_p \tag{2.4}$$

The first 3 columns of Table 2.4 group the development outcomes. Across the board, I see no effect of exposure to slavery. There seems to be a very small and positive effect on luminosity, but this is barely statistically significant. As a complement, Table 2.A.29 analyses the effect of exposure to slavery on schooling. Again, I see no significant effects across the board, with the exception of a barely significant positive effect on the number of children per school. Table 2.A.30 looks at disaggregated child mortality. Just as in Table 2.4, I see no effects on any of these variables. Table 2.A.31 analyses mortality rates, and reports no significant effects.

Table 2.4: Effect of Exposure to Institution of Slavery

	Development Outcomes			Votes for Chávez	
	(1) Luminosity (1998)	(2) Illiteracy (1990)	(3) Child Mortality (2000)	(4) Chávez Share (1998)	(5) Chávez Share (1998)
Distance Cocoa *Haciendas*	-1.260*	0.022	-0.023	-0.149	-0.089
	(0.688)	(0.046)	(0.018)	(0.116)	(0.099)
Observations	1098	1098	1098	1087	1087
Population	Yes	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes	Yes
Luminosity	No	No	No	No	Yes
Illiteracy	No	No	No	No	Yes
Additional Controls	No	No	No	Yes	Yes
Specification	2SLS	2SLS	2SLS	2SLS	2SLS

Notes: Distance to cocoa *haciendas* instrumented with cocoa suitability on all outcomes. Controls correspond to those from the regressions reported in Tables 2.1 and 2.3. Robust standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

Columns 4 and 5 of Table 2.4 look at the effect on the share of votes that Chávez received. The specification of column 4 is akin to that of column 1 of Table 2.3, wherein I do not control for luminosity and the illiteracy rate. Column 5 is akin to specification 4 in Table 2.3 where I do control for these other variables. Regardless of the specification, I observe no significant effect on Chávez's share.

To complement this set of results I redefine the measure of exposure to the institution of slavery to focus not on the *haciendas* where slaves were taken to, but rather the ports where they arrived. For a list of the ports of arrival, I rely on data from the Inter-American

Slave Trade Database, and the complementary Intra-American Slave Trade Database (Eltis, 1999). Despite not being exhaustive, this is a fairly commonly used resource (e.g., Nunn, 2008). As before, I define the main independent variable as the distance from the centroid of each civil parish to the closest port that used to be a port of arrival for slave journeys during the colonial period. As I do not have an instrument for the location of these ports, I only present correlational evidence.

Table 2.A.32 presents two sets of results. Panel A uses the baseline specification and finds no significant relations with the exception of column 3 that suggests that being closer to a slave port is related with a lower child mortality rate. Panel B changes the specification by dropping the distance to the closest contemporary port. Here, several relations emerge. Being closer to one of these ports is correlated with more luminosity, higher literacy and fewer votes for Chávez in the restricted specification of column 4. As suggested by Panel A, however, these relations are mostly due to these ports still being active today – overall, an unsurprising set of results. As an additional robustness check, I redefine the main independent variable of the section as the closest port of arrival or *hacienda*, and use the 2SLS strategy from equations 2.3 and 2.4. Results are unchanged and are reported in Table 2.A.33.

Taken together, these findings suggest that locations in Venezuela exposed to the slave trade during the colonial period are not causally worse-off, nor more likely to have supported Hugo Chávez in his initial presidential run. Stretching the strength of the evidence, it would seem these locations are even better-off today and less likely to have supported Chávez. The predominantly null results of this section set this paper apart from the classic studies in the literature (e.g., Acemoglu et al., 2012; Acharya et al., 2015, 2016b; Dell, 2010; Rocha et al., 2017). Instead, the comparatively stronger results of sections 2.5.1, 2.5.2, and 2.5.3 suggest that individuals with enslaved ancestry relocated (see Figure 2.4) and have persistently carried the burden of enslavement. This is expressed in lower economic development, lower human capital (proxied by illiteracy and low-productivity occupational specialisation), and worse health (see Table 2.1). Further, these conditions are not the result of recent economic downturns, but rather conditions that have persisted through time (see Table 2.2 and Figure 2.2). Hugo Chávez's initial candidacy was able to highlight these ills and gather the electoral support of these individuals (see Table 2.3).

2.6 Additional Robustness Checks

Changes to the similarity score threshold. I check for robustness of the results to changes along the names' similarity threshold generated by the matching algorithm I

implemented. For the main results of the paper I opted for a threshold of 85%–allowing for a false-positive rate (FPR) of 15%. I repeat the analyses for thresholds ranging from 80% to 99% i.e., 20% to 1% FPR. Results are presented in Figure 2.A.6. Panel (a) shows robustness of the first stage to changes along this thresholds. Panels (b), (c), and (d) show effects on the main development outcomes remain unchanged. Panel (e) also shows the results on Chávez's vote share are unchanged.

Choosing an alternative matching algorithm. As an alternative to the hybrid algorithm I use to match the slave surnames from Acosta Saignes (1984) to the contemporary list of surnames in Venezuela, I use a simpler algorithm that does not incorporate the phonetic representation of each n-gram.[37] For completeness and as above, I present results for a range of similarity scores in Figure 2.A.7. As above, I show robustness to changes along similarity thresholds ranging from 80% to 99%. Again, panel (a) shows robustness of the first stage; panels (b), (c), and (d) robustness of the main development outcomes; and panel (e) shows robustness of the effect on Chávez's vote share. Across the board, results remain unchanged, even if point estimates are slightly different.

Using maternal surnames. It is customary in Spanish-speaking countries to give newborns a paternal and a maternal surname. Usually, the surname that is passed down is the paternal surname, which tends to be the first of the two in individuals' names. The measure of enslaved ancestry is defined on individuals' first surnames. I check whether results are robust to defining the measure of enslaved ancestry on individuals' second surnames. All results are robust to this alternative definition, as Table 2.A.35 shows.

Winsorizing enslaved ancestry. To ensure the main results are not driven by outliers in the measure of enslaved ancestry, I winsorize the variable. I use three different thresholds at the top and bottom of the distribution of values: 1%, 5%, and 10%. Table 2.A.36 reports these results. Overall, coefficients retain their sign and statistical significance. The magnitude of the coefficients does change, however, with magnitudes becoming stronger with fewer outlier values.

Enslaved ancestry and recent migration. One relevant concern would be whether the locational sorting of individuals with enslaved ancestry is a recent phenomenon – undermining the historical component of the analysis. Obtaining Venezuelan migration data is difficult, but the 2001 census captures this information. I use the consolidated data from IPUMS to explore whether enslaved ancestry is related to measures of recent migration. A series of regressions reported in Table 2.A.37 suggest enslaved ancestry is uncorrelated with recent migration, regardless of whether it is internal or external migration. This allays concerns around selection on recent locational sorting based on

[37] The algorithm I implement for this analysis is the soundex algorithm.

individuals' surnames.

Spatially correlated standard errors. I check the sensitivity of the results to Conley (1999) standard errors. Rather than allowing the errors to be correlated for a few arbitrary distance thresholds, I estimate them for a whole range of values. These results are reported in Figure 2.A.8. Panel (a), (b), and (c) correspond to the development outcomes, and panel (d) corresponds to specification 4 of Table 2.3. As expected, the correction inflates the standard errors associated with the point estimates. Some specifications seem to be sensitive to this correction (e.g., panel (a) and (b)), while the results on child mortality and votes for Chávez seem unaffected by this correction.

Omitting potential bad controls. To address concerns around biases introduced by including potential intermediate outcomes as controls, Table 2.A.38 reports the results on development outcomes and Chávez's share with a smaller set of covariates. These covariates include state indicators, a parish's latitude, longitude and their square; the distance to the closest shore, and all XVIIIth century variables. All coefficients maintain their direction and have similar magnitudes. The statistical significance on the effect on Chávez's vote share disappears, however.

Alternative definition of instruments. Next, I redefine the instruments to test whether the results are driven by locations with high wheat suitability, low cocoa suitability, and their interaction. I present the first stage results in Table 2.A.39, and the second stage results in Table 2.A.40. Table 2.A.40 is divided into two panels. Panel A presents the results using the high-wheat, low-cocoa indicators, and their interaction, as instruments. Panel B only uses the interaction as an instrument since the first stage behind the results in Panel A is carried by the interaction term (see Table 2.A.39). Results are broadly in line with those from Section 2.5, but they are subject to a weak instruments problem from discretizing the instruments. These results reinforce the observation from Codazzi (1960) on subsistence farmers growing wheat in land that was not already used to produce other crops such as cocoa.

Correlational evidence and exploring heterogeneities. Table 2.A.41 presents reduced form regressions of enslaved ancestry on all the development outcomes, and Table 2.A.42 on votes for Chávez. These regressions could be subject to omitted variables bias, as discussed in Section 2.4 but are shown here for completeness. The coefficients point in the direction of the main results of the paper, though they differ in magnitude and levels of statistical significance. The even-numbered columns of Table 2.A.41 explore heterogeneities among locations close to *haciendas*. Interestingly, the strong negative relationship with luminosity is carried by parishes close to colonial *haciendas*. This is in contrast with the results from Columns 2-4 from Table 2.A.42. Here, locations far from

colonial *haciendas* show a positive relationship, while locations close to *haciendas* show a negative relation. At this stage, this analysis is only exploratory, but these heterogeneities could inform future identification strategies such as a spatial differences-in-differences analysis.

2.7 Conclusions and Limitations

This paper has studied the economic and political effects of the legacy of African slavery in Venezuela. By combining unique and novel data sources, and taking advantage of a unique onomastic norm for assigning surnames to enslaved people, I am able to identify individuals who have African enslaved ancestry in contemporary Venezuela. I followed a two-stage least-squares identification strategy, to identify observed effects including lower night light-density, higher illiteracy rates, and higher child-mortality rates, in areas with a higher concentration of people with enslaved ancestry. It is not the case that locations with higher enslaved ancestry have lower access to schools; this emphasises that the issues lie with the quality of what is being taught in these locations, rather than whether schools exist in these places. Also, it is not the case that adults are more likely to die in these locations. The effect can only be identified for children – the most vulnerable members of a community – rather than the population as a whole. I find suggestive evidence that the mechanism behind these effects is low human capital and low-productivity occupational specialisation. I present suggestive evidence that African enslaved ancestry is historically correlated with lower educational attainment, structural persistence in favour of low-productivity economic sectors such as subsistence agriculture, and low asset ownership.

Next, I identify an effect of African enslaved ancestry in support for Hugo Chávez during his first presidential run. This paper is the first to explore a historical explanation that causally identifies factors that contributed to the electoral victory of Hugo Chávez. The aforementioned measures for economic development do not fully account for the effect of enslaved ancestry on the votes for Chávez. Potential factors that could explain the rest of this effect include the persuasive appeal of Chávez as a populist candidate (De Anda Casas, 2025a; Frajman, 2014), explanations related to race and representation (e.g., Cannon, 2013; Herrera Salas, 2005; Ishibashi, 2003), or a combination thereof. Additionally, I find correlational evidence between the measure of enslaved ancestry and waves of Latinobarometer prior to Chávez's presidency that suggest that voters in locations with high enslaved ancestry distrusted politicians, thought politics was rigged, had low engagement with political affairs, and exhibited zero-sum thinking. These factors align with Chávez's anti-system rhetoric and contribute to explain why these individuals voted for him.

Individuals who have enslaved ancestry are not located in places where African slavery existed during the colonial period. This stylised fact is historically motivated and allows us to separate the effect of mobile factors of persistence, such as human capital, from place-based factors such as formal norms or laws. Following a similar identification strategy, I do not find any meaningful relationships between these development indicators and locations where slavery existed during the colony. These locations are also not more likely to have supported Hugo Chávez.

Altogether, these findings represent three key contributions. First, this is an empirical exploration of the long-term economic consequences of African slavery in Venezuela, which led to factors that contributed to the rise of *Chavismo*. It is the first paper in the literature on populism to provide such an explanation for the rise of a left-wing populist leader. Second, the uniqueness of the historical setting allows us to separate mobile and place-based factors of historical persistence: all of the effects are identified in locations where individuals carried with them the legacy of African enslavement, rather than in locations where enslaved labour was utilised during the colony. Third, these results emphasise enslavement's influence on those affected by it, and their kin. Venezuela was considered a fringe of the Spanish colonial empire (Mahoney, 2010). These results highlight the long shadow cast even by comparatively lighter colonial regimes.

The paper is subject to some limitations in its current form. These include, for instance, incorporating the new techniques developed by Conley and Kelly (2025) to account for spatially correlated effects and historical persistence. The Maisanta database provides a wealth of information that deserves further examination. Some of the adjustments explored in Güell, Mora, and Telmer (2013) could be incorporated when analysing surnames in Venezuela. For instance, additional robustness checks could include 'discounting' more common surnames, or trying to account for different cohorts. As it stands, the current persistence analysis relies on consolidated census waves from IPUMS. The Venezuelan censuses of 1873, 1881, 1920, 1926, 1936, 1941, and 1950 exist but have not been digitised yet. Once digitised, these additional data points could be incorporated into the analysis.

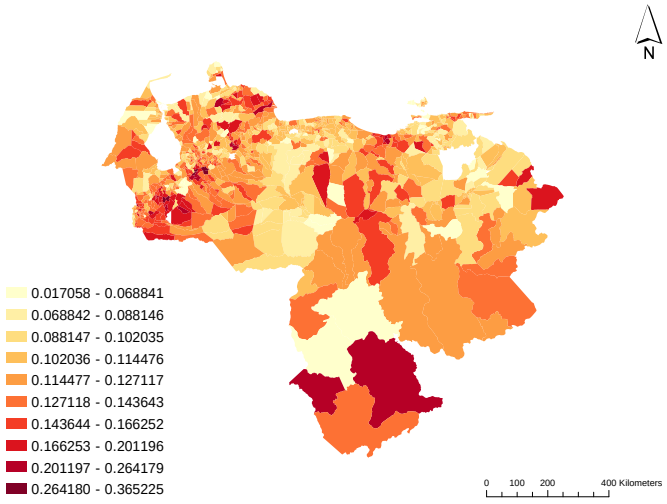
The paper has also remained agnostic about the several armed conflicts that took place in Venezuela during the XIXth and XXth centuries, and how these conflicts could have impacted patterns of internal migration. These conflicts include for instance the War of Independence (1811-1823) fought against Spain, the Federal War (1859-1863), or the Liberation Revolution (1901-1903). A broader, but important open question is disentangling the role of geographic determinants from mechanisms such as human capital on all outcome variables. Did individuals with low human capital establish themselves in land with low-productivity (via positive assortative matching), and are they poor today because

of the persistent low human capital, because of the low-productivity land, or a combination of both? Is geography destiny? I leave addressing these limitations for future versions of this paper, and future work.

Appendix

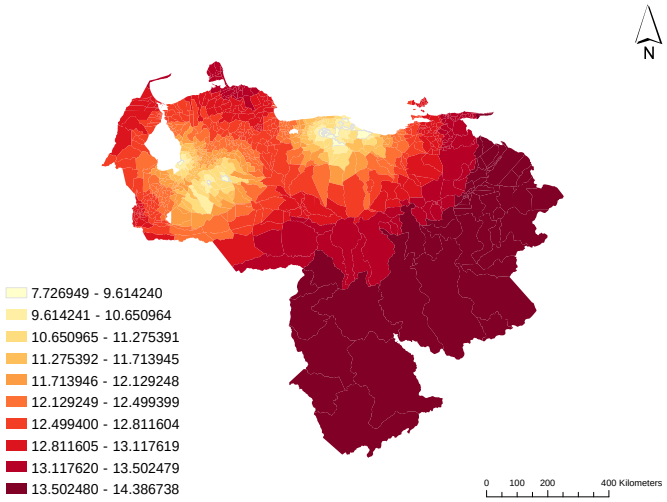
2.A Additional Figures and Tables

2.A.1 Appendix Figures



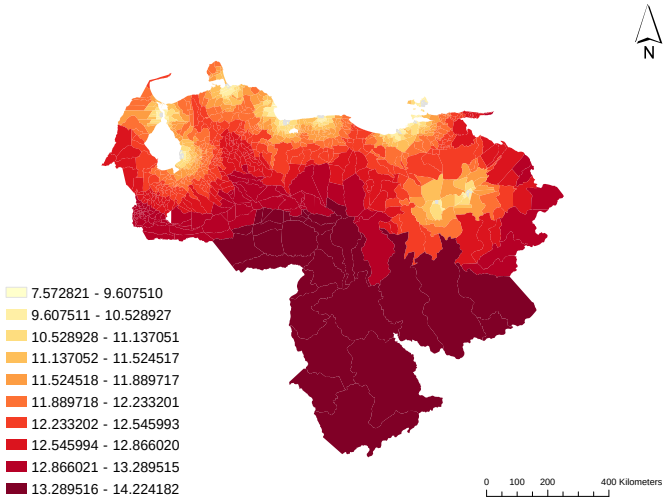
Notes: Map denoting proportion of individuals with enslaved ancestry in each civil parish of Venezuela. Venezuela's shapefile is from Porto Tapiquén (2020). All maps were processed in ArcGIS Pro.

Figure 2.A.1: Enslaved Ancestry



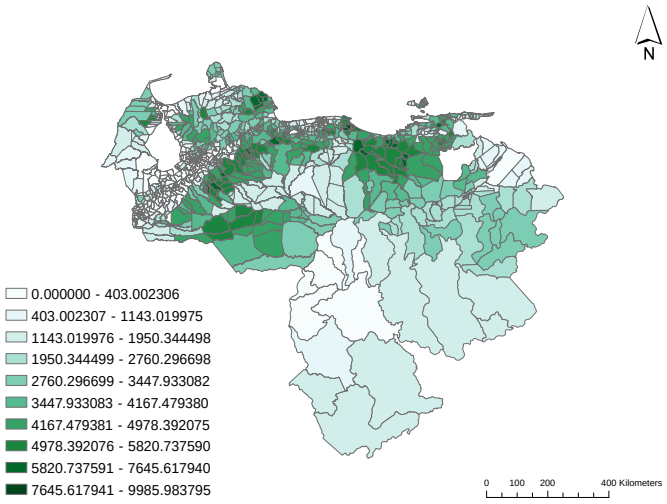
Notes: Map denoting distance from each civil parish to closes cocoa *hacienda* in Venezuela. Location of *haciendas* is from de Santos Cabrera (n.d.). All distances are transformed via the i.h.s.. Venezuela's shapefile is from Porto Tapiquén (2020). All maps were processed in ArcGIS Pro.

Figure 2.A.2: Distance to Cocoa *Haciendas*



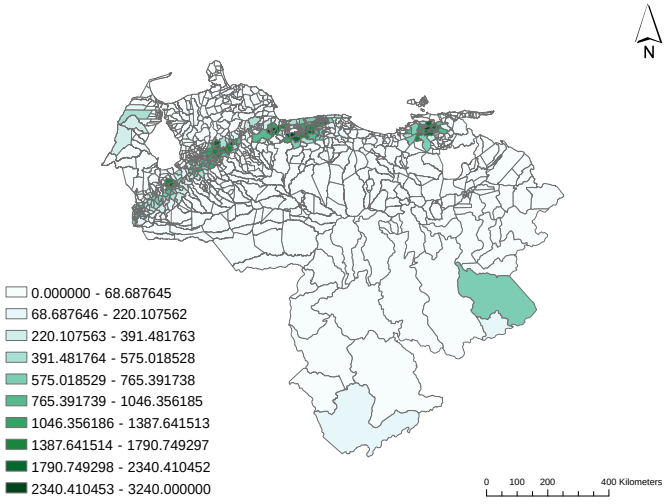
Notes: Map denoting distance from each civil parish to closes port of arrival of enslaved individuals. Ports are from the Trans-Atlantic and Inter-America slave trade databases (Eltis, 1999). All distances are transformed via the i.h.s.. Venezuela's shapefile is from Porto Tapiquén (2020). All maps were processed in ArcGIS Pro.

Figure 2.A.3: Distance to Ports Used in the Slave Trade



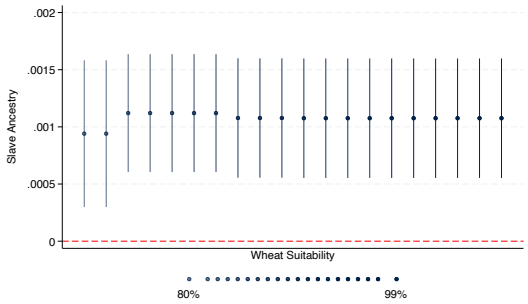
Notes: Map denoting average cocoa suitability at the level of the parish. Data is from the FAO as detailed in Section 2.3. Venezuela's shapefile is from Porto Tapiquén (2020). All maps were processed in ArcGIS Pro.

Figure 2.A.4: Cocoa Suitability in Venezuela

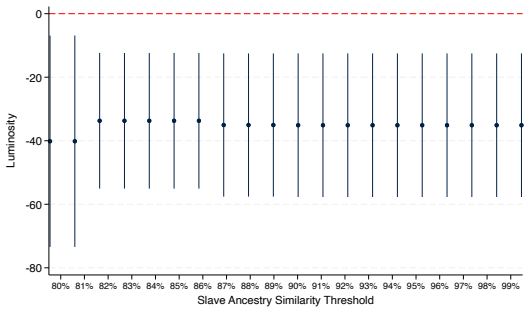


Notes: Map denoting average wheat suitability at the level of the parish. Data is from the FAO as detailed in Section 2.3. Venezuela's shapefile is from Porto Tapiquén (2020). All maps were processed in ArcGIS Pro.

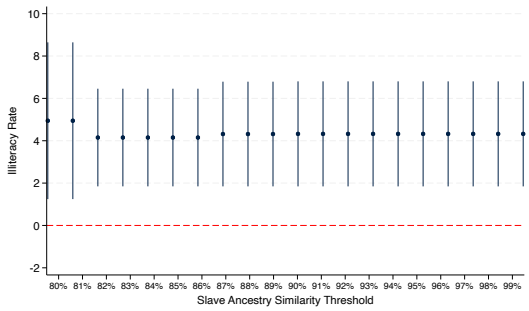
Figure 2.A.5: Wheat Suitability in Venezuela



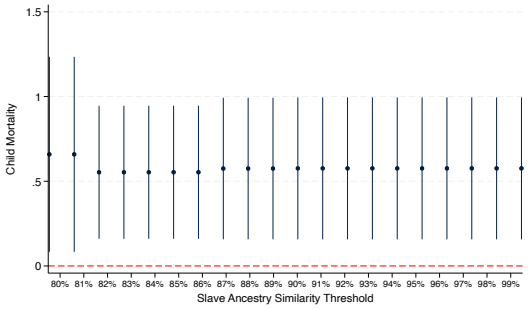
(a) First Stage



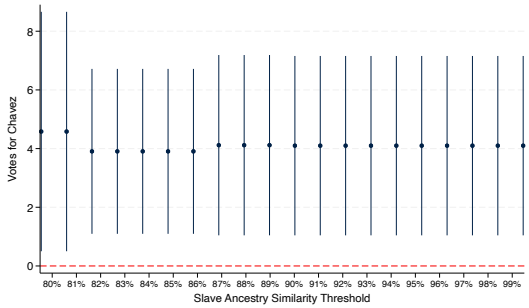
(b) Luminosity



(c) Illiteracy



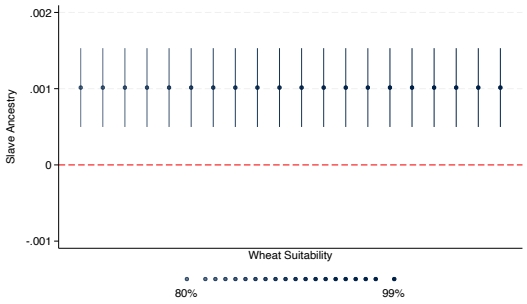
(d) Child Mortality



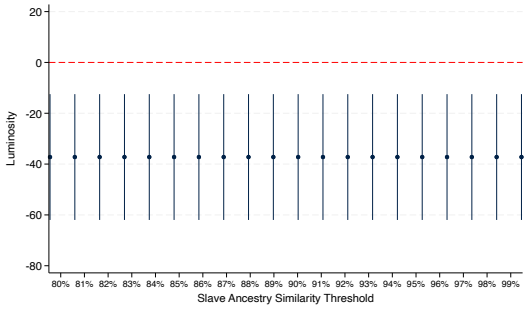
(e) Votes for Chávez

Notes: Figures probing the robustness of the main results after applying changes along the similarity threshold used to identify individuals with enslaved ancestry. Thresholds are changed from 80% to 99%. The main results of the paper use 85% as the threshold.

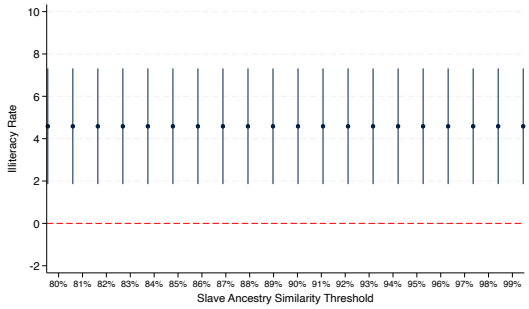
Figure 2.A.6: Robustness to Changes Along Similarity Threshold on Main Outcomes



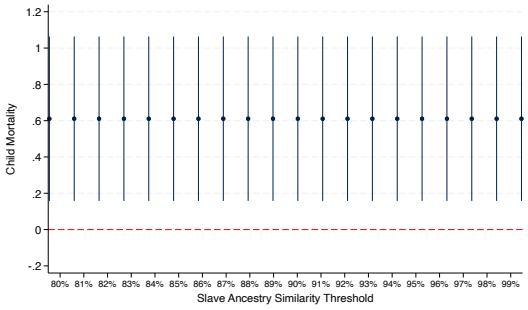
(a) First Stage



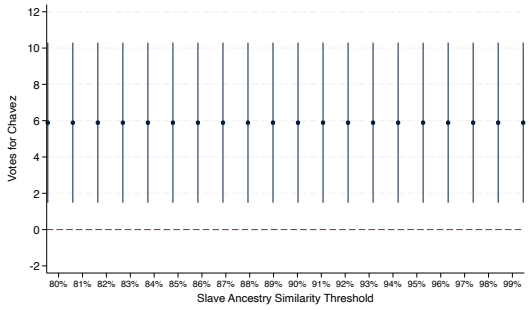
(b) Luminosity



(c) Illiteracy



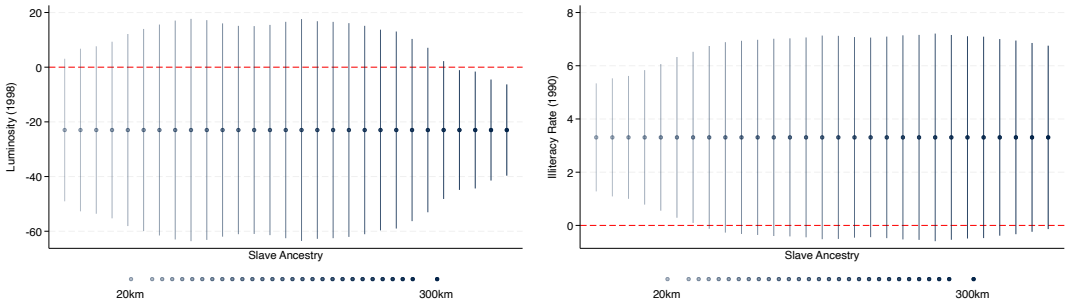
(d) Child Mortality



(e) Votes for Chávez

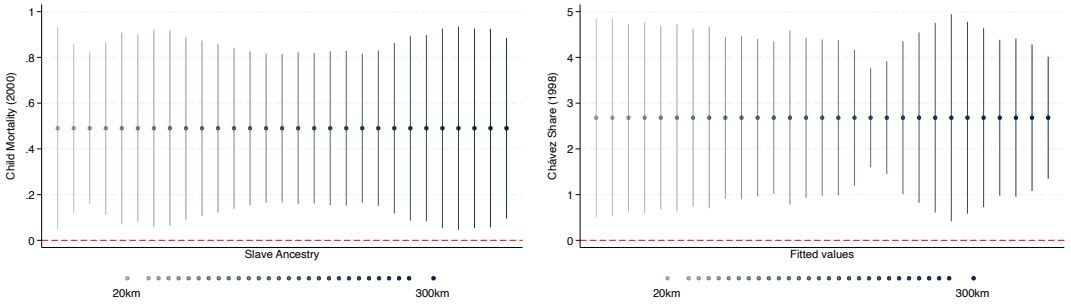
Notes: Figures probing the robustness of the main results after applying an alternative matching algorithm. This algorithm does not incorporate the phonetic representation of n-grams. These figures also probe the robustness of the results to alternative similarity thresholds from this algorithm, ranging from 80% to 99%.

Figure 2.A.7: Robustness to Using a Different Matching Algorithm and Changing Similarity Thresholds on Main Outcomes



(a) Luminosity

(b) Illiteracy



(c) Child Mortality

(d) Votes for Chávez

Notes: Figures probing the robustness of the main results to spatially correlated standard errors along distances ranging from 20km to 300km.

Figure 2.A.8: Robustness to Conley Standard Errors

2.A.2 Appendix Tables

Table 2.A.1: First Stage Regressions

	(1) Enslaved Ancestry	(2) Enslaved Ancestry	(3) Distance Cocoa *Haciendas*	(4) Distance Cocoa *Haciendas*
Wheat Suitability	0.001***	0.001***		
	(0.000)	(0.000)		
Cocoa Suitability			-0.024***	-0.017***
			(0.007)	(0.006)
Observations	1099	1098	1098	1091
Underidentification K-P Stat	13.826	10.349	11.843	6.936
Underidentification P-Val	0.000	0.001	0.001	0.008
Weak Idenficiation Stat	13.731	9.979	11.856	6.666
Population	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes
Contemporary Distances	Yes	Yes	Yes	Yes
Luminosity	No	Yes	No	Yes
Illiteracy Rate	No	Yes	No	Yes
Additional Controls	No	Yes	No	Yes

Notes: For the first stage results, enslaved ancestry is instrumented with wheat suitability and distance to cocoa *haciendas* with cocoa suitability. Baseline controls include state indicators, a dummy for parishes with indigenous population, population in 1990; latitude, longitude, and their squares; distance to XVIIIth century villas, monasteries, missions, farms; the area of a parish devoted to large scale agriculture in the XVIIIth century; distance to contemporary ports, distance to shoreline, proportion of land classified as urban, and proportion of land devoted to agriculture. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.2: Reduced Form Regressions of the Instruments on Main Outcomes

Panel A: Wheat Suitability	(1) Luminosity (1998)	(2) Illiteracy (1990)	(3) Child Mortality (2000)	(4) Chávez's Share (1998)
Wheat Suitability	-0.031**	0.004***	0.001***	0.005***
	(0.012)	(0.001)	(0.000)	(0.001)
Observations	1098	1098	1098	1087
Population	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes
Municipality Indicators	No	No	No	No
Latitude & Longitude	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes
Panel B: Cocoa Suitability	(1) Luminosity (1998)	(2) Illiteracy (1990)	(3) Child Mortality (2000)	(4) Chávez's Share (1998)
Cocoa Suitability	0.021	-0.000	0.000	0.003*
	(0.013)	(0.001)	(0.000)	(0.002)
Observations	1098	1098	1098	1087
Population	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes
Municipality Indicators	No	No	No	No
Latitude & Longitude	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes

Notes: Reduced form regressions of instruments on main outcomes. Panel A focuses on wheat suitability, and Panel B on cocoa suitability. Baseline controls include state indicators, a dummy for parishes with indigenous population, population in 1990; latitude, longitude, and their squares; distance to XVIIIth century villas, monasteries, missions, farms; the area of a parish devoted to large scale agriculture in the XVIIIth century; distance to contemporary ports, distance to shoreline, proportion of land classified as urban, and proportion of land devoted to agriculture. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.3: Effect of Enslaved Ancestry on Schooling

	(1) Primary School Enrollment	(2) Secondary School Enrollment	(3) Preschool Enrollment	(4) Children Per School
Enslaved Ancestry	0.210	1.242	0.653	731.709
	(0.493)	(2.537)	(0.400)	(3006.624)
Observations	1099	1099	1099	1099
Population	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes
Specification	2SLS	2SLS	2SLS	2SLS

Notes: Effect of enslaved ancestry instrumented with wheat suitability on schooling outcomes. Baseline controls include state indicators, a dummy for parishes with indigenous population, population in 1990; latitude, longitude, and their squares; distance to XVIIIth century villas, monasteries, missions, farms; the area of a parish devoted to large scale agriculture in the XVIIIth century; distance to contemporary ports, distance to shoreline, proportion of land classified as urban, and proportion of land devoted to agriculture. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.4: Effect of Enslaved Ancestry on Disaggregated Child Mortality Rates

	(1) Pre-natal Mortality (2000)	(2) Infant Mortality (2000)	(3) Older Children Mortality (2000)
Enslaved Ancestry	0.034**	0.056**	0.463**
	(0.015)	(0.027)	(0.198)
Observations	1098	1098	1098
Population	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes
Specification	2SLS	2SLS	2SLS

Notes: Effect of enslaved ancestry instrumented with wheat suitability on disaggregated child mortality rates. Baseline controls include state indicators, a dummy for parishes with indigenous population, population in 1990; latitude, longitude, and their squares; distance to XVIIIth century villas, monasteries, missions, farms; the area of a parish devoted to large scale agriculture in the XVIIIth century; distance to contemporary ports, distance to shoreline, proportion of land classified as urban, and proportion of land devoted to agriculture. Robust standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

Table 2.A.5: Effect of Enslaved Ancestry on Mortality Rates

	(1) Preventable Deaths	(2) Highly Preventable Deaths	(3) Moderately Preventable Deaths	(4) Non-Preventable Deaths
Enslaved Ancestry	-0.055	-0.024	-0.027	0.004
	(0.145)	(0.021)	(0.069)	(0.042)
Observations	1099	1099	1099	1099
Population	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes
Specification	2SLS	2SLS	2SLS	2SLS

Notes: Effect of enslaved ancestry instrumented with wheat suitability on mortality rates. Baseline controls include state indicators, a dummy for parishes with indigenous population, population in 1990; latitude, longitude, and their squares; distance to XVIIIth century villas, monasteries, missions, farms; the area of a parish devoted to large scale agriculture in the XVIIIth century; distance to contemporary ports, distance to shoreline, proportion of land classified as urban, and proportion of land devoted to agriculture. Robust standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

Table 2.A.6: Effect of Enslaved Ancestry on Votes for Chávez with Child Mortality Rate

	(1)	(2)
Enslaved Ancestry	3.906**	3.884**
	(1.709)	(1.723)
Observations	1087	1087
Luminosity	Yes	Yes
Illiteracy Rate	Yes	Yes
Child Mortality Rate	No	Yes
Population	Yes	Yes
Indigenous PP Dummy	Yes	Yes
State Indicators	Yes	Yes
Latitude & Longitude	Yes	Yes
XVIIIc. Variables	Yes	Yes
Land Cover	Yes	Yes
Additional Controls	Yes	Yes
Specification	2SLS	2SLS

Notes: Effect of enslaved ancestry instrumented with wheat suitability on votes for Chávez in 1998. Controls correspond to those detailed in Table 2.3 with the addition of child mortality ratio. Robust standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

Table 2.A.7: Persistence of Enslaved Ancestry and Economic Structure by Industry

Panel A				
	Primary Industries		Secondary Industries	
	(1) Agriculture	(2) Mining	(3) Manufacturing	(4) Construction
Enslaved Ancestry	2.185***	0.018	-0.621***	-0.231**
	(0.471)	(0.054)	(0.129)	(0.105)
Panel B				
	Tertiary Industries			
	(1) Electricity	(2) Retail	(3) Hotels and Restaurants	(4) Transport
Enslaved Ancestry	-0.058***	-0.318**	-0.021	-0.117**
	(0.016)	(0.126)	(0.035)	(0.054)
Panel C				
	Tertiary Industries			
	(1) Financial Services	(2) Public Admin	(3) Business Services	(4) Education
Enslaved Ancestry	-0.042**	-0.052	-0.085***	-0.283***
	(0.016)	(0.044)	(0.028)	(0.081)
Panel D				
	Tertiary Industries			
	(1) Health Services	(2) Other Services	(3) Private Hh Services	
Enslaved Ancestry	-0.116**	-0.163***	-0.040	
	(0.047)	(0.050)	(0.043)	

Notes: All specifications match those reported in Table 2.2. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.8: Gelbach Decomposition of Sequentially Added Covariates in Table 2.3

	(1)	(2)
Enslaved Ancestry	3.906**	3.884**
	(1.709)	(1.723)
	(1)	(2)
Luminosity	-0.882***	-0.878*
	(0.320)	(0.478)
Illiteracy (1990)	-0.240	-0.239
	(0.258)	(0.238)
Sum of Impacts	-1.122**	-1.098*
	(0.519)	(0.618)
Child Mortality (2000)		0.020
		(0.105)
Observations	1087	1087

Notes: Gelbach decomposition on the analysis from Table 2.3 accounting for the change in the coefficient of enslaved ancestry once luminosity, illiteraty, or child mortality are added as controls. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.9: Correlation Between Enslaved Ancestry and Distrusting the President

	Trust President			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	-77.820*	-72.562*	-77.657*	-71.680*
	(41.508)	(41.440)	(42.230)	(42.537)
Observations	709	695	709	695
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.10: Correlation Between Enslaved Ancestry and Whether Democracy is Established in Venezuela

	Democracy Established			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	-12.836*	-13.479*	-12.980*	-13.531*
	(6.902)	(7.051)	(6.885)	(7.026)
Observations	2133	2109	2133	2109
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.11: Correlation Between Enslaved Ancestry and Whether Politicians Offer Solutions

	Politicians Offer Solutions			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	-15.845	-19.655*	-14.548	-18.222*
	(10.670)	(10.665)	(10.713)	(10.696)
Observations	1457	1442	1457	1442
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.12: Correlation Between Enslaved Ancestry and Whether Voting Changes Things

	Voting Can Change Things			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	-23.475**	-23.523**	-23.044**	-23.144**
	(9.423)	(9.489)	(9.446)	(9.501)
Observations	2001	1973	2001	1973
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.13: Correlation Between Enslaved Ancestry and Trusting Political Parties

	Trust Political Parties			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	-26.404*	-23.205	-28.374*	-24.732
	(15.080)	(15.009)	(15.112)	(15.039)
Observations	2182	2154	2182	2154
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.14: Correlation Between Enslaved Ancestry and Getting Informed of Political Affairs Through the Radio

	Info. Pol. Affairs Radio			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	-22.205**	-20.202**	-21.064**	-19.315*
	(10.251)	(10.290)	(10.140)	(10.158)
Observations	1500	1485	1500	1485
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.15: Correlation Between Enslaved Ancestry and Getting Informed of Political Affairs Through Newspapers

	Info. Pol. Affairs Newspapers			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	-11.493	-7.521	-6.831	-4.338
	(10.585)	(10.551)	(10.374)	(10.442)
Observations	1500	1485	1500	1485
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.16: Correlation Between Enslaved Ancestry and Getting Informed of Political Affairs Through Television

	Info. Pol. Affairs TV			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	-23.636**	-23.692**	-19.274*	-20.244**
	(9.933)	(9.930)	(9.934)	(9.953)
Observations	1500	1485	1500	1485
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.17: Correlation Between Enslaved Ancestry and Getting Informed of Political Affairs Through Other Media

	Info. Pol. Affairs Other			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	-2.232	-2.126	-2.977	-2.945
	(2.904)	(2.946)	(2.998)	(3.039)
Observations	1500	1485	1500	1485
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.18: Correlation Between Enslaved Ancestry and Getting Informed of Political Affairs Through No Media

	Info. Pol. Affairs None			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	5.115**	5.146**	4.276*	4.424*
	(2.395)	(2.427)	(2.305)	(2.324)
Observations	1500	1485	1500	1485
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.19: Correlation Between Enslaved Ancestry and Getting Informed of Political Affairs Through Family

	Info. Pol. Affairs Family			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	-29.552***	-30.961***	-30.006***	-31.236***
	(8.153)	(8.272)	(8.061)	(8.197)
Observations	1500	1485	1500	1485
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.20: Correlation Between Enslaved Ancestry and Getting Informed of Political Affairs Through Friends

	Info. Pol. Affairs Friends			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	-12.274*	-12.582*	-13.804*	-13.998*
	(7.210)	(7.406)	(7.253)	(7.453)
Observations	1500	1485	1500	1485
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.21: Correlation Between Enslaved Ancestry and Getting Informed of Political Affairs Through Co-Workers

	Info. Pol. Affairs Co-Workers			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	-11.805**	-11.630**	-10.848*	-10.679*
	(5.514)	(5.627)	(5.593)	(5.711)
Observations	1500	1485	1500	1485
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.22: Correlation Between Enslaved Ancestry and Getting Informed of Political Affairs Through Co-Students

	Info. Pol. Affairs Co-Students			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	8.369**	8.976**	8.932**	9.249**
	(3.734)	(3.786)	(3.845)	(3.897)
Observations	1500	1485	1500	1485
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.23: Correlation Between Enslaved Ancestry and Likelihood to Participate in Demonstrations

	Participate Demonstrations			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	11.738	14.410	14.555	16.549*
	(9.726)	(9.781)	(9.775)	(9.793)
Observations	1464	1450	1464	1450
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.24: Correlation Between Enslaved Ancestry and Likelihood of Blocking Traffic as Demonstration

	Participate Blocking Traffic			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	3.742	4.957	4.897	5.543
	(8.338)	(8.250)	(8.476)	(8.358)
Observations	1465	1451	1465	1451
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.25: Correlation Between Enslaved Ancestry and Likelihood of Occupying Land as a Demonstration

	Participate Occupying Land			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	18.816***	20.212***	17.774**	19.077***
	(7.191)	(7.165)	(7.263)	(7.234)
Observations	1443	1429	1443	1429
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

Table 2.A.26: Correlation Between Enslaved Ancestry and Believing Success is Due to Connections

	Success Due to Connections			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	12.312	13.730	11.308	12.700
	(8.340)	(8.396)	(8.343)	(8.413)
Observations	2049	2022	2049	2022
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

Table 2.A.27: Correlation Between Enslaved Ancestry and Believing Success is Not Due to Hard Work

	Success Not Due to Hard Work			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	22.811***	25.398***	23.226***	25.589***
	(8.326)	(8.418)	(8.345)	(8.458)
Observations	2081	2054	2081	2054
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

Table 2.A.28: Correlation Between Enslaved Ancestry and Expressing Concern About Unemployment

	Concerned About Unemployment			
	(1)	(2)	(3)	(4)
Enslaved Ancestry	12.079*	12.725*	12.889*	13.769*
	(7.131)	(7.270)	(7.053)	(7.196)
Observations	2173	2148	2173	2148
State Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Basic Demographics	No	Yes	No	Yes
Assets & Services	No	No	Yes	Yes

Notes: Correlation between enslaved ancestry and outcome from 1996 and 1997 waves of Latinobarometer. Controls include state and wave indicators, as well as respondents' age, gender, education, and asset ownership as controls. Assets include access to hot water, sewage, drinking water, a second house, car, telephone, washer, computer, fridge, and television. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.29: Effect of Exposure to Slavery on Schooling

	(1) Primary School Enrollment	(2) Secondary School Enrollment	(3) Preschool Enrollment	(4) Children Per School
Distance Cocoa *Haciendas*	-0.023	0.152	0.014	-165.667*
	(0.027)	(0.280)	(0.018)	(97.487)
Observations	1099	1099	1099	1099
Population	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes
Specification	2SLS	2SLS	2SLS	2SLS

Notes: Effect of distance to cocoa *haciendas* instrumented by cocoa suitability on schooling. Baseline controls include state indicators, a dummy for parishes with indigenous population, population in 1990; latitude, longitude, and their squares; distance to XVIIIth century villas, monasteries, missions, farms; the area of a parish devoted to large scale agriculture in the XVIIIth century; distance to contemporary ports, distance to shoreline, proportion of land classified as urban, and proportion of land devoted to agriculture. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.30: Effect of Exposure to Slavery on Disaggregated Child Mortality Rates

	(1) Pre-natal Mortality (2000)	(2) Infant Mortality (2000)	(3) Older Children Mortality (2000)
Distance Cocoa *Haciendas*	-0.001	-0.003	-0.019
	(0.001)	(0.002)	(0.015)
Observations	1098	1098	1098
Population	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes
Specification	2SLS	2SLS	2SLS

Notes: Effect of distance to cocoa *haciendas* instrumented by cocoa suitability on disaggregated child mortality. Baseline controls include state indicators, a dummy for parishes with indigenous population, population in 1990; latitude, longitude, and their squares; distance to XVIIIth century villas, monasteries, missions, farms; the area of a parish devoted to large scale agriculture in the XVIIIth century; distance to contemporary ports, distance to shoreline, proportion of land classified as urban, and proportion of land devoted to agriculture. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.31: Effect of Exposure to Slavery on Mortality Rates

	(1) Preventable Deaths	(2) Highly Preventable Deaths	(3) Moderately Preventable Deaths	(4) Non-Preventable Deaths
Distance Cocoa *Haciendas*	-0.018	-0.002	-0.009*	-0.005
	(0.011)	(0.001)	(0.005)	(0.003)
Observations	1099	1099	1099	1099
Population	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes
Specification	2SLS	2SLS	2SLS	2SLS

Notes: Effect of distance to cocoa *haciendas* instrumented by cocoa suitability on disaggregated mortality rates. Baseline controls include state indicators, a dummy for parishes with indigenous population, population in 1990; latitude, longitude, and their squares; distance to XVIIIth century villas, monasteries, missions, farms; the area of a parish devoted to large scale agriculture in the XVIIIth century; distance to contemporary ports, distance to shoreline, proportion of land classified as urban, and proportion of land devoted to agriculture. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.32: Ports Exposed to Institution of Slavery and Main Outcomes

Panel A: Controlling for distance to contemporary ports	Development Outcomes			Votes for Chávez	
	(1) Luminosity (1998)	(2) Illiteracy (1990)	(3) Child Mortality (2000)	(4) Chávez Share (1998)	(5) Chávez Share (1998)
Distance Slave Ports	0.053	-0.001	0.001**	0.000	0.002
	(0.078)	(0.005)	(0.001)	(0.008)	(0.007)
Panel B: Not controlling for distance to contemporary ports	**Development Outcomes**			**Votes for Chávez**	
	(1) Luminosity (1998)	(2) Illiteracy (1990)	(3) Child Mortality (2000)	(4) Chávez Share (1998)	(5) Chávez Share (1998)
Distance Slave Ports	-0.215***	0.015***	0.000	-0.017**	-0.008
	(0.071)	(0.005)	(0.000)	(0.006)	(0.006)
Observations	1098	1098	1098	1087	1087
Population	Yes	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes	Yes
Luminosity	No	No	No	No	Yes
Illiteracy	No	No	No	No	Yes
Additional Controls	No	No	No	Yes	Yes
Specification	OLS	OLS	OLS	OLS	OLS

Notes: Effect of distance to ports of arrival on main outcomes. Controls in Panel A correspond to those in Table 2.4. Panel B drops the distance to contemporary ports. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.33: Effect of Exposure to Closest Port or *Hacienda*

	Development Outcomes			Votes for Chávez	
	(1) Luminosity (1998)	(2) Illiteracy (1990)	(3) Child Mortality (2000)	(4) Chávez Share (1998)	(5) Chávez Share (1998)
Distance Inst. Slavery	-1.243*	0.021	-0.023	-0.141	-0.096
	(0.678)	(0.046)	(0.018)	(0.103)	(0.089)
Observations	1098	1098	1098	1087	1087
Population	Yes	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes	Yes
Luminosity	No	No	No	No	Yes
Illiteracy	No	No	No	No	Yes
Additional Controls	No	No	No	Yes	Yes
Specification	2SLS	2SLS	2SLS	2SLS	2SLS

Notes: Effect of distance to cocoa *haciendas* or ports of arrival instrumented by cocoa suitability on main outcomes. Controls correspond to those in Table 2.4. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.34: Conditional Correlation Between Enslaved Ancestry and Distance to Cocoa *Haciendas*

	(1) Distance Cocoa Haciendas	(2) Distance Cocoa Haciendas	(3) Distance Cocoa Haciendas	(4) Distance Cocoa Haciendas	(5) Distance Cocoa Haciendas
Enslaved Ancestry	-0.815	-0.825	-1.274**	-1.202**	-1.242**
	(0.601)	(0.594)	(0.594)	(0.590)	(0.587)
Observations	1102	1098	1098	1102	1098
Population	Yes	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes	Yes
Land Cover	No	Yes	Yes	No	Yes
Latitude & Longitude	No	No	Yes	Yes	Yes
Contemporary Distances	No	Yes	No	Yes	Yes
Specification	OLS	OLS	OLS	OLS	OLS

Notes: Correlation between enslaved ancestry and distance to cocoa *haciendas* under increasingly strict specifications. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.35: Robustness to Measuring Enslaved Ancestry Through Maternal Surnames

	Development Outcomes			Chávez's Share	
	(1) Luminosity (1998)	(2) Illiteracy (1990)	(3) Child Mortality (2000)	(4) Chávez Share (1998)	(5) Chávez Share (1998)
Enslaved Ancestry (Maternal)	-32.800***	4.039***	0.538***	2.541*	3.353***
	(11.864)	(1.210)	(0.207)	(1.328)	(1.248)
Observations	1098	1098	1098	1087	1087
Population	Yes	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes	Yes
Additional Controls	No	No	No	Yes	Yes
Luminosity	No	No	No	No	Yes
Illiteracy Rate	No	No	No	No	Yes
Specification	2SLS	2SLS	2SLS	2SLS	2SLS

Notes: Robustness of main results to using maternal surnames as opposed to paternal surnames. Specifications correspond to those reported in Tables 2.1 and 2.3. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.36: Robustness to Winsorizing Measure of Enslaved Ancestry

Panel A: 1%	(1) Luminosity (1998)	(2) Illiteracy (1990)	(3) Child Mortality (2000)	(4) Chávez Share (1998)
Enslaved Ancestry	-40.807**	5.025***	0.669**	4.406**
	(16.360)	(1.780)	(0.298)	(2.125)
Observations	1098	1098	1098	1087
Population	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes
Specification	2SLS	2SLS	2SLS	2SLS
Panel B: 5%				
	(1) Luminosity (1998)	(2) Illiteracy (1990)	(3) Child Mortality (2000)	(4) Chávez Share (1998)
Enslaved Ancestry	-58.751**	7.235***	0.964**	6.388*
	(25.355)	(2.738)	(0.449)	(3.314)
Observations	1098	1098	1098	1087
Population	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes
Specification	2SLS	2SLS	2SLS	2SLS
Panel C: 10%				
	(1) Luminosity (1998)	(2) Illiteracy (1990)	(3) Child Mortality (2000)	(4) Chávez Share (1998)
Enslaved Ancestry	-80.633**	9.930**	1.323**	8.564*
	(36.769)	(3.980)	(0.640)	(4.632)
Observations	1098	1098	1098	1087
Population	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes
Specification	2SLS	2SLS	2SLS	2SLS

Notes: Robustness of main results to winsorizing measure of enslaved ancestry to curb the influence of outliers in results. Panel A winsorizes the top and bottom 1%, Panel B 5% and Panel C 10%. Specifications correspond to those reported in Tables 2.1 and 2.3. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.37: Enslaved Ancestry and Recent Migration

	(1) Migrant	(2) External Migrant	(3) Internal Migrant
Enslaved Ancestry	0.406	0.120	0.286
	(0.502)	(0.089)	(0.475)
Observations	157	157	157
Population	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes
Year Indicators	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes
Contemporary Land Coverage	Yes	Yes	Yes
Dist. Shore and Ports	Yes	Yes	Yes
Specification	OLS	OLS	OLS

Notes: Enslaved ancestry at the level of the municipality correlated with measures of recent migration from the 2001 census. This analysis attempts to allay concerns around recent selection based on surnames. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.38: Main Results Removing Potential Intermediate Outcomes from Regressions

	Development Outcomes			Chávez's Share
	(1) Luminosity (1998)	(2) Illiteracy (1990)	(3) Child Mortality (2000)	(4) Chávez Share (1998)
Enslaved Ancestry	-68.112***	4.956***	0.653**	1.605
	(20.581)	(1.630)	(0.280)	(1.244)
Observations	1103	1103	1103	1092
Population	No	No	No	No
Indigenous PP Dummy	No	No	No	No
State Indicators	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes
Land Cover	No	No	No	No
Additional Controls	No	No	No	No
Luminosity	No	No	No	No
Illiteracy Rate	No	No	No	No
Specification	2SLS	2SLS	2SLS	2SLS

Notes: Robustness of main results instrumenting enslaved ancestry with wheat suitability and removing all controls that could be intermediate outcomes. Controls include state indicators, XVIIIth century variables, latitude longitude and their squares. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.39: Discretizing Instruments First Stage Results

	(1) Enslaved Ancestry	(2) Enslaved Ancestry	(3) Enslaved Ancestry	(4) Enslaved Ancestry
High Wheat Suitability	0.003		0.002	
	(0.002)		(0.002)	
Low Cocoa Suitability	-0.003		-0.002	
	(0.002)		(0.002)	
High Wheat $\times$ High Cocoa	0.011***	0.011***	0.011***	0.011***
	(0.004)	(0.003)	(0.004)	(0.003)
Observations	1099	1099	1098	1098
Underidentification K-P Stat	19.830	16.324	15.876	13.404
Underidentification P-Val	0.000	0.000	0.001	0.000
Weak Idenficiation Stat	6.419	15.923	4.960	12.663
Overidentification P-Val	0.220	.	0.055	.
Population	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes
Contemporary Distances	Yes	Yes	Yes	Yes
Luminosity	No	No	Yes	Yes
Illiteracy Rate	No	No	Yes	Yes
Additional Controls	No	No	Yes	Yes

Notes: First stage regressions of enslaved ancestry on different instruments. These instruments include an indicator for an above-median split of wheat suitability, an indicator for a below-median split of cocoa suitability, and the interaction of both. The rest of the controls are in line with the analyses from Tables 2.B.2 and 2.3. Robust standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

Table 2.A.40: Discretizing Instruments Second Stage Results

Panel A	(1) Luminosity (1998)	(2) Illiteracy (1990)	(3) Child Mortality (2000)	(4) Chávez Share (1998)
Enslaved Ancestry	-30.381***	2.849***	0.337**	2.029*
	(9.634)	(0.927)	(0.142)	(1.094)
Observations	1098	1098	1098	1087
Population	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes
Specification	2SLS	2SLS	2SLS	2SLS

Panel B	(1) Luminosity (1998)	(2) Illiteracy (1990)	(3) Child Mortality (2000)	(4) Chávez Share (1998)
Enslaved Ancestry	-37.491***	3.087***	0.343**	1.611
	(11.397)	(1.048)	(0.155)	(1.114)
Observations	1098	1098	1098	1087
Population	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes
Specification	2SLS	2SLS	2SLS	2SLS

Notes: Robustness of main results to alternative instruments. Panel A uses three instruments: an indicator for an above-median split of wheat suitability, an indicator for a below-median split of cocoa suitability, and the interaction of both. Panel B only uses the interaction. Results are largely in line with those in the main body of the paper, but suffer from weak instruments as documented in Table 2.A.39. Controls correspond to those used in the analyses of Tables 2.1 and 2.3. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.41: OLS Results and Heterogeneities: Development Outcomes

	(1) Luminosity (1998)	(2) Luminosity (1998)	(3) Illiteracy (1990)	(4) Illiteracy (1990)	(5) Child Mortality (2000)	(6) Child Mortality (2000)
Enslaved Ancestry	-2.634**	2.759	0.181	-0.069	0.004	0.010
	(1.288)	(2.021)	(0.113)	(0.187)	(0.022)	(0.063)
Close to *Haciendas*		0.271		-0.007		0.001
		(0.189)		(0.018)		(0.004)
Enslaved Ancestry × Close to *Haciendas*		-8.097***		0.373		-0.009
		(2.610)		(0.247)		(0.064)
Observations	1098	1098	1098	1098	1098	1098
Population	Yes	Yes	Yes	Yes	Yes	Yes
Indigenous PP Dummy	Yes	Yes	Yes	Yes	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes	Yes	Yes
Land Cover	Yes	Yes	Yes	Yes	Yes	Yes
Specification	OLS	OLS	OLS	OLS	OLS	OLS

Notes: Correlation of enslaved ancestry on main development outcomes, complemented with analysis of heterogeneity by proximity to cocoa *haciendas*. Close to *haciendas* indicator defined as a median split on the distance to the closest *hacienda*. Controls correspond to those used in the analyses of Table 2.1. Robust standard errors in parentheses. * p-value $<$ 0.1, ** p-value $<$ 0.05, *** p-value $<$ 0.01.

Table 2.A.42: OLS Results and Heterogeneities: Chávez's Share (1998)

	(1) Chávez Share (1998)	(2) Chávez Share (1998)	(3) Chávez Share (1998)	(4) Chávez Share (1998)
Enslaved Ancestry	0.110	0.456**	0.456**	0.368*
	(0.151)	(0.231)	(0.226)	(0.207)
Close to *Haciendas*		0.053**	0.053**	0.049**
		(0.021)	(0.022)	(0.021)
Enslaved Ancestry × Close to *Haciendas*		-0.595*	-0.601*	-0.396
		(0.316)	(0.316)	(0.292)
Observations	1087	1092	1087	1087
Luminosity	Yes	No	No	Yes
Illiteracy Rate	Yes	No	No	Yes
Population	Yes	No	Yes	Yes
Indigenous PP Dummy	Yes	No	Yes	Yes
State Indicators	Yes	Yes	Yes	Yes
Latitude & Longitude	Yes	Yes	Yes	Yes
XVIIIc. Variables	Yes	Yes	Yes	Yes
Land Cover	Yes	No	Yes	Yes
Additional Controls	Yes	No	Yes	Yes
Specification	OLS	OLS	OLS	OLS

Notes: Correlation of enslaved ancestry on main electoral outcomes, complemented with analysis of heterogeneity by proximity to cocoa *haciendas*. Close to *haciendas* indicator defined as a median split on the distance to the closest *hacienda*. Controls correspond to those used in the analyses of Table 2.3 in increasingly strict specifications. Robust standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01.

2.B Additional Details on Data Sources

2.B.1 Historical Maps of Venezuela

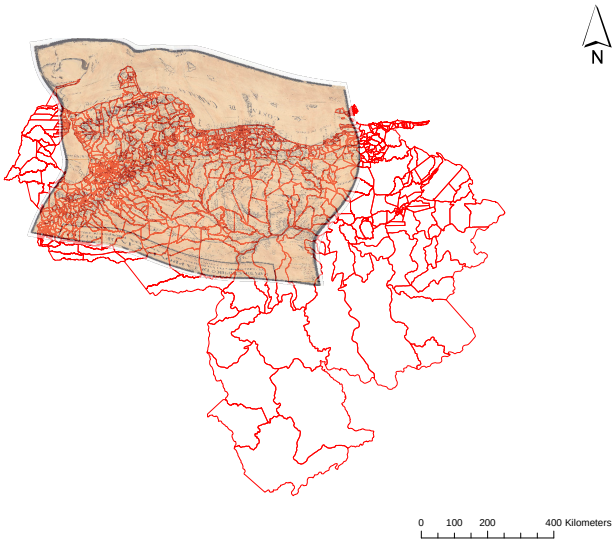
The map form which I base the location of cocoa *haciendas* and the layout of the country in the XVIIIth century comes from de Santos Cabrera (n.d.) and is reproduced below in Figure 2.B.1. I digitised this map manually in ArcGIS PRO. The map includes the location of cities, villas, monasteries, missions, military strongholds, farms, roads, lakes, and cocoa *haciendas*.



Notes: Map of colonial Venezuela from de Santos Cabrera (n.d.). This map was digitized to find the location of cocoa *haciendas* during the XVIIIth century. All maps were processed in ArcGIS Pro.

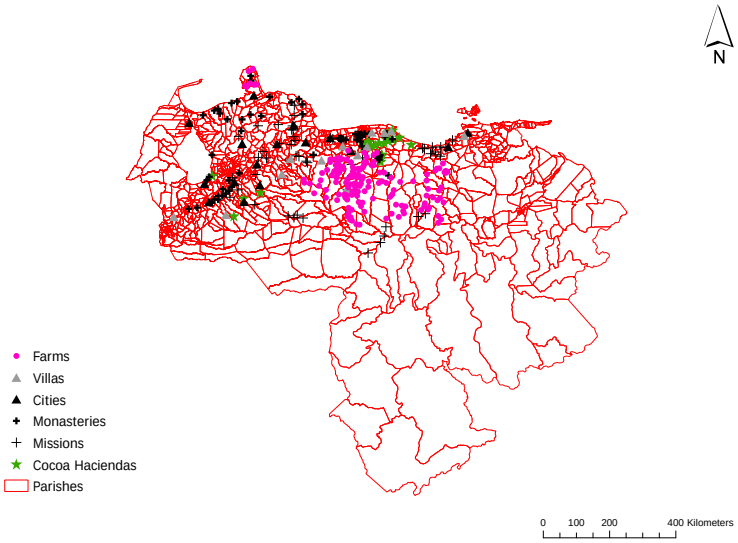
Figure 2.B.1: XVIIIth Century Map of Colonial Venezuela by de Santos Cabrera (n.d.)

The map from which I retrieve the proxy for the location of areas used for large-scale agricultural production during the colonial period is from Codazzi (1840). The map comes from a collection of lithographies that comprise the first atlas published for a South American country. The bottom panel of the lithography (Figure 2.B.4) shows Venezuela in the XIXth century divided into three areas: grasslands, jungles, and areas devoted to agriculture. Figure 2.B.5 shows the map superimposed on a contemporary map of Venezuela.



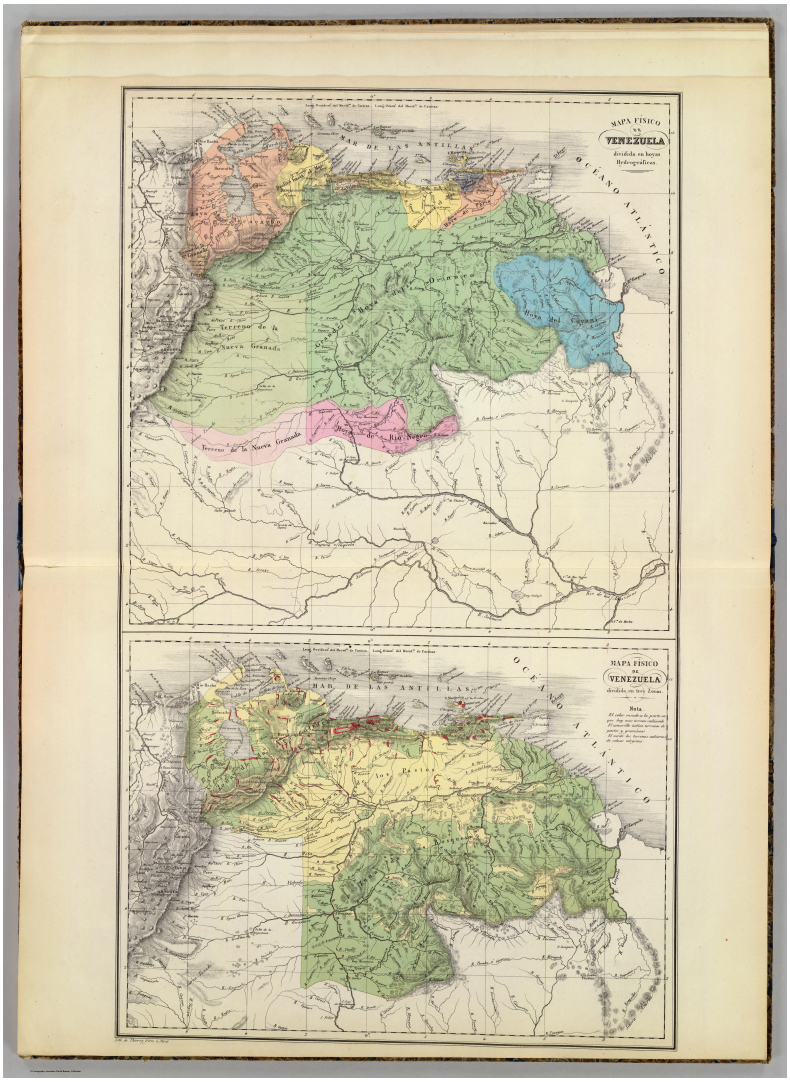
Notes: Map of colonial Venezuela from de Santos Cabrera (n.d.) after digitization superimposed on a contemporary map of Venezuela. Venezuela's shapefile is from Porto Tapiquén (2020). All maps were processed in ArcGIS Pro.

Figure 2.B.2: XVIIIth Century Map of Colonial Venezuela by de Santos Cabrera (n.d.) Superimposed on Current Map of Venezuela



Notes: Data from colonial map of Venezuela from de Santos Cabrera (n.d.) after digitization superimposed on a contemporary map of Venezuela. Venezuela's shapefile is from Porto Tapiquén (2020). All maps were processed in ArcGIS Pro.

Figure 2.B.3: Data Extracted from XVIIIth Century Map of Colonial Venezuela by de Santos Cabrera (n.d.)



Notes: Lithography from Codazzi (1840). This lithography was part of the first atlas of any South-American country. I digitized the bottom panel to proxy for the location of areas that were devoted to large-scale agriculture during the XVIIIth century. All maps were processed in ArcGIS Pro.

Figure 2.B.4: Map of Venezuela's Hydrography and Agricultural Zones from Codazzi (1840)

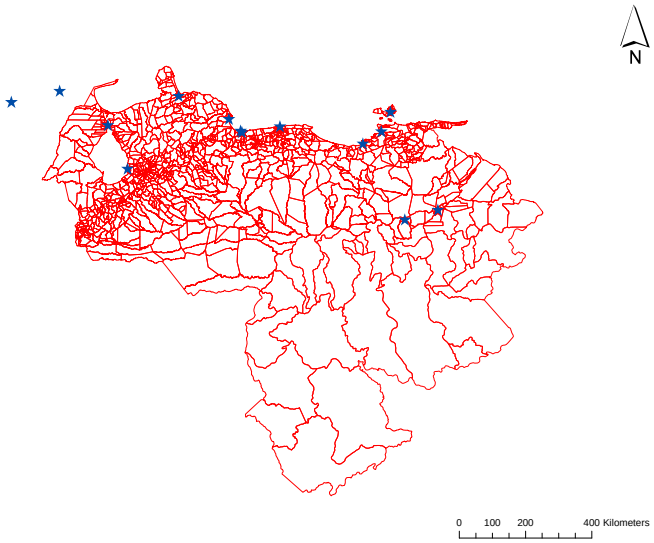


Notes: Bottom panel of the lithography from Codazzi (1840) superimposed on a contemporary map of Venezuela after digitization. I proxy for the location of areas that were devoted to large-scale agriculture during the XVIIIth century with the areas shown in red. All maps were processed in ArcGIS Pro.

Figure 2.B.5: Map of Venezuela's Hydrography and Agricultural Zones from Codazzi (1840) Superimposed on Current Map of Venezuela

2.B.2 Trans-Atlantic and Intra-American Slave Trade Databases

The Trans-Atlantic and Intra-American slave trade databases from Eltis (1999) have been widely used in past research (e.g., Nunn, 2008). Here, I join both databases and record ports of arrival in Venezuela or Colombia. I retrieve the coordinates for each port from Google Maps or the port's Wikipedia page. Figure 2.B.6 shows the location of all such ports.



Notes: Location of ports in Venezuela and Colombia used during the Slave Trade according to the Trans-Atlantic and Intra-American Slave Trade Databases (Eltis, 1999). Port locations were retrieved from each port's Wikipedia page or from Google Maps. Venezuela's shapefile is from Porto Tapiquén (2020). All maps were processed in ArcGIS Pro.

Figure 2.B.6: Ports Used for Trading Enslaved Individuals per the Trans-Atlantic and Intra-American Slave Trade Databases

2.B.3 African Slave Surnames in Acosta Saignés (1984)

In his excellent history of slavery in Venezuela, Acosta Saignes (1984) provides a list of the most common surnames that slaves brought into Venezuela received. The list of surnames is as follows:

ANAGACHI	CAFO	GANGARÁ	MALEMBA	NAGO
ANGOLA	CAMBINDO	GEGÉ	MALINKE	NDONGO
AQUEL	CAMBUTA	GOLOFO	MANDÉ	NGANGÁ
ARACHE	CAMOANDA	GUACHÍ	MANDELE	NGOLA
ARARÁ	CAMOCONDA	GUASA	MANDIMENCA	POPÓ
AROBÍ	CAMOJUNDA	GUAZA	MANDINGA	QUIBUTO
ASHANTI	CANDALA	GUINEA	MASINGA	QUINENE
BABUENDE	CANENE	GUUNGA	MASMOL	QUISONGO
BAGA	CARABALÍ	HAUSSA	MATAMBA	QUISULO
BALA	CASIMBA	IMBANGALA	MAYOMBE	SANTÉS
BALANTE	CATAGALA	ITALE	MBALA	SAPE
BAMBÍ	CATALOAQUI	JALUNGA	MENDÉ	SOSO
BANANA	CATENDE	JERERO	MINA	SUCHI
BANGUELA	CHALATA	JIRÍ	MINAS	SUNDI
BAÑÓN	CHARA	KANDA	MOBANGOMBE	TACANGA
BARIBA	CONGO	KANEN	MOJINGA	TAGUNCHO
BARILA	CULE	KANO	MONDA	TARÍ
BARROL	CUMBÁ	KANURI	MONDONGO	TARY
BASUNDI	CUNENE	KASSA	MOSANGA	TEMENE
BEMBA	DYOLA	KONGO	MOTOANGOMBE	TOGO
BINGA	EBLÁ	KPWES	MOTOMBO	ULOF
BIRONGO	EMBUILA	LIMBA	MPANGU	VRAFARA
BOLOE	ENTÓTERA	LOANGA	MUMBALA	WOLOF
BRAFARA	EWE-FON	LOANGO	MUMBONA	XOXO
BRAMA	FELUPE	LONGO	MUNDAMBA	YORUBA
BRAN	FOLOPO	LUANGO	MUNGUCHI	ZAPA
BRONG	FULA	LUCUMÍ	MUNYACA	ZAPE
CABENDA	FULO	LUENGA	MUSOMBO	
CABINDA	GANGA	LUF	MUSOSO	
CACHANGA	GANGÁ	MABALA	MUSUNDI	

2.B.4 The Tascón-Maisanta Database

Database History and Contents

The Tascón-Maisanta list was shared among Venezuelans in CDs which contained the *Batallón de Santa Inés* software. The CDs contained several .dbf files comprising the data, a .exe file that manipulates these different databases – the software itself –, a text file, and a presentation (see Figure 2.B.7 panel (a)). The text file (see Figure 2.B.7 panel (b)) and presentation contain general instructions about using the software, and a description of the regime's justification for distributing this data. Using the software is simple. A pop-up screen prompts the user the enter a specific voter ID (*Número de Cédula*) to perform a search (see Figure 2.B.7 panel (c)). One can also perform a search by typing somebody's full name. The software returns whether this individual signed against Chávez or not (Figure 2.B.7 panels (d) and (e), respectively), this individual's address, the details of the centre where this person is registered to vote, whether the person is alive, and whether they were recipients of two government programmes (*Misión Ribas* and *Vuelvan Caras*). As detailed on top of the screen, the databse contains the information of 12,394,109 registered voters.

Rather than performing a search, one can also access the data by clicking on the icon with the magnifier glass. This prompts the user to choose one of the 24 states in Venezuela and one of their municipalities (Figure 2.B.7 panel (f)). After choosing a municipality, a user can see summary statistics of the percentage of registered voters that signed against Chávez at the level of the voting centre (Figure 2.B.7 panel (g)). Further, one can choose a specific voting centre, and the software will display the list of voters: their IDs, their name, date of birth, age, address, and the rest of the aforementioned information (Figure 2.B.7 panel (h)). The user has the option of generating a .txt file with this information for printing (Figure 2.B.7 panel (i)).

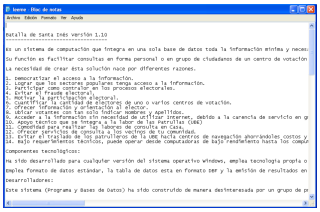
Ethical Use of the Tascón-Maisanta Data

Despite considering the Tascón-Maisanta list a historical data source, it is undeniable that it contains personal information that could potentially affect people who are still alive today. This commands a series of ethical considerations from researchers who want to use it.

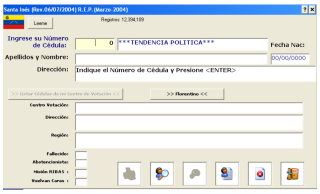
To maintain the ethical integrity of this project, all data manipulations were carried out offline and within a virtual machine installed locally with the sole purpose of manipulating the data. I extracted the .dbf files, saved them in a protected hard drive, joined the data in Stata, eliminated all personal information with the exception of the names of the voters



(a) CD Contents



(b) Text File



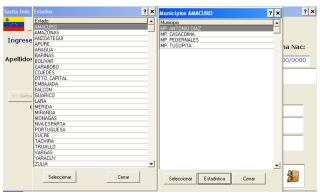
(c) User Interface



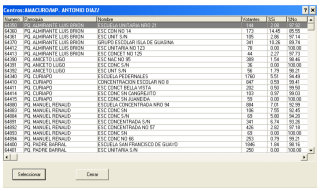
(d) Example of Results 1



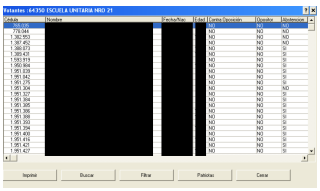
(e) Example of Results 2



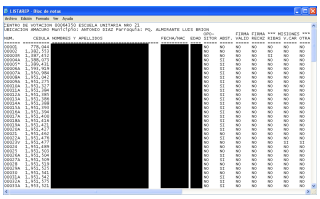
(f) Accessing Data Detail 1



(g) Accessing Data Detail 2



(h) Accessing Data Detail 3



(i) Printing Data

Notes: Panel (a) shows a screenshot of the contents of the CD containing the Batallón de Santa Inés software. Panel (b) shows a screenshot of the instructions for using the software. Panel (c) shows a screenshot of the software. Panels (d) and (e) show examples of results after searching for individuals on the software. Panels (f), (g), and (h) show screenshots of how users can access the Tascón-Maisanta list through the software. Panel (f) shows how users can print the data from within the software. To maintain anonymity of people in the list, the figures are redacted where they contain personal information.

Figure 2.B.7: Tascón-Maisanta List and Batallón de Santa Inés Software

and the numerical identifiers of the parishes where they live. After being unable to identify individuals with additional details such as their address or date of birth, I integrated the data to the rest of the pipeline.

Chapter 3

The Economic Consequences of Peace: The Economic Effects of Successfully Combating Drug Cartels in Mexico

Abstract

In 2013 the government of the state of Coahuila, Mexico, implemented a unique package of policies with the objective of combatting the presence of drug cartels. Two crucial dimensions included strengthening local police enforcement, and banning businesses that provided liquidity to cartels such as brothels, junk-yards, and casinos. Using synthetic controls, I test the effectiveness of this set of policies against 'intentional homicides' – the metric targetted by Coahuila's government – as well as all other crime-types. I find a modest reduction of crimes related with the presence of cartels, and no impact on crimes unrelated to cartels. I extend the analysis to assess the economic effects of these policies. I find no statistical evidence of economic dividends of peace beyond mechanical effects parsimoniously explained by the policies adopted, such as a reduction in informality.

Acknowledgements: I want to thank my supervisors Sanjay Jain and Paola Conconi for encouraging me to write this article. Thank you to Carlos González Pérez for discussing the synthetic controls methodology and brainstorming about sound robustness checks. Thank you to Julia Fischer for helpful comments and suggestions. Thank you to my parents, Abelardo and María del Carmen, who helped me gather contextual information when I could not travel to Mexico due to travel restrictions related to the Covid pandemic. All mistakes and omissions are my own responsibility.

3.1 Introduction

This paper explores two broad questions: what is an appropriate strategy to fight organized crime, thereby reducing violence and conflict? And, if successful, what are the economic consequences of achieving peace? I explore these questions in the context of Mexico. Specifically, I evaluate the effectiveness of a package of policies implemented in the state of Coahuila between 2012 and 2017 with the objective of reducing the presence of drug cartels. I test whether these policies were successful in decreasing 'intentional homicides' – Coahuila's government's explicit target – as well as other crimes; and I extend the analysis to test for effects on economic outcomes.

To put things into perspective, Mexico was the country with the highest number of journalists killed in 2022 (Reporters Without Borders, 2022) – surpassing countries at war such as Ukraine, Yemen, and Syria. Most of these deaths are connected with the presence of cartels. In this context, the state of Coahuila implemented an ambitious set of policies comprising over 80 new state laws. These policies targeted two crucial dimensions: empowering local police enforcement, as opposed to relying on federal police forces; and banning economic activities that provided the liquidity used by cartels to fund their operations. The latter entailed the closing down of casinos, brothels, and junk-yards, for instance.

Coahuila's government reported that the policies were highly successful, achieving an impressive reduction of 87% in the number of 'intentional homicides'[1] between 2012 and 2017. While the results in levels are indeed encouraging, similar trends were observed in states that did not adopt such a strict set of policies. This motivates evaluating the policies against a valid counterfactual. The evidence found in this paper suggests the policies were indeed successful at reducing the number of 'intentional homicides', as well as some but not all crime-types correlated with the presence of cartels. Thus, an evaluation against a valid counterfactual does paint a positive – albeit more modest – picture. Further, I find no suggestive evidence of a deep or long-lasting economic transformation in Coahuila beyond mechanical effects parsimoniously explained by the specific set of policies implemented by the government.

One of the key empirical contributions of this paper is evaluating these policies against a valid counterfactual. In order to construct a counterfactual, I rely on the synthetic controls methodology (Abadie, Diamond, & Hainmueller, 2010, 2015; Abadie & Gardeazabal, 2003).[2]

[1] I use literal translations from Spanish to refer to all crime-types. See Section 3.4 for details.

[2] Studying the origins and consequences of conflict and peace is empirically challenging. Establishing any causal relations (or the lack thereof) between conflict and economic development, ethnic segregation, polarisation or inequality, runs up against problems of reverse causality and omitted variables (see the review in Ray and Esteban (2017) for a discussion). The recent methodological development of synthetic controls

For inference and in order to push the causality of the findings, I apply the recent advances in Chernozhukov et al. (2021), calculating confidence intervals for effect estimates in all post-treatment periods when the synthetic control is calculated using cross-validated Lasso.[3] These results are complemented by applying a canonical synthetic differences-in-differences analysis (Arkhangelsky et al., 2021), which finds similar results. The paper is also a practical contribution for researchers as it provides code for applying the methodology of Chernozhukov et al. (2021) in Stata.

The influential work of Collier and Hoeffler (1998, 2004) sparked a great interest in understanding the drivers of conflict in society.[4] The opposite direction of causality exploring the economic consequences of conflict has also been well-documented.[5] In a recent book, Rohner (2024) emphasises several self-reinforcing pathways in which conflict causes longstanding economic damage. These cycles include low growth (Mueller & Tobias, 2016), low inter-personal trust and social capital (Fouka & Voth, 2023; Guiso, Sapienza, & Zingales, 2009; Rohner, Thoenig, & Zilibotti, 2013), or low human capital (León, 2012; Shemyakina, 2011). There is also strong evidence suggesting that improvements in education (Agarwal, Bahure, Bergonzoli, & Maji, 2024; Rohner & Saia, 2019), health (Berlanda, Cervellati, Esposito, Rohner, & Sunde, 2024; Cervellati, Sunde, & Valmori, 2017), and labour market outcomes (Blattman & Annan, 2016; Couttenier, Petrencu, Rohner, & Thoenig, 2019; Fetzer, 2020) can lead to reductions in conflict, potentially providing pathways to break from these cycles.

While most research has focused on aligning incentives in order to prevent conflict from happening, fewer papers have explored the reverse direction of causality and the complement to conflict i.e., examining the economic effects of achieving peace.[6] The work of Besley and Mueller (2012) is an exception. Here, the authors estimate the peace dividend of the process of pacification in Northern Ireland, proxied by an increase in house prices. In a similar vein, a recent paper by Fajardo-Steinhäuser (2024) explores the economic dividends of the pacification process involving the FARC guerilla in Colombia. Fajardo-Steinhäuser (2024) finds no evidence of economic peace dividends despite achieving an impressive reduction in violence. He rationalizes this finding by providing evidence that suggests the existence of a low state capacity equilibrium *à la* Besley and Persson (2010).

(Abadie et al., 2010, 2015; Abadie & Gardeazabal, 2003) allows researchers to study these questions in settings where data constraints impede inference via other methodologies.

[3] By allowing for more flexibility in the construction of the synthetic control, this approach improves the models' fit by a significant margin compared to models constructed under the canonical methodology from Abadie and Gardeazabal (2003).

[4] See e.g., Besley and Persson (2008, 2010, 2011); Dal Bo and Dal Bo (2011); Dube and Vargas (2013); Esteban and Ray (2008, 2011a, 2011b); Fearon and Laitin (2003); Miguel, Satyanath, and Sergenti (2004); Mitra and Ray (2014); Voigtländer and Voth (2012). Additionally, see Ray and Esteban (2017) for a review.

[5] See Blattman and Miguel (2010) and Rohner and Thoenig (2021) for reviews.

[6] This is a point clearly stated by Fajardo-Steinhäuser (2024).

The second empirical contribution of this paper is continuing to explore this direction of causality with regards to peacemaking (see Besley & Mueller, 2012; Fajardo-Steinhäuser, 2024), as the main objective of Coahuila's government was reducing the presence of cartels, rather than targetting economic indicators per se. However, while the paper's core set of results suggest moderate success, a series of caveats need to be acknowledged which should suggest caution when interpreting the findings of the paper. These caveats include the inability to corroborate the faithful implementation of the policies chosen by Coahuila's government due to a lack of data. The data on crimes reported by the Mexican government is also taken at face-value and might represent a lower-bound on criminal activity in the country. Further, it is not possible to validate it against other commonly-used sources on conflict such as the Uppsala Conflict Data Program (Davies, Pettersson, Sollenberg, & Öberg, 2025) or the Armed Conflict Locations and Events Data (ACLED) (Raleigh, Kishi, & Linke, 2023) as these sources report many fewer incidents.

The paper also leaves several open questions that could be addressed in future research. Peace efforts that lead to improvements in economic indicators could generate virtuous cycles that prevent future conflict – for instance, through human capital accumulation (e.g., Agarwal et al., 2024) or employment opportunities (e.g., Couttenier et al., 2019), as discussed above. Kickstarting these processes requires strong political will and coordination, which remain unexplored.[7] The literature focusing on post-conflict contexts is also clear on broader policies that should be encouraged in order to prevent future conflict (Mueller & Tobias, 2016). These include fostering investor confidence (e.g., Mueller, Piemontese, & Tapsoba, 2017), as well as building trust and inclusive political institutions (e.g., Appel & Loyle, 2012; Blattman, Hartman, & Blair, 2014). Relatedly, as highlighted in Rohner and Zhuravskaya (2023) another fruitful and policy-relevant area for future research is nation-building in post-conflict settings (e.g., Blouin & Mukand, 2019).

The rest of the paper is organised as follows. Section 3.2 provides some contextual information on organised crime in Coahuila and a description of the set of policies implemented by their government in late 2012. Section 3.3 provides an overview of the synthetic controls framework, starting with its conceptual definition and original formulation (Abadie et al., 2010; Abadie & Gardeazabal, 2003). This discussion also considers recent theoretical advances that can aid as tools for inference in empirical work.[8] Section 3.4 briefly discusses data sources and implementation. Section 3.5.1 analyses the impact of the set of policies on 'intentional homicides'. Section 3.5.2 extends the analysis to other crime-types. Section

[7] This resonates with Rohner and Thoenig (2021), who emphasise strong complementarities between peacebuilding and development that can only be realized if policies explicitly target them jointly.

[8] This overview relies heavily on Chernozhukov et al. (2021) and the recent synthesis in Abadie (2021). Readers familiar with these materials are encouraged to skim over the section.

3.5.3 considers the impact on economic indicators. Section 3.6 provides robustness checks. Finally, section 3.7 discusses limitations and concludes. Additional figures and tables are reported in Appendices 3.A and 3.B, respectively; while appendix 3.C reproduces the code used to compute confidence intervals throughout the paper.

3.2 Context

Drug cartels have had a long-time presence in Mexico. By the mid-2000s one of the most influential cartels in the country was *Los Zetas*, who operated in the states near the US border. The government of president Felipe Calderón (2006 - 2012) decided to combat organised crime in Mexico directly by declaring an open war on drug production and smuggling. Slowly, violence escalated in the entire country, turning Calderón's presidency into one of the most violent in Mexico's history. Panel (a) in Figure 3.1 plots the 12-month moving average monthly 'intentional homicides' in all states in Mexico. It is evident that this crime rose across the country in the lead-up to the year 2012.

Calderón's presidency mostly coincided with the government of Humberto Moreira Valdés in the state of Coahuila. In order to curb the presence of drug cartels – most notably, *Los Zetas* who controlled 90% of the state's territory, and *El Cártel del Pacífico* – these governments took a top-down approach that relied heavily on the presence of the Mexican army carrying out policing duties within the state. The set of policies adopted by these governments are related to an overall increase in violence as both Panel (a) and (b) from Figure 3.1 suggest. In January 2011, Humberto Moreira stepped down from his position to run for the presidency of the Executive Committee of his party, the *Partido Revolucionario Institucional* (PRI). Jorge Torres López became the interim governor of the state, until December 2011 when the government of Rubén Moreira Valdés, brother of Humberto Moreira, took office.

Between 2006 and 2012, there was an increase not just in 'intentional homicides', but also an alarming increase in other types of crime, including kidnapping, forced disappearance, extortion, violent theft from businesses, and violent car theft. Most of these crimes are strongly related to the presence of drug cartels. In particular, *Los Zetas* found in Coahuila a relatively undisputed territory that guaranteed them at least two entry-points into the United States through the cities of Acuña and Piedras Negras.[9] The *modus operandi* of this specific cartel is characterised by four distinct sets of actions: expanding their busi-

[9] *Los Zetas* slowly migrated into Coahuila from the neighbouring state to the east, Tamaulipas – mostly controlled by the *Cártel del Golfo* –, and which has at least 20 entry-points into the United States. The *Cártel del Pacífico*, which controlled the state of Durango and the city of Torreón in Coahuila, prevented them from taking more territory towards the west.

ness into the legal economy, becoming part of local communities through imposing and arranging marriages, generating fear, and violently eliminating competition from other cartels (Aguilar Valenzuela & Moreira Valdez, 2022). Some of the legal businesses the cartel started to expand towards included construction, real estate, restaurants, dance halls, hotels, casinos, bars, brothels, currency exchanges near the border, and second-hand car dealerships. The nature of these businesses allows cartels to launder money generated through other illegal activities such as drug-smuggling and human-trafficking.

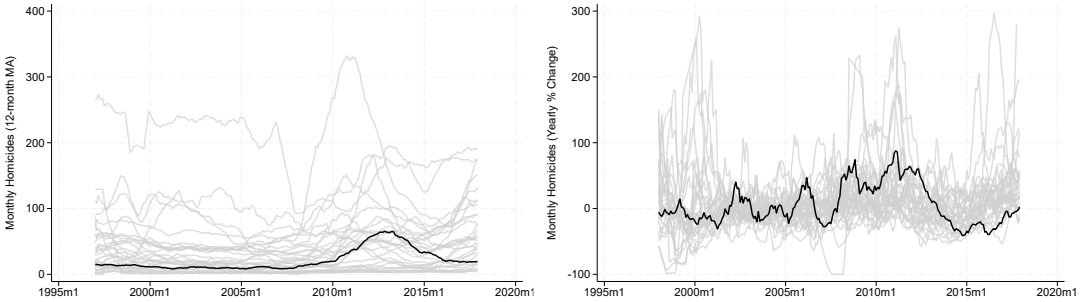
In a recent monograph, Aguilar Valenzuela and Moreira Valdez (2022) detail how drug cartels can slowly incorporate into their ranks criminals with the capacity to exercise brutal violence, whose actions are not even related to drug production and smuggling. These individuals are known as *halcones*. Sometimes, the revenues accrued from these other activities surpass those from drug production and smuggling. Aguilar Valenzuela and Moreira Valdez (2022) also explain that cartels do not only focus on the production of marijuana, but also other drugs such as crack, methamphetamine, fentanyl, or alcohol. This gradual permeation of cartels into society is what makes combating their presence such a challenge for policymakers.

Table 3.1: Coahuila's Multidimensional Strategy to Combat Organised Crime

Prevention	1. Generate employment opportunities 2. Increase number of schools 3. Promote culture and sports 4. Promote civic values 5. Legal reforms 6. Respecting human rights 7. Look for missing people
Combating Organised Crime	8. Combating violence generated by cartels 9. Outlaw casinos, cockfighting, brothels, etc. 10. Improve police capacity 11. Imprisoning major criminals 12. Improve coordination with federal authorities 13. Improve coordination
Support Actions	14. Improve available statistics 15. Improve government communication

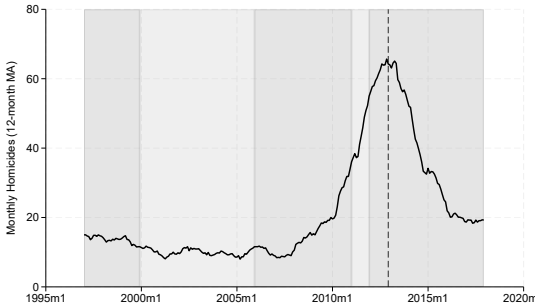
Notes: This table is adapted from Aguilar Valenzuela and Moreira Valdez (2022).

The government of Rubén Moreira Valdés implemented a multidimensional set of policies to combat cartels in Coahuila. The different margins that comprise these policies are summarised in Table 3.1, and have been considered as widely successful. Aguilar Valenzuela and Moreira Valdez (2022) discuss the context in which these policies were implemented and evaluate their effectiveness. This complex set of policies, which resulted in



(a) Monthly Homicides in Mexico by State

(b) Yearly Percentage Change of Monthly Homicides in Mexico by State



(c) 'Intentional Homicides' in Coahuila (12-month MA)

Notes: Panel (c) plots the 12-month moving average of 'intentional homicides' in the state of Coahuila form January 1997 to December 2017. The colours in the background denote the different governments the state had in that time period. From December 1993 to 1999 the state was governed by Rogelio Montemayor Seguy; from December 1999 to 2005, by Enrique Martínez y Martínez; from December 2005 to January 2011 by Humberto Moreira Valdés; from January 2011 to December 2011 by the interim governor Jorge Torres López; and from December 2011 to 2017, by Rubén Moreira Valdés, the brother of Humberto Moreira Valdés. All of these governments were from the same party: the *Partido Revolucionario Institutcional* (PRI).

Figure 3.1: Homicides in Mexico and the State of Coahuila

the introduction of over 80 new state laws, had two key margins: strengthening local police enforcement, and banning activities that contributed to financially sustain cartels. The latter policy entailed the banning of casinos, cockfighting, horse-racing, brothels, junk-yards, and prosecuting the illegal sale of alcohol and petrol, as well as strengthening the enforcement of laws such as those requiring the display of license plates and the official registry of all vehicles. Special care was also put in blocking cellphone signals in all state prisons. In an independent evaluation of these policies, Aguayo and Dayán (2020) also highlight policies directed at cultural change, such as destroying altars depicting *la Santa Muerte*,[10] banning *narco-corridos* from radio broadcasting, encouraging citizens to report criminal activities, as well as increasing employment opportunities, and promoting

[10] La Santa Muerte is a sacred personification of death depicted as a female skeleton. It is estimated that 5% of Mexico's population is devoted to its cult as a patron-saint. Its cult has been associated with organised crime, including drug trafficking and kidnapping. See the work of Chesnut (2012) and Kristensen (2015) for overviews on *la Santa Muerte*.

education and sport.

During the government of Rubén Moreira Valdés, Coahuila experienced a remarkable decline in 'intentional homicides' and other activities associated with cartels. Aguilar Valenzuela and Moreira Valdez (2022) document a 87% decrease in the levels of these crimes, as well as a 53% decrease in thefts, 42% decrease in kidnapping, and 70% decrease in extortions between 2012 and 2017. This change is captured in Figure 3.1's Panels (b) and (c). However, Panel (b) also highlights that Coahuila was not the only state that experienced a decline in 'intentional homicides' between 2013 and 2015. This motivates the need to evaluate the effectiveness of the policy relative to a valid counterfactual. The synthetic controls methodology is the appropriate tool to evaluate the success of these policies. The next section provides a methodological overview of synthetic controls, and details one of the approaches adopted in the paper to conduct inference.

3.3 An Overview of the Synthetic Controls Methodology

Following Abadie and Gardeazabal (2003), Abadie et al. (2010, 2015) and Abadie (2021), we are in a setting where we have data for $J+1$ units. Without loss of generality, unit $j=1$ receives a treatment. Units $j=2,...,J+1$ are what is called the *donor pool*. We observe data over T periods. The period when the intervention was implemented is denoted by T_0. Therefore, T is divided into a pre-intervention period $1 \leq t \leq T_0$, and a post-intervention period $T_0 < t \leq T_1$, where T_1 denotes the last period for which data is observed. This implies $T = T_0 + T_1$. Further, let Y_{jt} denote the outcome of interest for unit j at period t. Additionally, for each unit j, we observe a $k \times 1$ vector of predictors $\mathbf{X}_j$ which may include the pre-intervention values of Y_{jt}.

As in Abadie (2021), Y_{jt}^I is used to denote the potential outcome of unit j at time t under the intervention, and Y_{jt}^N to denote the potential outcome without the intervention. This is the standard potential outcomes framework under Rubin's model of causal inference. We are interested in finding the effect of a policy for unit 1 after it has been implemented. That is, we are interested in estimating:

$$\tau_{1t} = Y_{1t}^I - Y_{1t}^N,$$

for $t > T_0$. The fundamental problem of causal inference comes from the fact that we do not observe Y_{1t}^N for any $T_0 < t \leq T_1$, i.e., the counterfactual for the treated unit.

The synthetic control methodology tackles this challenge by calculating a weighted average of the units from the donor pool so that it resembles the treated unit as closely as

possible in the pre-intervention period[11] using a $J \times 1$ vector of weights, $\mathbf{W} = (w_2, .., w_{J+1})'$. The synthetic control is constructed by calculating the weighted average for all $t = 1, ..., T_1$. The synthetic control is the counterfactual of unit 1 for all $T_0 < t \leq T_1$. The difference between the observed outcome and the synthetic control in the post-intervention period corresponds to the effect of the implemented policy. Formally, the estimator is defined as:

$$\hat{Y}_{1t}^N = \mu + \sum_{j=2}^{J+1} w_j Y_{jt},$$
$$\hat{\tau}_{1t} = Y_{1t}^I - \hat{Y}_{1t}^N.$$

Recent methodological developments centre on finding different ways to estimate the constant μ and the weights w_j.[12] Essentially, this is a loss minimisation problem subject to a series of constraints that govern the values of μ and w_j.[13] For instance, by setting the weights so that they are positive, constant numbers that sum to one, and with no restrictions on the values of μ, we would end up with the standard difference-in-differences framework (J. D. Angrist & Krueger, 1999; J. D. Angrist & Pischke, 2009; Athey & Imbens, 2006). This approach sets:

$$w_j = \frac{1}{J} \qquad \forall j = 2, ...J+1, \quad \text{and}$$
$$\hat{\mu} = \frac{1}{T_0} \sum_{t=1}^{T_0} Y_{1t} - \frac{1}{J \cdot T_0} \sum_{t=1}^{T_0} \sum_{j=2}^{J+1} Y_{jt}.$$

These values yield:

$$\begin{aligned} \hat{Y}_{1t}^N &= \hat{\mu} + \frac{1}{J} \sum_{j=2}^{J+1} Y_{jt} \\ &= \left(\frac{1}{T_0} \sum_{t=1}^{T_0} Y_{1t} - \frac{1}{J \cdot T_0} \sum_{t=1}^{T_0} \sum_{j=2}^{J+1} Y_{jt} \right) + \frac{1}{J} \sum_{j=2}^{J+1} Y_{jt}, \end{aligned}$$

which can then be used to find $\hat{\tau}_{1t}$.

Another example is what Imbens and Doudchenko (2016) refer to as the 'canonical synthetic control method' – originally described by Abadie and Gardeazabal (2003) and

[11] Formally, the similarity could be defined over the entire pre-intervention period for which data is available $1 \leq t \leq T_0$, or for a subset of those periods.

[12] For detailed reviews of these methodologies see Imbens and Doudchenko (2016), Chernozhukov et al. (2021) and Abadie (2021).

[13] For a systematic analysis of what these constraints are, see Imbens and Doudchenko (2016).

Abadie et al. (2010, 2015). This setup assumes $\mu = 0$, and solves the following problem:

$$\hat{w} = \arg\min_{w} \sum_{t=1}^{T_0} \left(Y_{1t}^N - \sum_{j=2}^{J+1} w_j Y_{jt}^N \right)^2$$
$$\text{s.t.} \qquad w \geq 0, \quad \text{and}$$
$$\sum_{j=2}^{J+1} w_j = 1.$$

Note that this is a simplified version of the problem studied by Abadie et al. (2010, 2015), where the calculation of the optimal weights is carried out using additional covariates i.e., using the previously defined matrix $\mathbf{X}_j$. Imbens and Doudchenko (2016) refer to this special case of the problem as the 'constrained regression'. The problem is constrained because the vector of predictors $\mathbf{X}_j$ is set to only include pre-treatment outcomes Y_{jt} for $1 \leq t \leq T_0$, and no other variables.

A natural relaxation of the constraints imposed in the previous problem is to allow weights to take negative values. Chernozhukov et al. (2021) explore one way to do this by redefining the problem as:

$$(\hat{\mu}, \hat{w}) = \arg\min_{(\mu, w)} \sum_{t=1}^{T_0} \left(Y_{1t}^N - \mu - \sum_{j=2}^{J+1} w_j Y_{jt}^N \right)^2$$
$$\text{s.t.} \qquad ||w||_1 \leq 1,$$

where $||w||_1$ represents the l-1 norm of w. The authors refer to this method as constrained Lasso.

An even more general approach described by Chernozhukov et al. (2021) is defined as follows:

$$(\hat{\mu}, \hat{w}) = \arg\min_{(\mu, w)} \sum_{t=1}^{T_0} \left(Y_{1t}^N - \mu - \sum_{j=2}^{J+1} w_j Y_{jt}^N \right)^2 + \mathcal{P}(w),$$

where $\mathcal{P}(w)$ represents a general penalty function that increases the loss when weights have values different from zero. Several penalty functions have been explored in the literature. Some examples include Lasso (Tibshirani, 1996) (see Carvalho, Masini, & Medeiros, 2018; Hollingsworth & Wing, 2022; Li & Bell, 2017), or Lava (Chernozhukov, Hansen, & Liao, 2017).

While the original research describing the synthetic controls methodology relied on applying permutation tests for inference, there have been important advances in developing formal inference tests (e.g., Chernozhukov et al., 2021; Li, 2020). In particular,

Chernozhukov et al. (2021) extend the potential outcomes framework outlined above, and define the following data-generating process:

$$\begin{aligned} Y_{1t}^N &= \hat{Y}_{1t}^N + u_t \\ Y_{1t}^I &= \hat{Y}_{1t}^N + \tau_t + u_t, \end{aligned}$$

for all $t = 1, ..., T_1$, with $\mathbf{E}(u_t) = 0$, and $\{u_t\}$ such that it is a centred stationary stochastic process. This framework allows to test hypotheses on the the effect of the intervention in the post-treatment period $\tau = (\tau_{T_0+1}, ..., \tau_{T_1})'$ of the following form:

$$H_0 : \tau = \tau^0,$$

where $\tau^0 = \left(\tau_{T_0+1}^0, ..., \tau_{T_1}^0\right)'$. The authors propose rewriting the available data under the null hypothesis as $\mathbf{Z} := \mathbf{Z}\left(\tau^0\right) = (Z_1, ..., Z_{T_1})$, where:

$$Z_t = \begin{cases} \left(Y_{1t}^N, ..., Y_{J+1t}^N, X_{1t}', ..., X_{J+1t}'\right)', & t \leq T_0 \\ \left(Y_{1t}^I - \tau_t^0, ..., Y_{J+1t}^N, X_{1t}', ..., X_{J+1t}'\right)', & t > T_0. \end{cases}$$

Under $\mathbf{Z}$, $\hat{Y}_{1t}^N$ is estimated and then used to compute the residuals $\hat{u} = (\hat{u}_1, ..., \hat{u}_T)$. After having found the residuals for each time period, and in order to test the hypotheses defined above, the authors suggest finding a summary statistic that is later used to compute the corresponding p-values for each hypothesis. The test statistic is defined as follows:

$$S(\hat{u}) = S_q(\hat{u}) = \left(\frac{1}{\sqrt{T_*}} \sum_{t=T_0+1}^{T_1} |\hat{u}_t|^q \right)^{1/q},$$

where q denotes the norm being used, and T_* the number of post-intervention periods being considered. The authors consider using block permutations of the residuals in order to compute the p-values, but they clarify that the type of permutation algorithm does not affect the finite sample validity of the procedure as long as residuals are exchangeable. Instead, different permutation regimes allow us to calculate more precise p-values. Following the authors' notation, π is defined as a particular permutation from a set Π, and the permuted p-value is defined as follows:

$$\hat{p} = 1 - \hat{F}\left(S(\hat{u})\right),$$

where $\hat{F}(x) = \frac{1}{|\Pi|} \sum_{\pi \in \Pi} \mathbf{1}\{S(\hat{u}_\pi) < x\}$, and $|\Pi|$ denotes the cardinality of the set Π, and $\hat{u}_\pi$ the vector of residuals computed under permutation $\pi \in \Pi$.

For simplicity, and following the applied example in Chernozhukov et al. (2021), the permutation set adopted in this exercise is a singleton defined over the identity permutation. The p-values are computed for each post-intervention period individually ($t^* \in \{T_{0+1}, ..., T_1\}$), such that each post-intervention period is added to the full set of pre-intervention outcomes on its own. Furthermore, the norm adopted is the l-1 norm, and the full set of pre-intervention outcome measures is used in $\mathbf{X}_j$. This simplifies the calculations, yielding the following definition for the p-values associated with each hypothesis:

$$\hat{p} = \frac{1}{T_0 + 1} \sum_{t \in \underline{t} \cup t^*} \mathbf{1}\{|\hat{u}_t| \geq |\hat{u}_{t^*}|\},$$

with $\underline{t}$ denoting the full set of pre-intervention periods $\{1, ..., T_0\}$. Under these assumptions, the p-value is essentially defined as the proportion of residuals that are larger in magnitude than those for the specific post-intervention period being tested under the null $\mathbf{Z}$. This is the approach that I adopt for the analyses reported in Section 3.5 and Appendix 3.A.

Finally, in order to construct pointwise $(1 - \alpha)$ confidence intervals for each post-intervention period, the authors suggest doing it by test inversion using the following algorithm:

Algorithm 1: (Pointwise Confidence Intervals)

1. Choose a grid of G candidate values $\tilde{\mathscr{T}}_t = \{\tilde{\tau}^0_{1t}, ..., \tilde{\tau}^0_{Gt}\}$.
2. For each $\tilde{\tau}^0 \in \tilde{\mathscr{T}}_t$, define $\mathbf{Z}$ and compute $H_0 : \tau_t = \tilde{\tau}^0_t$, calculating the corresponding p-value, $\hat{p}(\tilde{\tau}^0_t)$, as above.
3. Return the $(1 - \alpha)$ confidence set: $\mathscr{C}_{(1-\alpha)t} = \left\{\tilde{\tau}^0_t \in \tilde{\mathscr{T}}_t : \hat{p}(\tilde{\tau}^0_t) > \alpha\right\}$.

3.4 Data Sources and Implementation

Crime data. For all reported crimes, I use a monthly time series database of criminal incidents compiled by the Mexican government and available online through the government's website.[14] This database aggregates the number of crimes reported in each of the 32 states of the country and Mexico City between January 1997 and December 2017. Each crime is categorised under a specific 'modality'. Within each 'modality', there are several crime-types and within each type, the government identifies sub-types. For

[14] Two popular datasets used to analyse conflict are the Uppsala Conflict Data Program (Davies et al., 2025) or the Armed Conflict Locations and Events Data (ACLED) (Raleigh et al., 2023). Unfortunately, ACLED only has data for Mexico from 2018 onwards. While Uppsala's does have data on armed conflicts in Mexico from 1989 to 2021, only 82 instances are recorded for the state of Coahuila between these years. Even in the most violent years, the database records only 18 instances of conflict for 2012, and 16 for 2013.

instance, the 'modality' of homicides is divided into 'culpable homicide' [*homicidio culposo*] and 'intentional homicide' [*homicidio doloso*]. Each of these types of homicides are then divided into sub-types classifying how the crimes were committed e.g., with a 'firearm', 'bladed weapon', 'other' or 'no data'.

Throughout the paper, I refer to crime-types using literal translations of their Spanish names. Crimes classified as *culposo* ('culpable') denote a lack of intent, as opposed to *doloso*, which conveys pre-meditation and intentionality. In this way, 'culpable homicides' would be akin to manslaughter, and 'intentional homicides' to murder. Similarly, 'culpable lesions' [*lesiones culposas*] and 'intentional lesions' [*lesiones dolosas*] would be close to accidental harm, and grievous bodily harm, respectively.

Given the variability of the data, I aggregate each crime at the second classification level i.e., the level of the crime-type. Some examples of the different crime-types include the aforementioned 'culpable' and 'intentional homicides', but also kidnapping, 'culpable' and 'intentional lesions', or common theft with- and without-violence. In order to smoothen the time series, I calculate the 12-month moving average for each crime-type, as suggested by Abadie (2021).

Economic outcomes. I retrieve a set of high frequency state-level economic indicators from Mexico's National Statistics Institute (INEGI). These variables comprise indices for primary, secondary, tertiary, and total economic production; total exports; total economically active population, total non-economically-active population, total non-economically-active available population; total occupied population, total occupied population by income brackets; total informally employed population; total unemployed population, and the unemployment rate. These outcomes are reported on a quarterly basis. As above, to smoothen the series, I transform all of them to a yearly moving average.

Implementation. Following recent advances in the synthetic controls literature, I opt to implement the most general optimization approach as described in Chernozhukov et al. (2021) and choose Lasso (Tibshirani, 1996) as a penalty function. This approach has been used by Li and Bell (2017), Carvalho et al. (2018), and more recently by Hollingsworth and Wing (2022). The approach places few restrictions on the weights assigned to donor units. Namely, it allows weights to take negative values and relaxes restricting their sum. To allay concerns around over-fitting *a priori*, I rely on a 10-fold cross-validation procedure (Hollingsworth & Wing, 2022), which I implement using Stata's lassopack (Ahrens, Hansen, & Schaffer, 2020). For calculating confidence intervals, I rely on the methodology outlined in Chernozhukov et al. (2021) and reviewed in Section 3.3. I provide the programmes written for Stata to estimate the synthetic control and calculate the policy impact, as well as computing confidence intervals for all post-treatment periods, in Appendix 3.C. As a

complement to these results I use the canonical synthetic controls method (Abadie et al., 2010; Abadie & Gardeazabal, 2003) and replicate the analysis under a synthetic differences-in-differences framework (Arkhangelsky et al., 2021; Clarke, Pailañir, Athey, & Imbens, 2024). All results are aligned with those using cross-validated Lasso. All calculations were carried out in StataSE 18.5.

3.5 Main Results

3.5.1 'Intentional Homicides'

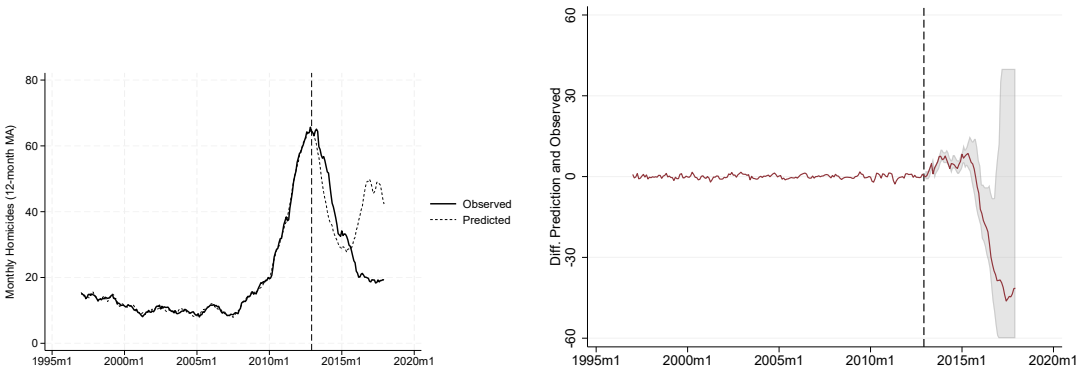
To measure the impact that the set of policies had on reducing 'intentional homicides', I generate the synthetic control by running a 10-fold cross-validated Lasso on the time series of 'intentional homicides' from January 1997 up until January 2013. Panel (a) in Figure 3.2 shows that the model is capable of tracking the observed data very closely. The vertical line indicates the application of the policy. From the vertical line onward, the observed and predicted data differ. The difference between these two lines is interpreted as the effect of the policy and is presented in Panel (b) of Figure 3.2. The figure suggests that the set of policies was indeed successful in reducing 'intentional homicides' in Coahuila. Column 1 of Table 3.A.1 reports the donor units and their corresponding coefficients (weights) when calculating the synthetic control.

The reduction in 'intentional homicides' relative to the counterfactual took several months to happen, as the difference between the observed and the estimated data becomes negative after October 2015. For the rest of the months for which the data is available the difference remains negative and it grows in magnitude reaching a reduction of up to 35 monthly homicides compared to the counterfactual. Panel b of Figure 3.2 also adds 90% confidence intervals calculated using Algorithm 1. This figure suggests that the observed reduction in 'intentional homicides' indeed was statistically different from zero, relative to the estimated counterfactual – albeit for a short period of time. This evidence also confirms that Coahuila had an unprecedented increase in 'intentional homicides', just as the new government arrived in office.[15]

3.5.2 Other Crime-Types

This section expands the analysis to consider whether the policies had an impact on other types of crimes. Tables 3.A.2 and 3.A.3 in the Appendix present cross-correlation tables for

[15] There are several ways of probing the reliability of these results, as discussed in Abadie et al. (2010, 2015). I devote Section 3.6 to those analyses.



(a) Observed Data and Synthetic Control

(b) Difference Observed Data and Synthetic Control with 90% CIs

Notes: Panel (a) shows the series of 'intentional homicides' in Coahuila, as well as the synthetic control created using the rest of the states in Mexico. Panel (b) shows the difference between these two series: the effect of the policies, along with 90% confidence intervals calculated using Algorithm 1 from Chernozhukov et al. (2021). Coarseness of grid defined as $1/21$ of the grid's size. Finer grids generate smoother confidence intervals.

Figure 3.2: Policy Effect on Monthly 'Intentional Homicides'

all types of crime in Coahuila. The first table uses the full sample, while the latter focuses on the dates between 2008 and 2012. This period coincides with the dates in which the yearly percentage changes in monthly homicide rates in the state were highest, as seen in Panel (b) of Figure 3.1. I find correlations greater than $\rho = 0.500$ between 'intentional homicides' and kidnapping ($\rho = 0.559$), extortion ($\rho = 0.519$), violent car theft ($\rho = 0.880$), violent bank robbery ($\rho = 0.543$), violent robbery ($\rho = 0.946$), non-violent robbery ($\rho = 0.593$), rape ($\rho = 0.629$), other sexual crimes ($\rho = 0.523$), and threats ($\rho = 0.956$). Most of these crimes can be easily associated with the presence and operation of drug cartels, as discussed in Section 3.2. If the set of policies enacted by the government of Coahuila were holistically successful we would expect to see a decline in these crimes as well.

The policy impacts on the rest of crime-types are presented in Figure 3.A.1. Results across other crime-types suggest a mixed picture. Out of the crimes that are highly correlated with 'intentional homicides', only violent car theft and non-violent robbery show a decrease relative to the counterfactual. Kidnapping, extortion, violent bank robbery, violent robbery, child abuse, and other sexual crimes do not suggest changes relative to the counterfactual, whereas threats suggest a steep increase.

Figure 3.A.3 complements the analysis as it presents the Mean-Squared Predicted Error (MSPE) ratios relative to the benchmark model of 'intentional homicides' for the rest of the types of crime simply as a comparison of the models' fit.[16] Models with a negative

[16] Note that the vertical axes of all Panels in Figure 3.A.3 are logarithmic to allow for an easier comparison between models. The need for a logarithmic scale to visualise the results speaks of the great heterogeneity in

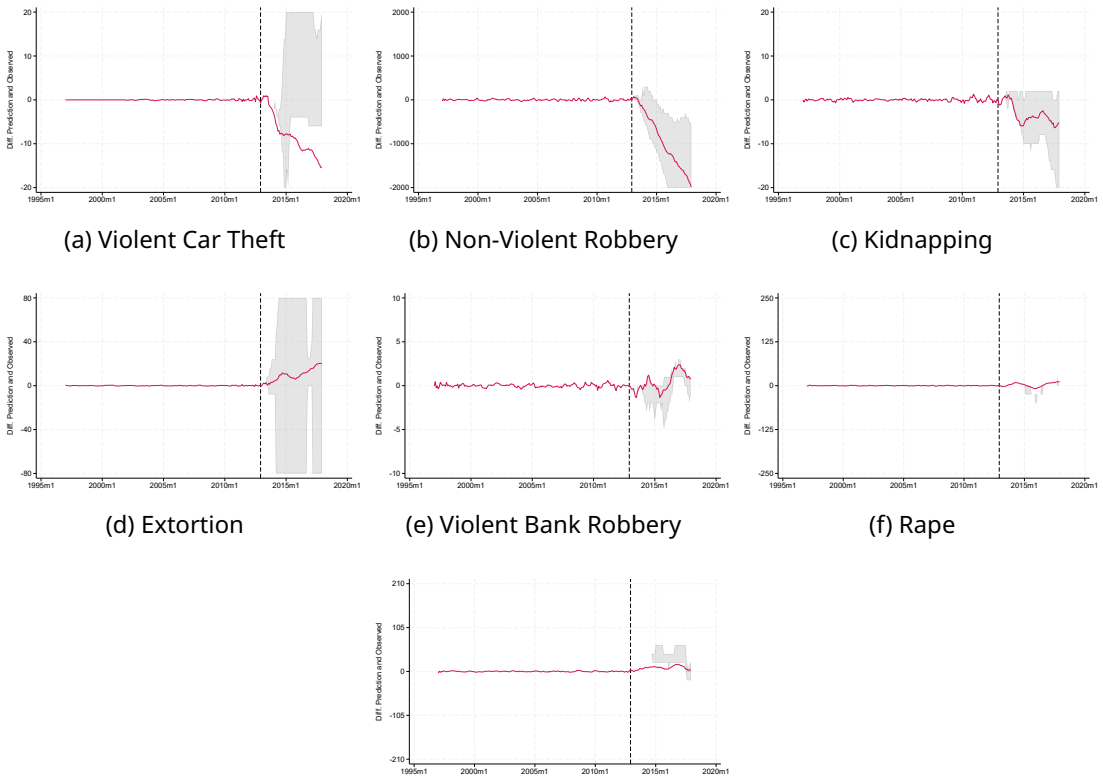
natural log MSPE ratio fit the data better than the benchmark model for 'intentional homicides', while positive values reflect a worse performance than the benchmark. For interpretation, I opt to give higher credibility to models whose MSPE ratio is 10 or lower (2.302 in the logarithmic scale). This cautionary cut-off would lead us to only consider the models for 'culpable homicides', 'culpable lesions' (*lesiones culposas*, akin to accidental harm), kidnapping, breach of trust, dispossession, extortion, violent and non-violent car theft, violent and non-violent bank robbery, livestock theft, rape, child abuse, and other sexual crimes.

Altogether, on top of the impact on 'intentional homicides', these analyses suggest a statistically significant decrease in some activities associated with drug cartels, i.e., violent car theft, and non-violent robbery. At the same time, I find no evidence of a decrease in kidnapping, extortion, violent bank robbery, rape, and other sexual crimes. Figure 3.A.1 also suggests crime-types weakly correlated with 'intentional homicides' were in general unaffected by the policy (e.g., breach of trust, child abuse, and 'culpable homicides' perhaps surprisingly). Some of these findings contrast with the results discussed in Aguilar Valenzuela and Moreira Valdez (2022), where the authors document a stark reduction in 'intentional homicides', kidnapping, extortion, and theft. The evidence presented by these authors corresponds to a comparison in levels across time, not taking into account an overall decline in criminal activity close to 2015. By providing a valid counterfactual, the findings in this section paint a more modest, albeit still positive, picture of the success of this set of policies.

3.5.3 The Economic Effects of Peace

Next, I apply the synthetic controls methodology to a set of economic outcomes. Figure 3.4 shows the effect on these outcomes relative to their synthetic counterfactuals. Figure 3.4 is complemented by Figure 3.A.2 in the appendix, where I calculate confidence intervals for each of these economic indicators following Algorithm 1. Figure 3.4 suggests some movement on these indicators that could be mechanically related with the implementation of the package of policies. For instance, as a consequence of the banning of casinos, brothels, and junk-yards, we could observe a reduction in the number of informal workers, an increase in the unemployment rate, or an increase in the number of unoccupied individuals. However, the post-treatment series for all of these indicators are statistically indistinguishable from zero as suggested in Figure 3.A.2. This lack of significance across the board might arise from over-fitting the models, since the number of observations is much

performance that exists between these models.



(a) Violent Car Theft (b) Non-Violent Robbery (c) Kidnapping

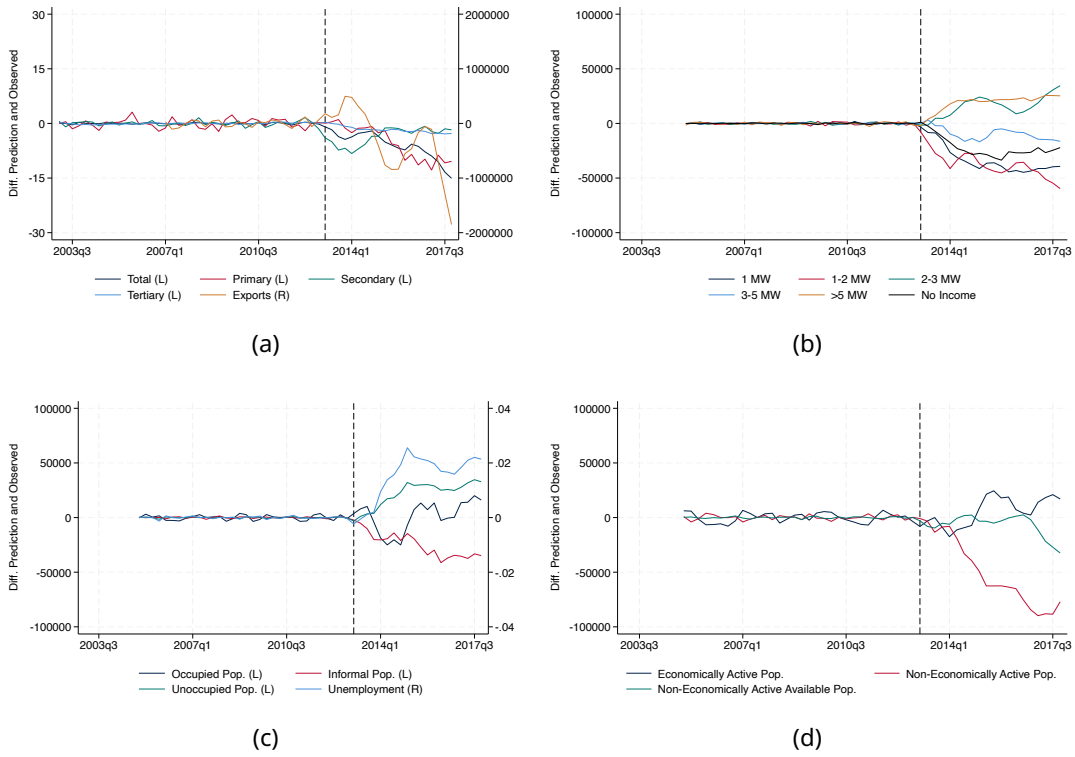
(d) Extortion (e) Violent Bank Robbery (f) Rape

(g) Other Sexual Crimes

Notes: The figure presents the effect of the set of policies implemented in Coahuila on several crime-types closely related to the presence of drug cartels. Each panel corresponds to one specific crime-type. The figures include confidence intervals calculated using Algorithm 1 from Chernozhukov et al. (2021). The policy was successful at reducing some crime-types associated with cartels, such as violent car theft (a) or non-violent robbery (b). The policies were not successful at reducing other crime-types such as kidnapping (c), extortion (d), violent bank robbery (e), rape (f), and other sexual crimes (g). The figures include 90% confidence intervals. Coarseness of grid defined as $1/21$ of the grid's size. Finer grids generate smoother confidence intervals.

Figure 3.3: Policy Effect on Other Crime-Types with Confidence Intervals

smaller than for the crime outcomes. When checking for effects on these outcomes using canonical synthetic controls under a synthetic differences-in-differences framework,[17] I find similar results in the trends, but I do observe statistical significance for some outcomes (see Figure 3.6). The analysis captures a decrease in the number of workers earning between 0 and 2 times the minimum wage, and a small reduction in informality, complemented by an increase in workers earning over 2 times the minimum wage. Likewise, I observe a small decline in tertiary-sector activities. All of these changes can be rationalized as a consequence of the specific policies adopted in Coahuila, rather than as a consequence of reducing cartel-related violence.



Notes: The Figure shows the impact of the policies implemented in Coahuila on a series of economic outcomes. The Figure is complemented with Figure 3.A.2 that shows confidence intervals for each of these indicators calculated using Algorithm 1 from Chernozhukov et al. (2021). This Figure suggests some mechanical effects from the policies, such as a reduction in informality or an increase in unemployment (c). However, these effects are not statistically different from zero.

Figure 3.4: Policy Effect on Economic Outcomes

3.6 Robustness Checks

In this section, I conduct a series of tests to probe the robustness of the results presented in Section 3.5. I also include more details on the synthetic differences-in-differences

[17] See Section 3.6 for more detail on robustness checks.

framework I adopt to complement the main results.

3.6.1 Placebo Policy Dates

I assign placebo policy dates and evaluate the capacity of the model to track the actual data, particularly in the periods after the placebo policy date and before the actual policy date. Figure 3.A.4 implements these placebo policy dates. The observed data is represented with a continuous line, while the synthetic controls are represented with dashed lines. The vertical dotted lines represent the dates of implementation of the real and placebo policies.

The models with placebo policy dates perform worse at tracking the actual data than the model with the correct date. The models whose policy dates are 2010 and 2008 predict a peak in homicides much lower than the actual peak. The difference between the actual data and these counterfactuals remains positive even after 2015, which could lead us to conclude that the policy was actually not successful at reducing homicides relative to the counterfactual. A similar situation arises when we analyse the synthetic controls generated with policy dates set in 2006 and 2004. In these cases, the models predict a peak in homicides higher than the actual data. In both of the latter cases, the predicted and observed data begin to differ right after 2008. The date coincides with an increase in violence associated with drug cartels due to the government's change in strategy to combat drug production and smuggling. Ultimately, these differences arise from the set of states and the respective weights that are being 'chosen' by the Lasso–namely, from including states such as Aguascalientes, Campeche, Michoacán, Yucatán, or Chiapas, which are not included in the base model. Table 3.A.1 in the Appendix details the included states and employed weights in each of these models.

3.6.2 Swapping Treated and Donor Units

Another test of the model's performance is to assign treatment to units from the donor pool while moving the treated unit to the donor pool and look for an effect similar to what is observed on the treated unit. These placebo policies are presented in Figure 3.A.5. As is evident from Panel (a) in the Figure, the model's performance varies greatly across the set of units. To narrow down the sample of placebo tests being considered, Abadie et al. (2010, 2015) suggest calculating the mean squared predicted error (MSPE) for each of these models, finding the ratio of the MSPE of the placebo units and the unit that did receive the treatment, and setting an upper threshold for this value to be considered part of the sample. Panels (b), (c) and (d) explore setting this limit at 10, 5 and 2, respectively.

Naturally, as the threshold becomes stricter, the sample is reduced to models that fit the data better. Already in Panel (b) we can observe that the policy impact for the state of Coahuila is one of the strongest effects being identified, particularly for the post-2016 period. Panels (c) and (d) confirm this observation, with no other state showing a stronger effect than Coahuila in Panel (d). Since there are 11 control states left in this exercise apart from Coahuila, this would suggest that under a random permutation, the probability of finding an effect as large as the one observed for Coahuila is approximately 8%, which would be under a comfortable level of statistical significance.

3.6.3 Spillovers to Donor Units

One final concern related to how synthetic controls are built is related to the donor pool considered when setting up the model. In general, we would like to ensure that there are no spillovers between the treated unit and the units comprising the donor pool. However, in most applications, the units whose outcomes most resemble those of the treated unit, could be subject to spillovers after the implementation of the policy. An example would be units that are geographically close to each other, where there is free mobility across units. In this particular setting, a valid concern would be the presence of the same drug cartels or the existence of conflict between competing cartels. Given that the model is agnostic to this potential pitfall, the counterfactual would be biased in these situations. We could observe either positive or negative spillovers. For the case of positive spillovers, we would be overestimating the counterfactual and thus underestimating the effect of the policy. Negative spillovers could arise from a cartel increasing the intensity of its operations in another state after the implementation of the policies designed to curb their activities in Coahuila.

In order to tackle this concern, I first take a data-driven approach to restrict the donor pool before the synthetic control is generated. I do this by finding the linear correlation coefficient between the time series for monthly homicides in Coahuila and the other 31 states in the donor pool. I then calculate their percentiles according to the absolute value of these coefficients and restrict the sample using different cut-off percentiles. The results are presented in Figure 3.A.6. Panel (a) shows the different counterfactuals, and Panel (b) the effect associated with each of these. Naturally, setting a stricter restriction comes at the expense of the model's performance. By excluding from the donor pool states whose time series are closely related to that of Coahuila, the model's ability to track the performance of Coahuila's data through time worsens. This is quantified in Panel (c) from Figure 3.A.6, which shows the MSPE ratio of the models with a restricted donor pool relative to that

with the full pool.

As expected, the models built on a pool which excludes the top 10% and 15% of the units most closely related to Coahuila's data find effects that are slightly greater in magnitude than those found with the model with the full pool. In turn, the model in which the top 20% is excluded finds a smaller effect, but with a much worse performance relative to the other two. The model excluding the top quintile and the rest of those considered have an MSPE ratio greater than 2 and even 5, which should be a warning about over-trimming our donor pool. For robustness, Figures 3.A.7 and 3.A.8 report the same analyses as in Figures 3.A.1 and 3.A.3 while trimming the top 10% of the donor pool. Results are robust to trimming the sample.

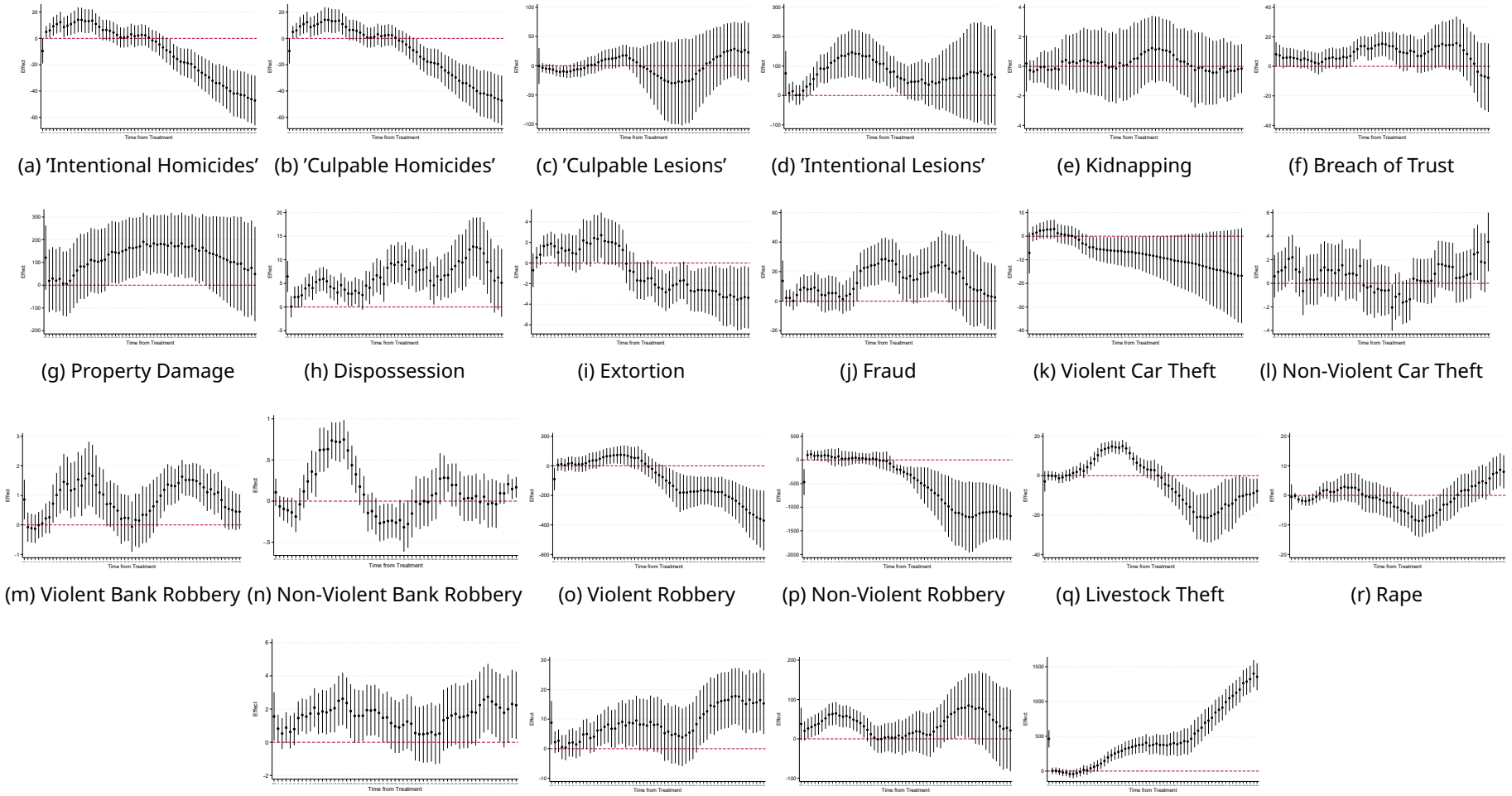
Next, I consider geographic spillovers to adjacent states. I recompute the basic results from Figure 3.2 by excluding the states that share a border with Coahuila from the donor pool – namely, Chihuahua, Durango, Zacatecas, San Luis Potosí, and Nuevo León. Figure 3.A.9 reports the results from this exercise. The estimated effect is remarkably close to the effect computed on the full pool of states.

3.6.4 Comparison to Canonical Synthetic Controls

To check the overall robustness of the approach followed so far, I rerun all analyses on crime-type using the original methodology in Abadie and Gardeazabal (2003) – the canonical synthetic controls method. Specifically, I rely on the recent application from Arkhangelsky et al. (2021) to perform inference on all post-treatment periods using a synthetic differences-in-differences approach.[18] I report all results in Figures 3.5 and 3.6 for all crimes-types and economic outcomes, respectively. Each panel in each figure corresponds to a different outcome. The first coefficient from each panel is the ATT for the entire post-treatment period. Results are broadly in line with those reported in Section 3.5. The main disadvantage of using the canonical framework compared to the approach followed to produce the results from Section 3.5 is in the closeness of the synthetic control's fit. For instance, for the baseline result on 'intentional homicides', the canonical framework produces a MSPE of 129.27, while the MSPE using cross-validated Lasso is 0.31.[19]

[18] I implement this alternative analysis using the package sdid_event in Stata (Clarke et al., 2024).

[19] I implemented the canonical framework using the allsynth package in Stata (Wiltshire, 2025). For this application, I restricted the sample to 11 years' worth of pre-treatment observations i.e., from January 2002 to December 2012.



(a) 'Intentional Homicides' (b) 'Culpable Homicides' (c) 'Culpable Lesions' (d) 'Intentional Lesions' (e) Kidnapping (f) Breach of Trust

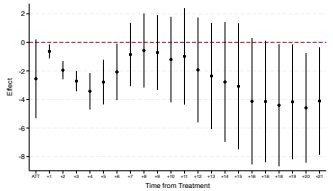
(g) Property Damage (h) Dispossession (i) Extortion (j) Fraud (k) Violent Car Theft (l) Non-Violent Car Theft

(m) Violent Bank Robbery (n) Non-Violent Bank Robbery (o) Violent Robbery (p) Non-Violent Robbery (q) Livestock Theft (r) Rape

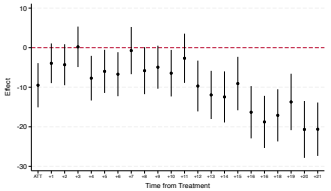
(s) Child Abuse (t) Other Sexual (u) Threats (v) Other

Notes: The figure shows the effect of the policy on all crime-types calculated using the canonical synthetic controls methodology in a synthetic differences-in-differences framework following Arkhangelsky et al. (2021). Calculations were performed in Stata using the sdid_event package (Clarke et al., 2024). The first coefficient in each figure corresponds to the ATT for the entire post-treatment period. The figure includes 95% confidence intervals. Results are aligned with those from section 3.5.

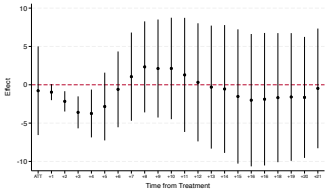
Figure 3.5: Synthetic Difference-in-Differences using Canonical Synthetic Controls on All Crime-Types



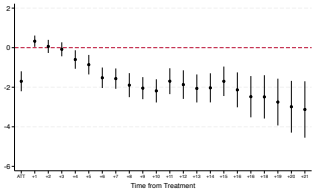
(a) Total Economic Activity



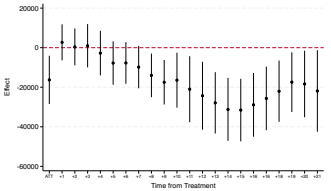
(b) Primary Activities



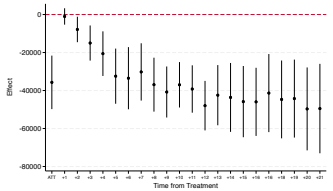
(c) Secondary Activities



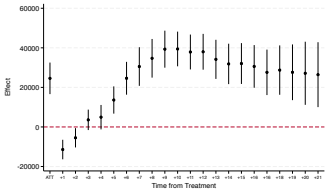
(d) Tertiary Activities



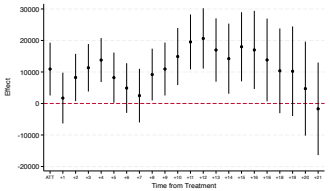
(e) People Earning $<$ 1 MW



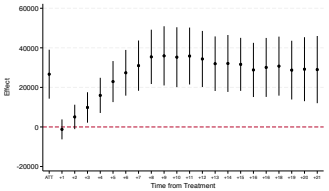
(f) People Earning 1-2 MW



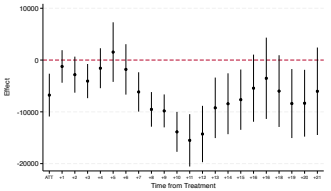
(g) People Earning 2-3 MW



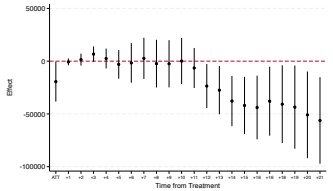
(h) People Earning 3-5 MW



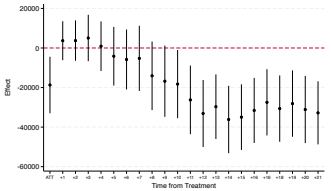
(i) People Earning $>$ 5 MW



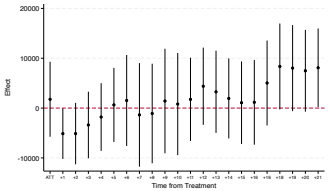
(j) People Earning No Income



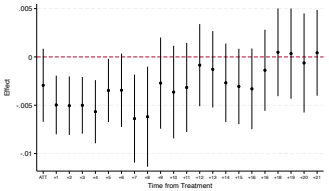
(k) Occupied Population



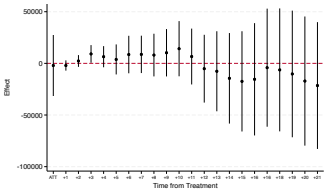
(l) Informal Population



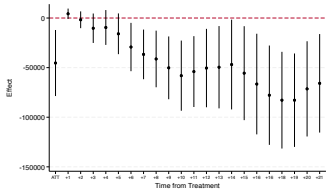
(m) Unoccupied Population



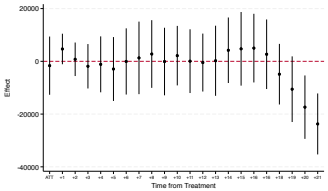
(n) Unemployment Rate



(o) Economically Active Pop.



(p) Non-Ec. Active Pop.



(q) Non-Ec. Active Available Pop.

Notes: The figure shows the effect of the policy on all crime-types calculated using the canonical synthetic controls methodology in a synthetic differences-in-differences framework following Arkhangelsky et al. (2021). Calculations were performed in Stata using the sdid_event package (Clarke et al., 2024). The first coefficient in each figure corresponds to the ATT for the entire post-treatment period. The figure includes 95% confidence intervals.

Figure 3.6: Synthetic Difference-in-Differences using Canonical Synthetic Controls on All Economic Outcomes

3.7 Conclusions and Limitations

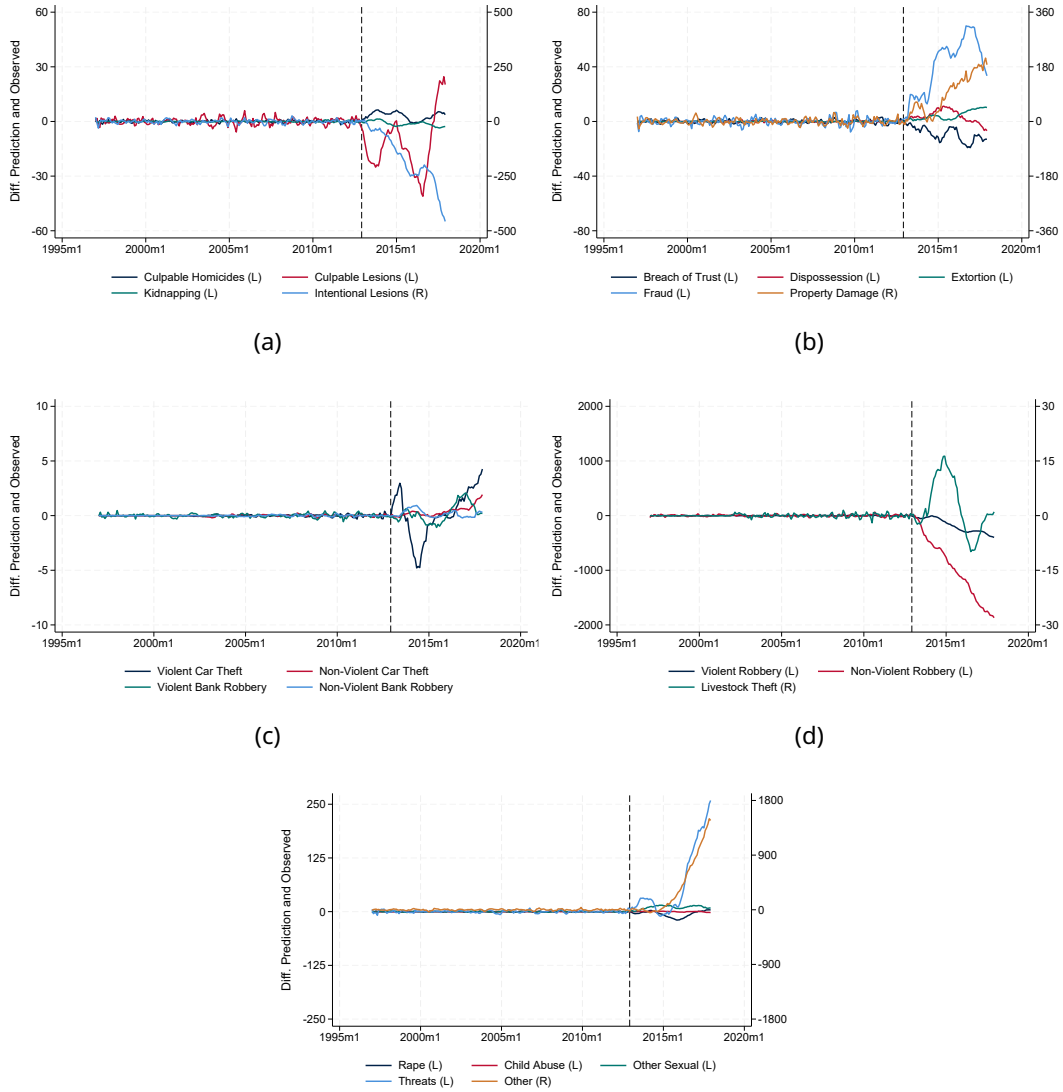
This paper explores two main questions: whether the bottom-up strategy implemented by the government of Coahuila to combat drug cartels at the end of 2012 was effective in reducing the incidence of various types of crimes; and whether this strategy had any economic spillovers. The paper relies on the synthetic controls methodology to assess these questions. In particular, it relies on recent advances that allow researchers to compute confidence intervals for each observation in the post-treatment period. The paper provides a technical contribution by adapting this methodology for users of Stata.

The findings of the paper suggest that there was a statistically significant reduction in 'intentional homicides' during the post-treatment period. When extending the analysis to other crime-types that were not explicitly targeted by the government of Coahuila, I find suggestive evidence of further reductions in crimes correlationally related to the presence of drug cartels, such as violent car theft and non-violent robbery. However, I see no impact on other crime-types closely related to cartels such as kidnapping or extortion. Further, I document no impact on crime-types unrelated to drug cartels. Overall, these findings suggest that the set of policies were modestly successful in reducing crime associated with cartels. Consistent with the literature, I observe no statistically significant impacts on economic indicators.

These findings should be strongly caveated on grounds of data availability and integrity. So far, this article has taken at face-value the implementation of the set of policies outlined by Coahuila's government. Being unable to corroborate the implementation of these policies is a key limitation. Also, being unable to validate the data on crimes is another limitation. This is particularly salient for instance for crimes classified under *other*, for which the data documents a steep increase following 2012 (see Figure 3.A.10). Further, the synthetic controls methodology forces us to remain relatively agnostic about phenomena closely related to conflict, such as migration, as well as potential spillovers to other geographic locations. Finally, notwithstanding the results presented, this paper has remained largely agnostic about the potential underlying mechanisms that might help generalize these findings to other contexts.

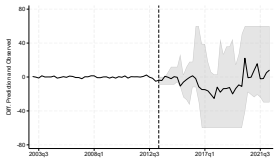
Appendix

3.A Additional Figures

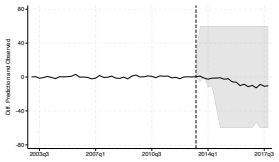


Notes: The figure shows the effect of the policy on all crime-types apart from 'intentional homicides'. Each series corresponding vertical axis is denoted in parenthesis as part of each panel's legend.

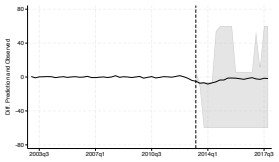
Figure 3.A.1: Policy Effect on Other Crime-Types



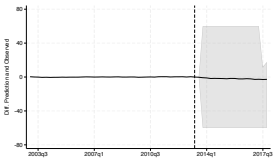
(a) Total Economic Activity



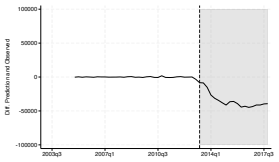
(b) Primary Activities



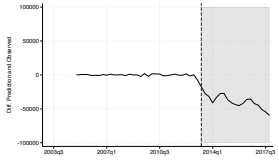
(c) Secondary Activities



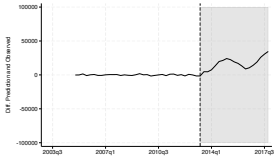
(d) Tertiary Activities



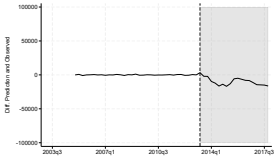
(e) People Earning $<$ 1 MW



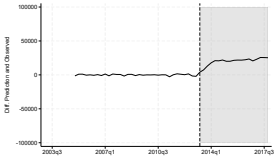
(f) People Earning 1-2 MW



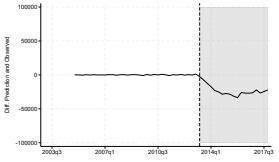
(g) People Earning 2-3 MW



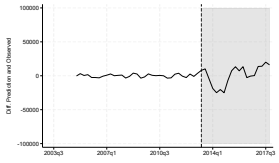
(h) People Earning 3-5 MW



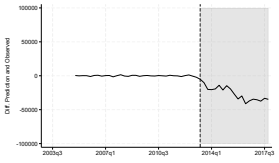
(i) People Earning $>$ 5 MW



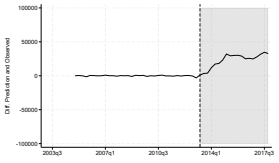
(j) People Earning No Income



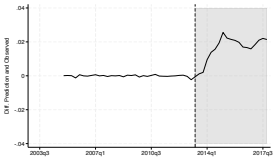
(k) Occupied Population



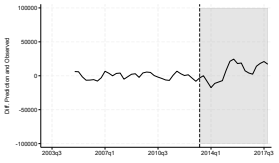
(l) Informal Population



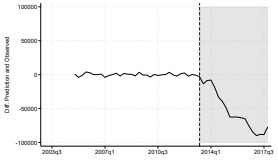
(m) Unoccupied Population



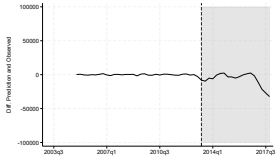
(n) Unemployment Rate



(o) Economically Active Pop.



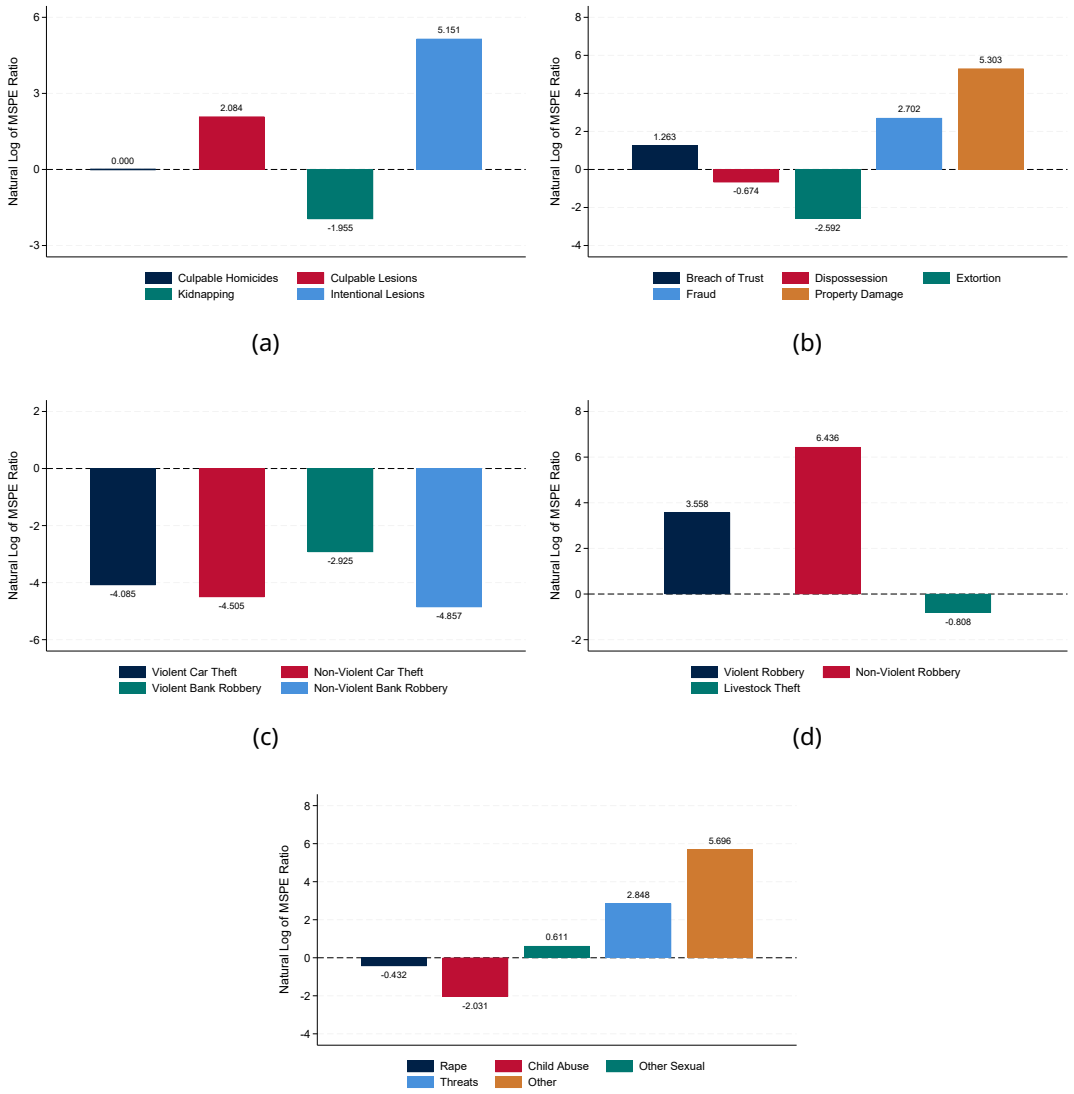
(p) Non-Ec. Active Pop.



(q) Non-Ec. Active Available Pop.

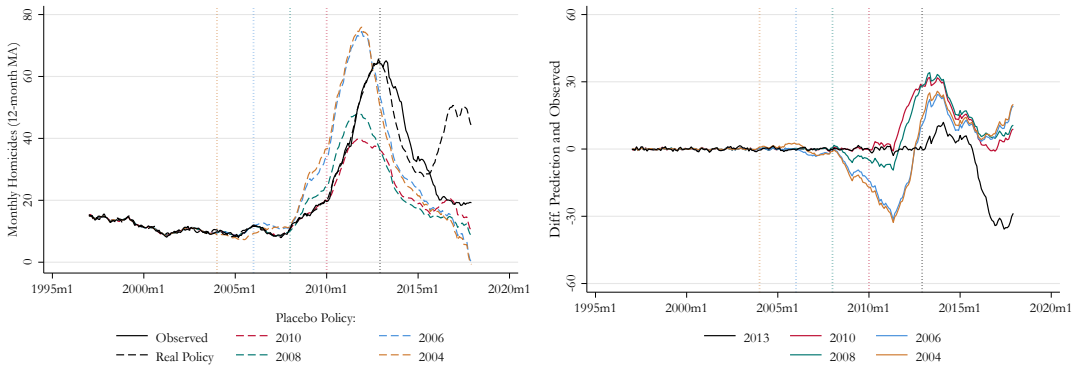
Notes: Policy effect on all economic outcomes with confidence intervals calculated using Algorithm 1 from Chernozhukov et al. (2021). The figures include 90% confidence intervals. Coarseness of grid defined as $1/20$ of the grid's size. Finer grids generate smoother confidence intervals.

Figure 3.A.2: Policy Impact on Economic Outcomes with Confidence Intervals



Notes: Figure denotes the natural logarithm of the MSPE Ratios of each model relative to the MSPE for 'intentional homicides'. These ratios denote a comparison of the models' fits. Models with negative log-ratios fit the data better than the 'intentional homicides' benchmark, while positive values denote a worse fit.

Figure 3.A.3: Natural Logarithm of MSPE Ratios Relative to MSPE for 'Intentional Homicides'

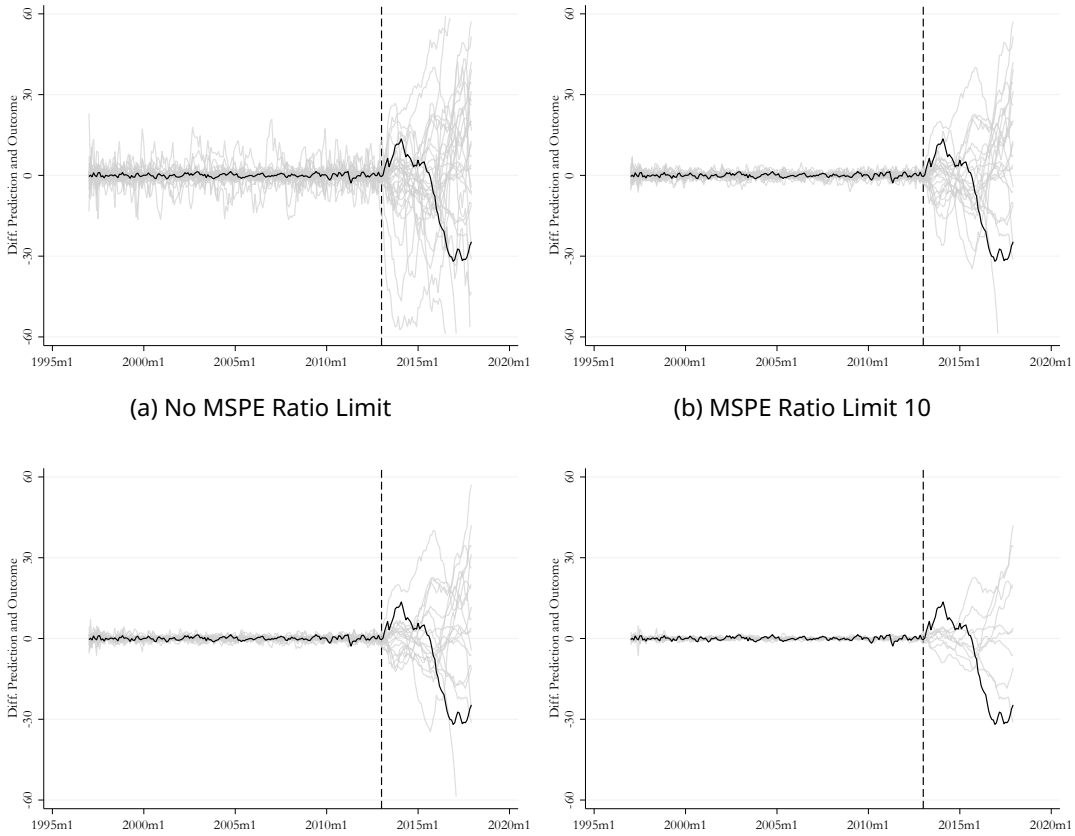


(a) Observed Data and Synthetic Controls

(b) Difference Observed Data and Synthetic Controls

Notes: The figure shows the performance of the synthetic control using placebo policy dates. Having fewer observations in the pre-treatment period worsens the performance of the models. Using an earlier date would lead us to conclude that ultimately the policy was not effective. This is mainly due to from including donor units such as Aguascalientes, Campeche, Michoacán, Yucatán, or Chiapas, which are not included in the base model, as Table 3.A.1 documents.

Figure 3.A.4: Model Evaluation Through Placebo Policy Dates



(a) No MSPE Ratio Limit

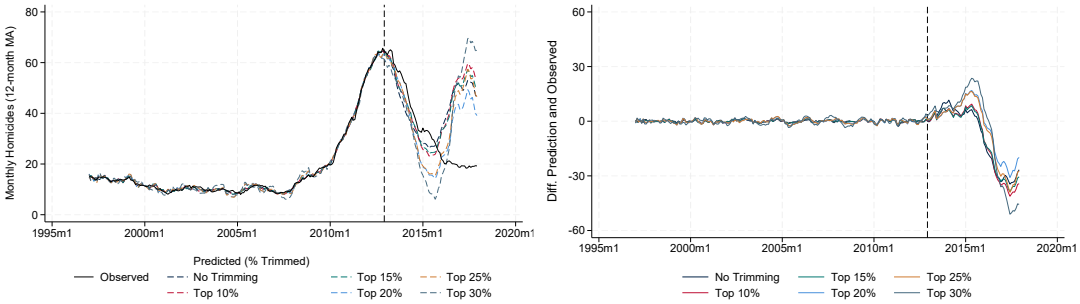
(b) MSPE Ratio Limit 10

(c) MSPE Ratio Limit 5

(d) MSPE Ratio Limit 2

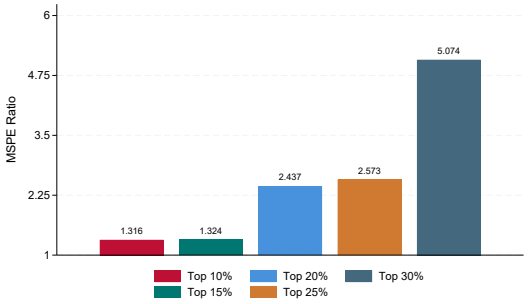
Notes: The figure shows effects on placebo policies on the control states compared to Coahuila. Each panel restricts the sample using different MSPE ratio thresholds. In panel (d), the sample is reduced to 11 states and the effect seen for Coahuila is the largest one. This suggests that the probability of estimating an effect as large as the one for Coahuila under a random permutation in this sample is close to 8%.

Figure 3.A.5: Placebo Policies on Control Units by MSPE Ratio Limits



(a) Observed Data and Synthetic Controls

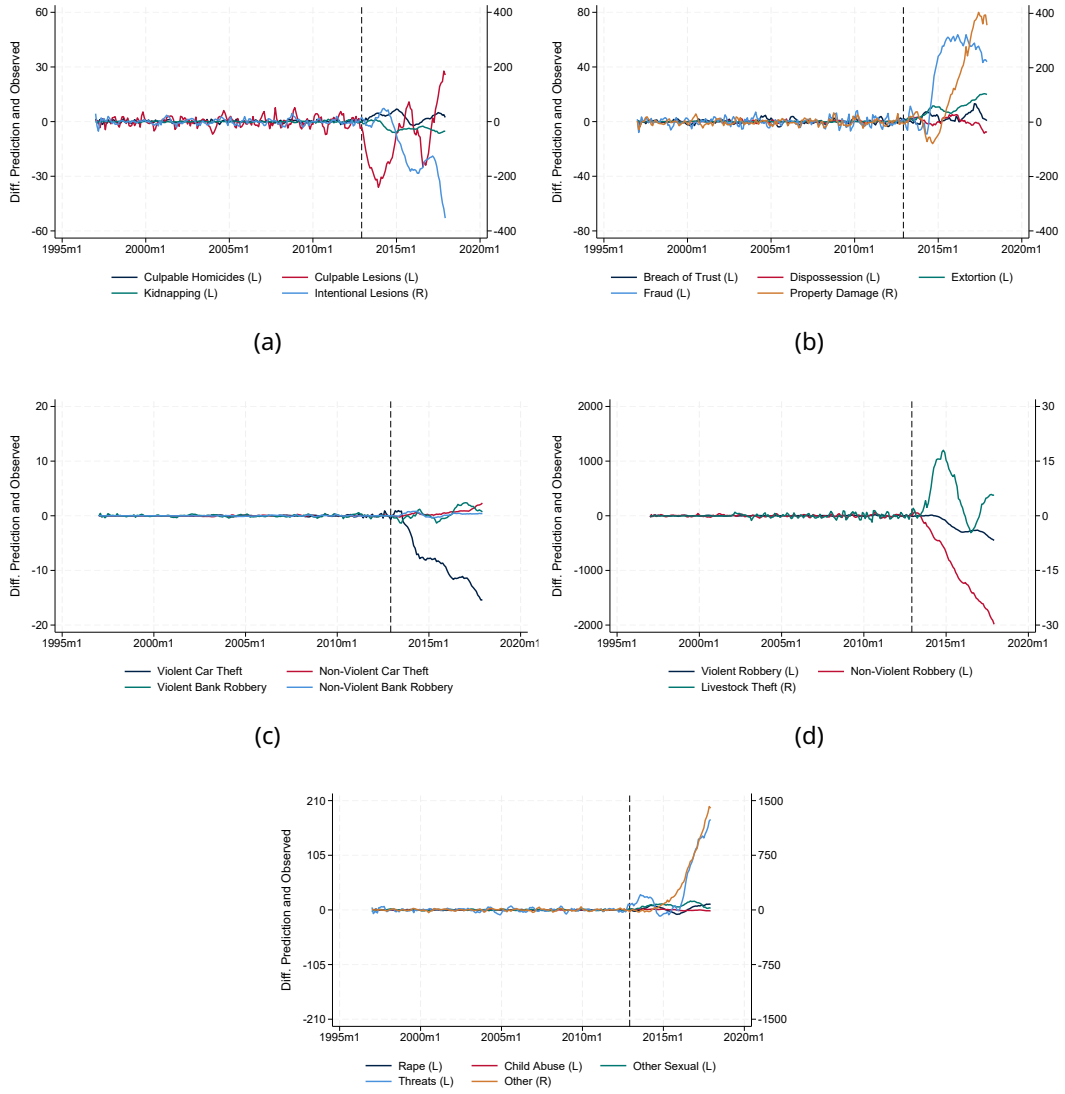
(b) Difference Observed Data and Synthetic Controls



(c) MSPE Ratios

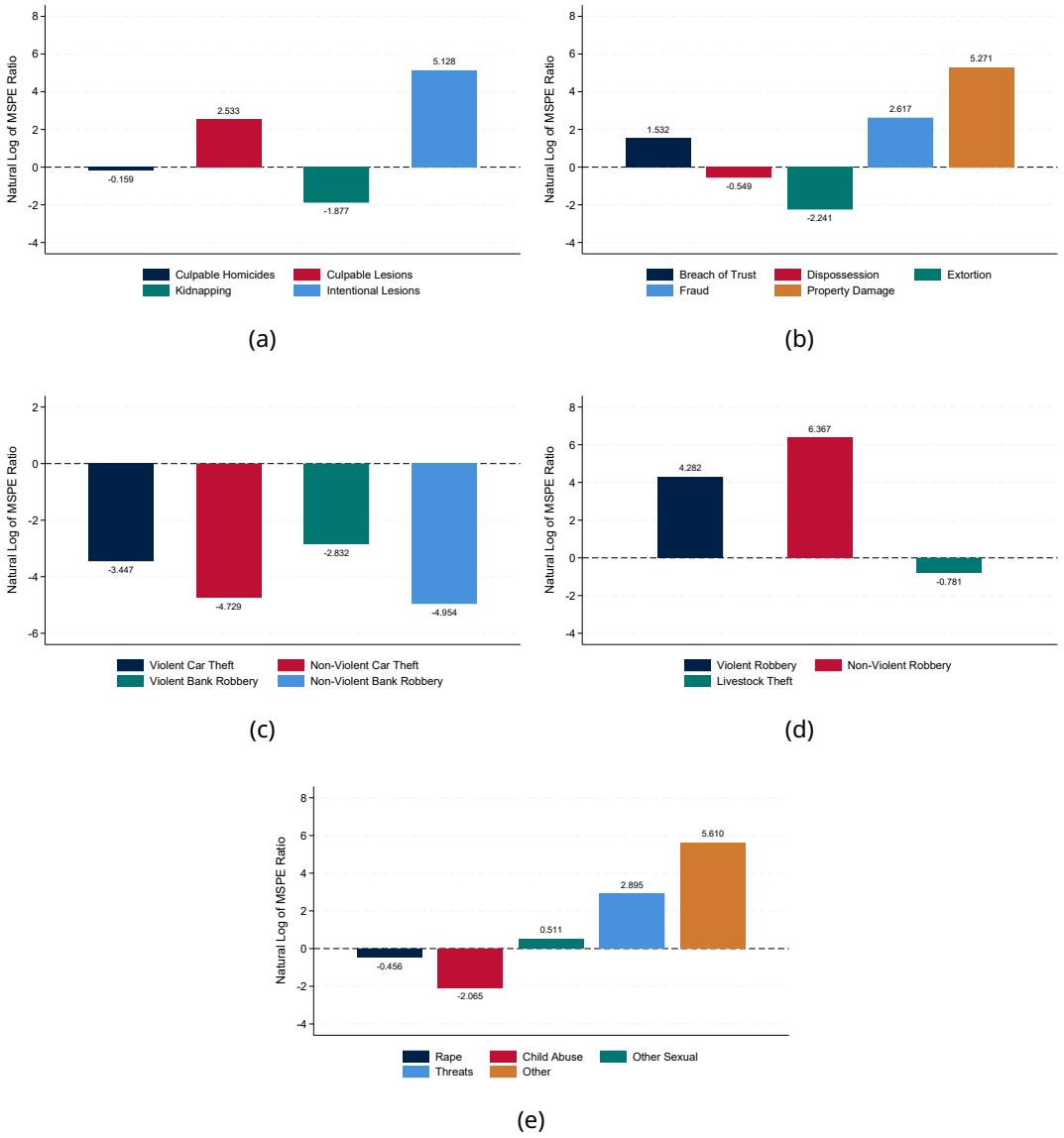
Notes: The figure compares the fit of several synthetic controls calculated using different donor pools. The donor pools are trimmed in 5% increments dropping states that most closely resemble the trajectory of Coahuila. This approach attempts to eliminate states with potential spillovers. Panel (a) shows the observed data and the synthetic controls constructed. Panel (b) shows the difference between these series – the effect under different donor pools. Panel (c) shows the MSPE ratios of each of these synthetic controls against the base model. All values are positive, which suggest a worse fit compared to the base model.

Figure 3.A.6: Synthetic Control Performance When Restricting Donor Pool



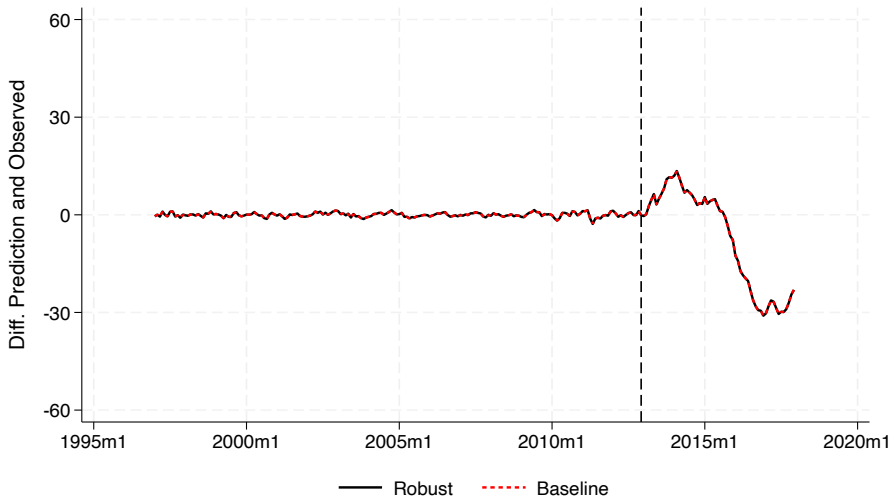
Notes: Effect on other crime-types apart from 'intentional homicides' when the synthetic controls is calculated restricting the donor pool by trimming the top 10% of the sample to avoid spillovers. The rest of the notes correspond to those from Figure 3.A.1.

Figure 3.A.7: Policy Effect on Other Crime-Types (Trimming Top 10% of Donor Pool)



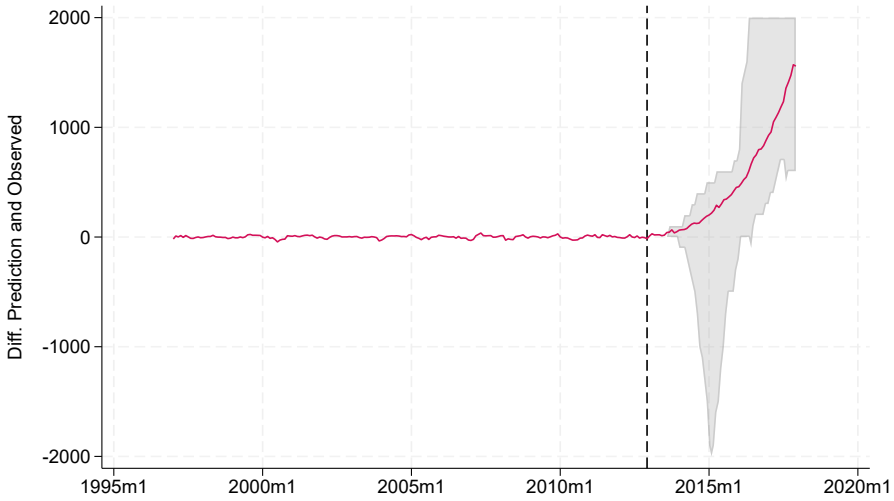
Notes: Natural logarithms of the MSPE ratios of models compared to the base model of 'intentional homicides' when the synthetic controls are built by trimming the top 10% of the sample to avoid spillovers. The rest of the notes correspond to those from Figure 3.A.3.

Figure 3.A.8: Natural Logarithm of MSPE Ratios Relative to MSPE for 'Intentional Homicides' (Trimming Top 10% of Donor Pool)



Notes: The figure shows the basic result from Figure 3.2 by excluding the states that share a border with Coahuila from the donor pool – namely, Chihuahua, Durango, Zacatecas, San Luis Potosí, and Nuevo León. Excluding from the donor pool states that Coahuila shares a border with does not alter the results.

Figure 3.A.9: Synthetic Control Performance Omitting Geographically Adjacent Donors



Notes: The figure shows the policy's effect and its respective 90% confidence intervals for crimes categorized under 'other'. Coarseness of grid defined as $1/20$ of the grid's size. Finer grids generate smoother confidence intervals.

Figure 3.A.10: Policy Effect on Crimes Categorized Under 'Other' with Confidence Intervals

3.B Additional Tables

Table 3.A.1: Weights

	(1) Base Model	(2) Placebo Date: 2010	(3) Placebo Date: 2008	(4) Placebo Date: 2006	(5) Placebo Date: 2004
Baja California Norte	0.017	-0.014		-0.028	
Baja California Sur	-0.730	-0.493	-0.384	-0.590	-0.547
Chihuahua	0.010	0.047	0.072	0.107	0.080
CDMX	0.108	0.103	0.095	0.019	0.087
Colima	0.255		-0.138	-0.351	-0.089
Durango	-0.043	-0.015	-0.051	-0.038	-0.071
Guanajuato	0.191	0.083		0.035	0.015
Guerrero	0.046	0.017	0.014	0.016	0.009
Hidalgo	-0.118		0.012		
Jalisco	0.061	0.029	0.016	0.087	
Morelos	-0.060	-0.010	-0.050		
México	0.097	-0.047	-0.036	-0.069	-0.002
Nayarit	-0.006	0.163	0.092	0.170	0.049
Nuevo León	0.070		0.052	0.156	0.109
Oaxaca	-0.028		0.006	0.008	
Puebla	0.005				
Querétaro	-0.090		0.010	0.016	
Quintana Roo	0.071		-0.008	-0.052	
San Luis Potosí	-0.123	0.092	0.105	0.146	0.072
Sinaloa	0.014	-0.075	-0.008	-0.036	-0.002
Tabasco	-0.554	-0.245	-0.048	-0.077	
Tamaulipas	0.092	0.103	0.116	0.133	0.008
Tlaxcala	-0.079	0.000	0.044	0.017	
Veracruz	0.204	0.010	0.005	-0.012	
Zacatecas	0.136	0.065	0.082	-0.015	
Aguascalientes		0.041		0.011	
Campeche		-0.018		-0.046	
Michoacán		-0.015	-0.002	0.004	0.001
Sonora		0.019	0.023		
Yucatán		0.023	0.045	0.081	0.061
Chiapas				-0.001	
Observations	192	156	132	108	84

Notes: The table reports the coefficients associated with the series of each donor unit when computing the synthetic control as chosen by a 10-fold cross-validated Lasso for the state of Coahuila. Column 1 reports the baseline specification. Columns 2-5 report the corresponding coefficients when the synthetic control is calculated on placebo treatment dates. The number of observations correspond to the number of pre-treatment observations.

Table 3.A.2: Cross-Correlation Crime-Types in Coahuila (1997 - 2017)

Variables																							
Intentional Homicides	1.000																						
Culpable Homicides	0.508	1.000																					
Culpable Lesions	0.076	0.230	1.000																				
Intentional Lesions	0.592	0.762	0.385	1.000																			
Kidnapping	0.569	0.684	0.212	0.573	1.000																		
Breach of Trust	0.438	0.510	0.623	0.742	0.502	1.000																	
Property Damage	0.582	0.705	0.621	0.805	0.699	0.858	1.000																
Dispossession	0.416	0.399	0.571	0.542	0.520	0.853	0.861	1.000															
Extortion	0.817	0.687	0.118	0.701	0.720	0.571	0.722	0.550	1.000														
Fraud	0.045	0.210	0.396	0.557	0.017	0.585	0.407	0.409	0.123	1.000													
Violent Car Theft	0.920	0.274	0.028	0.427	0.318	0.325	0.381	0.286	0.638	0.022	1.000												
Non-Violent Car Theft	0.427	0.376	0.507	0.467	0.320	0.596	0.665	0.717	0.480	0.220	0.389	1.000											
Violent Bank Robbery	0.327	0.610	0.024	0.488	0.701	0.243	0.532	0.352	0.601	0.013	0.045	0.272	1.000										
Non-Violent Car Robbery	0.423	0.481	0.233	0.503	0.455	0.533	0.628	0.595	0.705	0.211	0.277	0.598	0.532	1.000									
Violent Robbery	0.899	0.621	-0.050	0.604	0.605	0.284	0.537	0.296	0.832	-0.011	0.734	0.337	0.599	0.463	1.000								
Non-Violent Robbery	0.727	0.800	0.335	0.749	0.766	0.653	0.824	0.611	0.797	0.104	0.571	0.588	0.534	0.522	0.721	1.000							
Livestock Theft	0.647	0.628	0.351	0.691	0.600	0.777	0.853	0.809	0.818	0.311	0.416	0.608	0.510	0.707	0.645	0.732	1.000						
Rape	0.210	0.112	0.278	0.251	0.133	0.394	0.330	0.480	0.209	0.228	0.247	0.408	0.105	0.221	0.121	0.213	0.210	1.000					
Child Abuse	0.533	0.550	-0.104	0.475	0.299	0.154	0.339	0.082	0.670	0.012	0.397	0.185	0.329	0.483	0.591	0.423	0.485	-0.054	1.000				
Other Sexual	0.532	0.441	0.565	0.684	0.469	0.843	0.828	0.824	0.606	0.533	0.399	0.593	0.314	0.578	0.411	0.512	0.763	0.582	0.294	1.000			
Threats	0.596	0.386	0.391	0.626	0.480	0.662	0.728	0.693	0.611	0.471	0.451	0.463	0.422	0.536	0.521	0.407	0.680	0.471	0.334	0.921	1.000		
Other	0.141	0.151	0.463	0.359	0.193	0.536	0.529	0.605	0.219	0.438	0.012	0.373	0.231	0.344	0.090	0.066	0.463	0.503	0.085	0.819	0.838	1.000	

Notes: The table shows pairwise correlations across crime-types in Coahuila for the full time series.

Table 3.A.3: Cross-Correlation Crime-Types in Coahuila (2008 - 2012)

Variables																						
Intentional Homicides	1.000																					
Culpable Homicides	0.168	1.000																				
Culpable Lesions	-0.174	-0.624	1.000																			
Intentional Lesions	0.176	0.672	-0.687	1.000																		
Kidnapping	0.559	0.755	-0.413	0.245	1.000																	
Breach of Trust	-0.031	0.592	-0.676	0.908	0.143	1.000																
Property Damage	0.327	0.750	-0.740	0.871	0.516	0.874	1.000															
Dispossession	-0.187	0.713	-0.655	0.739	0.259	0.847	0.806	1.000														
Extortion	0.519	0.667	-0.597	0.475	0.675	0.450	0.747	0.567	1.000													
Fraud	-0.575	-0.191	0.141	0.194	-0.582	0.146	-0.206	-0.024	-0.643	1.000												
Violent Car Theft	0.880	-0.228	0.019	0.054	0.117	-0.137	0.092	-0.397	0.248	-0.351	1.000											
Non-Violent Car Theft	-0.219	-0.623	0.626	-0.433	-0.566	-0.294	-0.472	-0.306	-0.472	0.038	0.032	1.000										
Violent Bank Robbery	0.543	0.642	-0.398	0.361	0.801	0.401	0.711	0.488	0.778	-0.702	0.180	-0.248	1.000									
Non-Violent Car Robbery	-0.417	0.179	-0.225	0.044	-0.119	0.191	0.040	0.396	0.269	-0.044	-0.415	0.075	-0.016	1.000								
Violent Robbery	0.946	0.422	-0.323	0.345	0.737	0.173	0.553	0.059	0.664	-0.619	0.715	-0.330	0.750	-0.371	1.000							
Non-Violent Robbery	0.593	0.818	-0.694	0.667	0.821	0.593	0.887	0.596	0.883	-0.502	0.245	-0.586	0.855	-0.017	0.793	1.000						
Livestock Theft	0.264	0.560	-0.661	0.682	0.402	0.785	0.904	0.755	0.721	-0.352	0.076	-0.277	0.745	0.117	0.482	0.774	1.000					
Rape	0.629	0.068	-0.194	0.250	0.154	0.118	0.337	0.115	0.546	-0.380	0.646	-0.062	0.307	-0.027	0.533	0.403	0.317	1.000				
Child Abuse	-0.036	0.370	-0.626	0.715	-0.055	0.801	0.703	0.793	0.543	0.028	-0.031	-0.207	0.269	0.436	0.062	0.449	0.700	0.437	1.000			
Other Sexual	0.523	0.268	-0.433	0.777	0.114	0.679	0.726	0.372	0.383	-0.014	0.543	-0.075	0.400	-0.247	0.585	0.559	0.656	0.458	0.531	1.000		
Threats	0.956	0.310	-0.316	0.427	0.555	0.242	0.561	0.049	0.625	-0.499	0.828	-0.261	0.631	-0.359	0.960	0.726	0.476	0.659	0.190	0.718	1.000	
Other	-0.383	-0.623	0.245	0.006	-0.878	0.162	-0.166	-0.024	-0.467	0.430	0.016	0.594	-0.490	0.033	-0.494	-0.545	-0.004	0.038	0.294	0.239	-0.294	1.000

Notes: The table shows pairwise correlations across crime-types in Coahuila during the pre-treatment period 2008-2012.

3.C Code

The programme synth_ci2 calculates confidence intervals for a specific synthetic control over the post-treatment period in the dataset. The programme requires the data to be set as a time series and for the time vairable to be called *period*. The programme requires the following inputs:

- *outcome*: outcome on which the synthetic control will be calculated
- *unit*: ID number of the treated unit
- *date*: date the post-period starts
- *trim*: an indicator for whether the donor pool should be trimmed by the percentile of the correlation of their values of the outcome against the treated unit
- *lowlim*: bottom-percentile of the donor pool that should be dropped
- *uplim*: top-percentile of the donor pool that should be dropped
- *min*: grid's lower-limit
- *delta*: determines how fine the grid is
- *max*: grid's upper-limit
- *sign*: level of statistical significance

```
cap program drop synth_ci2
program synth_ci2

// Args: outcome, unit, date, trim, lowlim, uplim, min, delta, max, sign

args outcome unit date winsorize l_lim u_lim min delta max sign

// Setup of dates
quietly summ period
local tend = r(max) // Period ends
local tbegin = r(min) // Period starts
local T01 = `tend' - `tbegin' // Number of time periods
local T0 = `date' - `tbegin' // Number of policy period
local T1 = `T01' - `T0' // Number of periods post int

// Create outcome and control vars
global outcomeunit "`outcome'`unit'"

if `winsorize' == 1{
    quietly corr `outcome'* if period <= `date' // Correlation before policy date
    matrix cors = r(C)
    matrix c = cors[1...,`unit'] // column is affected unit
    mata: st_matrix("abs_c",abs(st_matrix("c"))) // Take abs value of corrs

    local row_names: rownames c
    matrix rownames abs_c = `row_names' // Assign vars to rownames

    cap frame drop var_mat
```

```
    frame create var_mat
    frame change var_mat

    svmat abs_c
    gen obs = _n
    tostring obs, replace
    gen name = "‘outcome’" + obs
    drop if abs_c1 == 1 // Drop correlation with itself
    xtile percentiles = abs_c1, n(20) // Split into percentiles
    drop if percentiles <= ‘l_lim’ | percentiles >= ‘u_lim’ // Drop using limits
    mkmat abs_c1, matrix(selected_donors) rownames(name) // Create selected_donors matrix

    frame change default
    local donor_list : rownames selected_donors // Local in "default"

}

if ‘winsorize’ == 0{
    local ents "1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32"
    local donor_ents : list ents - unit // Drop affected unit from donor pool
    local donor_list "" // Local in "default"
    foreach x of local donor_ents{
        local to_include "‘outcome’‘x’"
        local donor_list : list donor_list | to_include
    }

}

global donors = "‘donor_list’"

// Create results frame and link to default
cap frame drop results
cap drop results
frame put period, into(results)
frlink 1:1 period, frame(results)
if ‘winsorize’ == 1{
    quietly frame results: gen ci_lb_$outcomeunit = .
    quietly frame results: gen ci_ub_$outcomeunit = .
}
if ‘winsorize’ == 0{
    quietly frame results: gen ci_lb_nt_$outcomeunit = .
    quietly frame results: gen ci_ub_nt_$outcomeunit = .
}

// Outcome vector
cap matrix drop Y1
mkmat $outcomeunit, matrix(Y1)

// Controls matrix
cap matrix drop Y0
mkmat ‘donor_list’, matrix(Y0)

* ---------
disp "Finished setting everything up. Will start loop now!"
// Want to run loop over all post intervention date outcome obs
forvalues post = 1(1)‘T1’{
    disp "Begin post-intervention period: ‘post’ / ‘T1’."
    // This goes inside the loop
    local t = ‘T0’ + ‘post’ // t is counter of post-intervention periods.
    global t = ‘t’ + ‘tbegin’
    // Matrix of outcome at all pre-int period and t post-int period
    cap matrix drop Y1ci
    cap matrix drop add_outcome
    matrix Y1ci = Y1[1..‘T0’,.]
    matrix add_outcome = Y1[‘t’,1]
    matrix Y1ci = (Y1ci \ add_outcome)

    // Matrix of covars at all pre-int period and post-int period t
    cap matrix drop Y0ci
    cap matrix drop add_vars
    matrix Y0ci = Y0[1..‘T0’,1...] // Extract matrix (1:T0,all)
```

```
matrix add_vars = Y0[‘t’,1...] // Extract row t
matrix Y0ci = (Y0ci \ add_vars) // Append extra row to matrix

// These will be the inputs into the elastic net.
// These inputs change with every iteration of "post"
* ----------------------

quietly numlist "‘min’(‘delta’)‘max’", ascending
local grid r(numlist) // Local var called grid with values
local gridlength : word count ‘r(numlist)’ // Length of grid

// Create empty matrix for grid values and p-values
cap matrix drop pvector
matrix pvector = J(‘gridlength’,2,.)

disp "About to jump into the grid!"

local grid_id = 1 // Counter for the p-value vector
foreach x of numlist ‘min’(‘delta’)‘max’ {
    disp "Start grid loop: ‘grid_id’. Post intervention period ‘post’ / ‘T1’."
    // New frame "model" for calculations
    cap frame drop model
    frame create model
    frame change model
    disp "Stage 1/2. Grid loop: ‘grid_id’ / ‘gridlength’."

    // Add values of matrices as variables
    quietly svmat Y1ci, n(col)
    quietly svmat Y0ci, n(col)
    quietly gen obsid = _n // Observation ID

    // Perturb outcome at post int obs by value x of grid
    disp "Perturb outcome"
    replace $outcomeunit = $outcomeunit - ‘x’ if obsid == ‘T0’ + 1

    // Calculate Elastic Net
    disp "Calculate Elastic Net"
    quietly cvlasso $outcomeunit $donors, alpha(1.0) seed(123) postresults lopt

    // Predict outcome
    cap drop predicted
    quietly predict predicted, xb

    // Get residuals
    cap drop uhat
    quietly gen uhat = $outcomeunit - predicted

    // Absolute values of residuals and
    // indicator for extreme values
    quietly gen abs_uhat = abs(uhat)
    quietly gen indicator = (abs_uhat >= abs_uhat[‘T0’ + 1])

    // Assign proportion of extreme observations to element of second col. of
    // matrix of p-values, and grid value to first col.
    matrix pvector[‘grid_id’,1] = ‘x’

    quietly summ indicator
    matrix pvector[‘grid_id’,2] = r(mean)

    disp "The p-value for gridvalue ‘x’ is ‘r(mean)’."

    frame change default

    disp "End grid loop: ‘grid_id’ / ‘gridlength’. Post intervention period ‘post’ / ‘T1’."
    local grid_id = ‘grid_id’ + 1 // Loop ends here

}

// Create new frame "model2" and drop "model"
cap frame drop model2
frame create model2
```

```
    frame change model2
    cap frame drop model
    disp "Start stage 2/2."

    quietly svmat pvector

    // Create new var that takes values of grid for which
    // p-value is greater than significance level
    quietly gen cival = pvector1 if pvector2 > `sign'

    // Find lower and higher bound for CI and assign
    // to variables in results frame
    summ cival
    global min = r(min)
    global max = r(max)
    disp "Cool. Min of CI is `r(min)'"
    disp "Cool. Max of CI is `r(max)'"
    cap frame results: replace ci_lb_nt_$outcomeunit = $min if period == $t
    cap frame results: replace ci_ub_nt_$outcomeunit = $max if period == $t

    cap frame results: replace ci_lb_$outcomeunit = $min if period == $t
    cap frame results: replace ci_ub_$outcomeunit = $max if period == $t


    frame change default
    cap frame drop model2

    disp "End post-intervention period: `post' / `T1'."

}

cap frget ci_lb_nt_$outcomeunit, from(results)
cap frget ci_ub_nt_$outcomeunit, from(results)

cap frget ci_lb_$outcomeunit, from(results)
cap frget ci_ub_$outcomeunit, from(results)

end
```

An example of how the program can be used:

```
// Set trimming thresholds
global trim_top 21 // No trimming
global trim_low 0 // No trimming low end

synth_ci2 ma_homicidios_dol 8 m(2012m12) 0 $trim_low $trim_top -40 1 20 0.10
```

Concluding Remarks

This thesis is comprised of three independent articles in the field of political economy. To conclude the thesis, I summarise the key findings and contributions of each of these papers, and indicate limitations and productive areas for future research.

The first article explores the role of direct political communication in consolidating the regime of Hugo Chávez in Venezuela. I study the introduction of *Aló Presidente*, which began as a weekly unscripted radio programme. I rely on two key sources of variation to study its influence: variation in the coverage of the radio station that carried the programme, and variation over time in the content broadcast through this station – including the introduction of *Aló Presidente*. I identify a robust increase in Chávez's vote share attributable to the influence of *Aló Presidente*. I find evidence suggesting that the possible mechanisms underpinning this effect could be that Chavez created a narrative around democracy, and successfully used a personalised style of communication. I am able to rule out other potential explanations, such as the dis-intermediation of political parties, or the by-passing of traditional means of communication.

The paper contributes to the literature at the intersection of mass media and populism (see DellaVigna & Gentzkow, 2010; Guriev & Papaioannou, 2022; Zhuravskaya et al., 2020). It innovates by focusing on the introduction of a direct means of communication between a politician and their electorate, exploring its electoral returns. It also contributes to the literature on Chávez's regime and its legacy (e.g., Knight & Tribin, 2019, 2022). It is, to my knowledge, the first paper in economics to study the role played by political communication during the transition to an autocratic regime following Chávez's original electoral victory. Broadly, the paper's findings also connect the work of Dornbusch and Edwards (1991) and Edwards (2010) on Latin American populism with Funke et al. (2023), pointing at institutional deterioration as a potential mechanism behind the poor economic performance following the election of populist candidates. The tension between direct and constitutional democracy also connects these findings at a broader level with the literature on democratic backsliding (Levitsky & Ziblatt, 2018).

The key limitation of this paper is that reception of RNV after the introduction of *Aló Presidente* entails a large 'treatment-package' with several dimensions. These dimensions include, for instance, exposure to Chávez[20] or exposure to the narrative created by Chávez. Even exposure to Chávez has several dimensions: being exposed to the president, the fact

[20] My colleague Alena Wabitsch studies this question in her job-market paper, *The Messenger Matters*, in the context of macroeconomic policy communication. In my mind, this is a key policy-relevant issue, especially as deep-fake videos become more prevalent.

that he was a military man, his charisma, or his rhetoric.[21] The message itself could be separated into its constituent parts: what Chávez said or the way Chávez spoke.[22] Isolating these margins in experimental settings and exploring when, why, how they work, and who they work on might be a productive area for future research. There is also scope to apply more natural language processing techniques to our analyses and test frameworks of strategic (political) communication (see e.g., Di Tella, Kotti, Le Pennec, & Pons, 2025).

The second paper studies the potential role of colonial enslavement on influencing contemporary political preferences, and its underlying persistence mechanisms. I focus on the legacy of African enslavement in Venezuela. Venezuela's unique history allows us to empirically identify locations where enslaved labour was utilized during the colonial period, and individuals whose ancestors were enslaved. To identify the former, I digitise new historical sources that locate the cocoa *haciendas* where enslaved people were forced to work during the XVIIIth century. To identify enslaved ancestry, I rely on the colonial convention of registering enslaved people with their ethnic group or place of origin as surname, and the fact that these surnames have persisted through time.

I test the effect of both of these variables on contemporary outcomes, including measures of local development and votes for Hugo Chávez. I find that locations where Venezuelans have more enslaved ancestors are economically worse-off and more likely to have voted for Chávez. Correlationally, survey respondents from these locations also expressed distrust in politicians and political parties, were disengaged from politics, and exhibited zero-sum thinking. Historical census data suggests these locations persistently had lower human capital and specialised in low-skill occupations. In contrast, locations exposed to the practice of enslavement during the colonial period are not economically worse-off today, and not more likely to have voted for Chávez. Aligned with historical accounts, I find evidence suggesting that enslaved individuals migrated from cocoa *haciendas* following the abolition of slavery during the XIXth century.

This paper makes three contributions. First, it explores historical factors that contributed to the rise of *Chavismo*, pointing to human capital and occupational specialisation as their mechanisms of persistence. This adds a political dimension to the work of Althoff and Reichardt (2024) or Nunn (2008), and explores complementary mechanisms to the recent work of Becker and Voth (2023). Second, it contributes to the literature on historical persistence (e.g., Acemoglu et al., 2012, 2001; Dell, 2010; Michalopoulos & Papaioannou,

[21] Thanks to Noam Yuchtman for pointing this out.

[22] To paraphrase a recent conversation I had with Gabriel León-Ablán and Santiago Oliveros in the context of *Aló Presidente*: maybe it was but 'a tale told by an idiot, full of sound and fury, signifying nothing', and the reason individuals supported certain ideas was simply because Chávez was behind them. It might have not been about the message, but maybe in *how* the message was conveyed i.e., *the sound and fury* – the way Chávez spoke.

2016; Voigtländer & Voth, 2012) by testing mobile and place-based factors of persistence in the same context, which yields a new and unique set of findings. Third, from a historical perspective, the paper emphasises the long shadow of enslaved labour even in areas considered to be at the periphery of a colonial empire, despite that area being endowed with comparatively more liberal institutional arrangements (Mahoney, 2010). The setting contrasts in this regard with the work of Dell (2010), Acemoglu et al. (2012), or Rocha et al. (2017), and is similar to Valencia Caicedo (2019).

As detailed in the paper's conclusions, the current limitations of this project are mainly related to data availability. There are seven Venezuelan census waves that have not been digitised yet, and other lithographies similar to Codazzi (1840) that can capture exposure to armed conflict during the XIXth century, like the War of Independence. As discussed in the robustness section, while the paper attempts to establish a causal relationship between enslaved ancestry and exposure to the practice of enslavement on all outcomes via instrumental variables, there could still be space to apply other identification strategies.

The current set of findings also raises several questions. These include for instance separating place-based determinants (e.g., agricultural suitability) from mobile or place-based mechanisms.[23] Another question is quantifying whether the effects I capture are due to poverty and the persistence of low human capital on their own, or whether there is an active racial component, as suggested by other disciplines, that could be separately identified.[24] Exploring political outcomes beyond Chávez's first election could also be interesting, since he adopted race as an important part of his rhetoric post-2000.

The third paper explores the economic dividends of peace. It is an evaluation of an original set of policies adopted by the state government of Coahuila in Mexico in 2013 with the objective of combating drug cartels. The comprehensive policy-set centred on two crucial dimensions: strengthening local police enforcement, and banning businesses that provided liquidity to cartels, such as brothels, junk-yards or casinos. Coahuila's government claims the policies were highly successful, but a formal evaluation had not been carried out yet.

I rely on the synthetic controls methodology to construct valid counterfactuals and evaluate the policies (Abadie, 2021; Abadie & Gardeazabal, 2003). I check for their effectiveness against all crime-types reported by the Mexican government, and extend the analysis to high-frequency economic indicators. For inference, I adopt two complementary strategies

[23] This leads to avenues of research like disentangling whether current underdevelopment is due to the low human capital of the people who migrated into specific locations, the geographic conditions of these locations that kept them available at the time of the migration shock, or a combination of both. I want to thank Paola Conconi and Sanjay Jain for inspiring this question.

[24] Thanks to Elias Papaioannou, Andreas Fuchs, and Gianni De Fraja for suggesting this line of inquiry.

that yield similar results: I apply the recent advances in Chernozhukov et al. (2021) to compute confidence intervals for the calculated effects in all post-treatment periods, and a synthetic differences-in-differences approach (Arkhangelsky et al., 2021). I find the set of policies moderately reduced some crime-types associated with the presence of cartels, but not all. I see no impact on crime-types uncorrelated with cartels. I am unable to identify transformative economic dividends, beyond mechanical artefacts parsimoniously explained by the policies themselves, such as a reduction in informality.

This paper contributes to the literature by examining the economic dividends of peace. While the causes and consequences of conflict have been well documented (see e.g., Ray and Esteban (2017) and Rohner and Thoenig (2021), respectively), the economic returns of pacification have remained relatively understudied. The paper is also an empirical contribution that evaluates the effectiveness of a unique set of policies. The success of the approach adopted by this state could inspire similar strategies with the objective of fighting organised crime and reducing violence.

The key limitations of this paper also relate to data availability. One crucial dimension is implementation fidelity and the difficulty of quantifying it in this context. This would point to the mechanisms underpinning the results captured in the analyses. Clearly identifying these mechanisms would be a crucial component for productive research in this literature, to improve the generalizability of the findings, and to confidently inspire the adoption of new policies. More work needs to be done to explore pacification processes and how they can lead to better economic performance.[25] Another component that deserves more attention in these processes is the role of politicians' buy-in, commitment, and credibility.

[25] One exception is the work of Besley and Mueller (2012).

List of References

Abadie, A. (2021). Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*, *59*(2), 391–425. doi: 10.1257/jel.20191450

Abadie, A., Athey, S., Imbens, G., & Wooldridge, J. (2022). When Should You Adjust Standard Errors for Clustering? *Working Paper*, 1–71. Retrieved from http://arxiv.org/abs/1710.02926

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, *105*(490), 493. Retrieved from http://search.proquest.com/docview/741578133/

Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*, *59*(2), 495–510.

Abadie, A., & Gardeazabal, J. (2003). The Economic Costs of Conflict: A Case Study of the Basque Country. *The American Economic Review*, *93*(1), 113–132.

Abebe, G., Caria, A. S., Fafchamps, M., Falco, P., Franklin, S., & Quinn, S. (2021, 5). Anonymity or Distance? Job Search and Labour Market Exclusion in a Growing African City. *The Review of Economic Studies*, *88*(3), 1279–1310. Retrieved from https://doi.org/10.1093/restud/rdaa057 doi: 10.1093/restud/rdaa057

Acemoglu, D., García-Jimeno, C., & Robinson, J. A. (2012). Finding Eldorado: Slavery and long-run development in Colombia. *Journal of Comparative Economics*, *40*(4), 534–564.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *The American Economic Review*, *91*(5), 1369–1401.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *The Quarterly Journal of Economics*, *117*(4), 1231–1294.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2005). Chapter 6 Institutions as a Fundamental Cause of Long-Run Growth. In *Handbook of economic growth* (Vol. 1, pp. 385–472). Elsevier B.V.

Acemoglu, D., & Robinson, J. A. (2012). *Why nations fail: the origins of power, prosperity and poverty*. New York: Crown Publishers.

Acharya, A., Blackwell, M., & Sen, M. (2015). A Culture of Disenfranchisement: How American Slavery Continues to Affect Voting Behavior. *Working Paper*.

Acharya, A., Blackwell, M., & Sen, M. (2016a). Explaining Causal Findings Without Bias: Detecting and Assessing Direct Effects. *The American Political Science Review*, *110*(3), 512–529.

Acharya, A., Blackwell, M., & Sen, M. (2016b). The Political Legacy of American Slavery. *The Journal of Politics*, *78*(3), 621–641.

Acosta Saignes, M. (1984). *Vida de los esclavos negros en Venezuela* (3rd ed.). Valencia: Vadell Hermanos.

Adena, M., Enikolopov, R., Petrova, M., Santarosa, V., & Zhuravskaya, E. (2015, 5). Radio and the Rise of the Nazis in Prewar Germany. *The Quarterly Journal of Economics*, *130*(4), 1885–1940. Retrieved from https://www-jstor-org.ezproxy-prd.bodleian.ox.ac.uk/stable/26372641

Agarwal, M., Bahure, V., Bergonzoli, K., & Maji, S. (2024). Education and Domestic Violence: Evidence from a School Construction Program in India. *Working Paper*. Retrieved from https://www.sitesideas.org/category/working-papersâĂćtheREPECwebsite:https://ideas.repec.org/s/awm/wpaper.html

Aguayo, S., & Dayán, J. (2020). *"Reconquistando" la laguna: Los zetas, el estado y la sociedad organizada, 2007-2014*. Mexico City: El Colegio de México.

Aguilar Valenzuela, R., & Moreira Valdez, R. (2022). *Jaque Mate al Crimen Organizado. Coahuila: Una estrategia multidimensional para la paz*. Mexico City: Editorial Planeta.

Ahrens, A., Hansen, C. B., & Schaffer, M. E. (2020). lassopack: Model selection and prediction with regularized regression in Stata. *The Stata Journal*, *20*(1), 176–235. doi: 10.1177/1536867X20909697

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters*, *80*(1), 123–129.

Althoff, L., & Reichardt, H. (2024, 11). Jim Crow and Black Economic Progress after Slavery. *The Quarterly Journal of Economics*, *139*(4), 2279–2330. Retrieved from https://doi.org/10.1093/qje/qjae023 doi: 10.1093/qje/qjae023

Anderson, M. L. (2008). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, *103*(484), 1481–1495.

Angrist, J. D., & Krueger, A. B. (1999). Empirical Strategies in Labor Economics. In *Handbook of labor economics* (Vol. 3, pp. 1277–1366). Elsevier B.V.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: an empiricist's companion*. Princeton, New Jersey ; Woodstock, Oxfordshire.

Angrist, N., Djankov, S., Goldberg, P. K., & Patrinos, H. A. (2021, 4). Measuring human capital using global learning data. *Nature*, *592*(7854), 403–408.

Appel, B. J., & Loyle, C. E. (2012). The economic benefits of justice: Post-conflict justice and foreign direct investment. *Journal of Peace Research*, *49*(5), 685–699. doi: 10.1177/0022343312450044

Arcila Farías, E. (1966). *El régimen de la encomienda en Venezuela* (2nd ed.). Caracas: Universidad Central de Venezuela, Facultad de Economía, Instituto de Investigaciones.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic Difference-in-Differences. *The American Economic Review*, *111*(12), 4088–4118. doi: 10.1257/aer.20190159

Assouad, L. (2020). Charismatic Leaders and Nation Building. *IDEAS Working Paper Series from RePEc*.

Athey, S., & Imbens, G. (2017). The Econometrics of Randomized Experiments. In A. V. Banerjee & E. Duflo (Eds.), *Handbook of economic field experiments* (chap. 3). Amsterdam: North-Holland.

Athey, S., & Imbens, G. W. (2006). Identification and Inference in Nonlinear Difference-in-Differences Models. *Econometrica*, *74*(2), 431–497.

Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized Random Forests. *The Annals of Statistics*, *47*(2), 1148–1178.

Baron, R. M., & Kenny, D. A. (1986). The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182.

Barzel, Y., & Silberberg, E. (1973). Is the act of voting rational? *Public Choice*, *16*(1), 51–58. doi: 10.1007/BF01718806

Becker, S. O., & Voth, H.-J. (2023). From the "death of God" to the rise of Hitler. *Working Paper*, *18543*. Retrieved from http://www.jvoth.com/shallow.pdf

Bellodi, L., Morelli, M., & Vannoni, M. (2024). A Costly Commitment: Populism, Economic Performance, and the Quality of Bureaucracy. *American Journal of Political Science*, *68*(1), 193–209. doi: 10.1111/ajps.12782

Benjamini, Y., Krieger, A. M., & Yekutieli, D. (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, *93*(3), 491–507.

Berlanda, A., Cervellati, M., Esposito, E., Rohner, D., & Sunde, U. (2024). Medication Against Conflict. *Journal of Development Economics*, *171*. doi: 10.1016/j.jdeveco.2024.103306

Bernini, A. (2023). The voice of radio in the battle for equal rights: Evidence from the U.S. South. *Economics & Politics*, *35*(1), 163–226. Retrieved from https://onlinelibrary

.wiley.com/doi/abs/10.1111/ecpo.12213

Besley, T., & Mueller, H. (2012). Estimating the Peace Dividend: The Impact of Violence on House Prices in Northern Ireland. *The American Economic Review*, *102*(2), 810–833. doi: 10.1257/aer.102.2.810

Besley, T., & Persson, T. (2008). Wars and State Capacity. *Journal of the European Economic Association*, *6*(2-3), 522–530.

Besley, T., & Persson, T. (2010). State Capacity, Conflict, and Development. *Econometrica*, *78*(1), 1–34.

Besley, T., & Persson, T. (2011). Fragile States and Development Policy. *Journal of the European Economic Association*, *9*(3), 371–398.

Björkegren, D. (2019). The adoption of network goods: Evidence from the spread of mobile phones in Rwanda. *The Review of Economic Studies*, *86*(3), 1033–1060.

Blattman, C., & Annan, J. (2016). Can Employment Reduce Lawlessness and Rebellion? A Field Experiment with High-Risk Men in a Fragile State. *American Political Science Review*, *110*(1), 1–17. doi: 10.1017/S0003055415000520

Blattman, C., Hartman, A. C., & Blair, R. A. (2014). How to promote order and property rights under weak rule of Law? An experiment in changing dispute resolution behavior through Community Education. *American Political Science Review*, *108*(1), 100–120. doi: 10.1017/S0003055413000543

Blattman, C., & Miguel, E. (2010). Civil War. *Journal of Economic Literature*, *48*(1), 3–57.

Blouin, A., & Mukand, S. W. (2019). Erasing ethnicity? Propaganda, nation building, and identity in rwanda. *Journal of Political Economy*, *127*(3), 1008–1062. Retrieved from http://search.proquest.com/docview/2277026453/ doi: 10.1086/701441

Borusyak, K., & Hull, P. (2021). Non-Random Exposure to Exogenous Shocks: Theory and Applications. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3696212

Borusyak, K., & Hull, P. (2023). Nonrandom Exposure to Exogenous Shocks. *Econometrica*, *91*(6), 2155–2185.

Brewer Carías, A.-R. (2015). *La mentira como política de estado : crónica de una crisis política permanente : Venezuela, 1999-2015*. Caracas, Venezuela.

Brito Figueroa, F. (1966). *Historia economica y social de Venezuela : una estructura para su estudio*. Caracas.

Brito Figueroa, F. (1973). *El problema tierra y esclavos en la historia de Venezuela*. Caracas: Teoría y Praxis.

Bursztyn, L., Cantoni, D., Funk, P., Schönenberger, F., Schönenberger, S., & Yuchtman, N. (2024). Identifying the Effect of Election Closeness on Voter Turnout: Evidence from Swiss Referenda. *Journal of the European Economic Association*, *22*(3), 876–914.

Caesmann, M., Caprettini, B., Voth, H.-J., & Yanagizawa-Drott, D. (2021). Going Viral:

Propaganda, Persuasion and Polarization in 1932 Hamburg. *IDEAS Working Paper Series from RePEc*.

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2019). Regression Discontinuity Designs Using Covariates. *The Review of Economics and Statistics*, *101*(3), 442–451.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust data-driven inference in the regression-discontinuity design. *The Stata Journal*, *14*(4), 909–946.

Cannon, B. (2013). *Hugo Chávez and the Bolivarian Revolution: Populism and democracy in a globalised age*. Manchester: Manchester University Press.

Carrasquero, J. V., & Welsch, F. (2001). Revolución en democracia o retorno al caudillismo. In J. V. Carrasquero, T. Maingón, & F. Welsch (Eds.), *Venezuela en transición: elecciones y democracia 1998-2000* (1st ed.). Red Universitaria de Estudios Políticos de Venezuela-RedPol: CDB Publicaciones.

Carvalho, C., Masini, R., & Medeiros, M. C. (2018). ArCo: An artificial counterfactual approach for high-dimensional panel time-series data. *Journal of Econometrics*, *207*(2), 352–380.

Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2019). *A Practical Introduction to Regression Discontinuity Designs*. Cambridge Elements: Quantitative and Computational Methods for Social Science. doi: 10.1017/9781108684606

Cattaneo, M. D., & Titiunik, R. (2022). Regression Discontinuity Designs. *Annual Review of Economics*, *14*(1), 821–851.

CEPT. (2007). *Final Acts of the CEPT T-DAB Planning Meeting* (Tech. Rep.). Constanţa: CEPT.

Cervellati, M., Sunde, U., & Valmori, S. (2017). Pathogens, Weather Shocks and Civil Conflcits. *The Economic Journal*, *127*(607), 2581–2616. doi: 10.1111/ecoj.12430

Chen, Y., & Yang, D. Y. (2019). The Impact of Media Censorship: 1984 or Brave New World? *The American Economic Review*, *109*(6), 2294–2332.

Chernozhukov, V., Hansen, C., & Liao, Y. (2017). A Lava Attack on the Recovery of Sums of Dense and Sparse Signals. *The Annals of Statistics*, *45*(1), 39–76.

Chernozhukov, V., Wüthrich, K., & Zhu, Y. (2021). An Exact and Robust Conformal Inference Method for Counterfactual and Synthetic Controls. *Journal of the American Statistical Association*, *116*(536), 1849–1864.

Chesnut, R. A. (2012, 1). *Devoted to Death: Santa Muerte, the Skeleton Saint.* Oxford University Press. Retrieved from https://doi.org/10.1093/acprof:oso/9780199764662.001.0001 doi: 10.1093/acprof:oso/9780199764662.001.0001

Chi, G., Fang, H., Chatterjee, S., & Blumenstock, J. E. (2022). Microestimates of wealth for all low- and middle-income countries. *Proceedings of the National Academy of Sciences*, *119*(3). doi: 10.1073/pnas.2113658119

Chiang, C.-F., & Knight, B. (2011). Media bias and influence: Evidence from newspaper

endorsements. *Review of Economic Studies*, *78*(3), 795–820. doi: 10.1093/restud/rdq037

CIESIN. (2011). *Foresight Project on Migration and Global Environmental Change, Report MR4: Estimating Net Migration by Ecosystem and by Decade,1970-2010* (Tech. Rep.). London: CIESIN.

CIESIN. (2017). *Global Population Density Grid Time Series Estimates.* Palisades, New York: NASA Socioeconomic Data and Applications Center (SEDAC). Retrieved from https://doi.org/10.7927/H47M05W2

Clarke, D., Pailañir, D., Athey, S., & Imbens, G. (2024). On synthetic difference-in-differences and related estimation methods in Stata. *The Stata Journal*, *24*(4), 557–598. doi: 10.1177/1536867X241297914

Codazzi, A. (1840). Mapa fisico de Venezuela divida en hoyas hydrograficas. Mapa fisico de Venezuela dividida en tres zonas. In *Atlas fisico y politico de la republica de venezuela dedicado por su autor, el coronel de ingenieros agustin codazzi al congreso constituyente de 1830.* Caracas. Retrieved from https://www.davidrumsey.com/rumsey/download.pl?image=/D0059/5226005.sid

Codazzi, A. (1960). *Obras escogidas*. Caracas: Dirección de Cultura y Bellas Artes, Departmento de Publicaciones.

Codazzi, A., Baralt, R. M., & Díaz, R. (1940). *Resumen de la geografía de Venezuela (Venezuela en 1841)*. Caracas: Taller de artes gráficas, Escuela técnica industrial.

Collier, P., & Hoeffler, A. (1998). On economic causes of civil war. *Oxford Economic Papers*, *50*(4), 563–573.

Collier, P., & Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford Economic Papers*, *56*(4), 563–595.

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, *92*(1), 1–45.

Conley, T. G., & Kelly, M. (2025). The standard errors of persistence. *Journal of International Economics*, *153*, 104027. doi: 10.1016/j.jinteco.2024.104027

Couttenier, M., Petrencu, V., Rohner, D., & Thoenig, M. (2019). The Violent Legacy of Conflict: Evidence on Asylum Seekers, Crime, and Public Policy in Switzerland. *The American Economic Review*, *109*(12), 4378–4425. doi: 10.1257/aer.20170263

Cox, G. W. (1988). Closeness and Turnout: a Methodological Note. *The Journal of Politics*, *50*(3), 768–775.

Cox, G. W., & Katz, J. N. (1996). Why Did the Incumbency Advantage in U.S. House Elections Grow? *American Journal of Political Science*, *40*(2), 478–497.

Cox, G. W., & Munger, M. C. (1989). Closeness, Expenditures, and Turnout in the 1982 U. S. House Elections. *The American Political Science Review*, *83*(1), 217–231.

Dal Bo, E., & Dal Bo, P. (2011). Workers, Warriors, and Criminals: Social Conflict in General Equilibrium. *Journal of the European Economic Association*, *9*(4), 646–677.

Davies, S., Pettersson, T., Sollenberg, M., & Öberg, M. (2025). Organized violence 1989–2024, and the challenges of identifying civilian victims. *Journal of Peace Research*, *62*(4).

Davis, J. M. V., & Heller, S. B. (2017). Using Causal Forests to Predict Treatment Heterogeneity: An Application to Summer Jobs. *The American Economic Review*, *107*(5), 546–550.

De Anda Casas, A. (2025a). Aló Presidente: Direct Political Communication and Electoral Outcomes in Chávez's Venezuela. *Working Paper*.

De Anda Casas, A. (2025b). Poverty and Low Human Capital: The legacy of African slavery in Venezuela and the deep historical roots of Chavismo. *Working Paper*.

Delignette-Muller, M. L., & Dutang, C. (2015). fitdistrplus: An R package for fitting distributions. *Journal of Statistical Software*, *64*(4), 1–34.

Dell, M. (2010). The Persistent Effects of Peru's Mining Mita. *Econometrica*, *78*(6), 1863–1903.

Dell, M., Lane, N., & Querubin, P. (2018). The Historical State, Local Collective Action, and Economic Development in Vietnam. *Econometrica*, *86*(6), 2083–2121.

DellaVigna, S., & Gentzkow, M. (2010). Persuasion : Empirical Evidence. *Annual Review of Economics*, *2*, 643–669.

DellaVigna, S., & Kaplan, E. (2007). The Fox News Effect : Media Bias and Voting. *The Quarterly Journal of Economics*, *122*(3), 1187–1234.

de Santos Cabrera, J. (n.d.). *Mapa Topographico de la Provincia de Venezuela alías Caracas, con parte del nuevo reino de Granada probincia de Maracaibo.* [17–?]. Retrieved from https://bibliotecavirtual.defensa.gob.es/BVMDefensa/es/consulta/registro.do?id=115573

Desposato, S. W., & Petrocik, J. R. (2003). The Variable Incumbency Advantage: New Voters, Redistricting, and the Personal Vote. *American Journal of Political Science*, *47*(1), 18–32.

Di Tella, R., Kotti, R., Le Pennec, C., & Pons, V. (2025). Keep Your Enemies Closer: Strategic Platform Adjustments During U.S. And French Elections. *The American Economic Review*, *115*(8), 2488–2528. doi: 10.2139/ssrn.4519265

Dornbusch, R., & Edwards, S. (1991). *The Macroeconomics of Populism in Latin America*. Chicago: University of Chicago.

Downs, A. (1957). *An Economic Theory of Democracy*. New York: Harper & Row.

Dube, O., & Vargas, J. F. (2013). Commodity Price Shocks and Civil Conflict: Evidence from Colombia. *The Review of Economic Studies*, *80*, 1384–1421.

EBU. (2003). *Technical Bases for T-DAB Services Network Planning and Compatibility with Existing Broadcasting Services* (Tech. Rep.). Geneva, Switzerland: EBU Technical Texts.

EBU. (2013). *Technical Bases for T-DAB Services Network Planning and Compability with Existing*

Broadcasting Services (Tech. Rep.). Geneva, Switzerland: EBU Technical Texts.

Edwards, S. (2010). *Left behind Latin America and the false promise of populism*. Chicago: University of Chicago Press.

Eichengreen, B. J. (2018). *The populist temptation: economic grievance and political reaction in the modern era*. New York.

Ellner, S. (2008). *Rethinking Venezuelan politics: class, conflict, and the Chávez phenomenon*. Boulder, Colorado: Lynne Rienner.

Eltis, D. (1999). *The trans-Atlantic slave trade a database on CD-ROM.* Cambridge: Cambridge University Press.

Elvidge, C. D., Baugh, K. E., Dietz, J. B., Bland, T., Sutton, P. C., & Kroehl, H. W. (1999). Radiance calibration of DMSP-OLS low-light imaging data of human settlements. *Remote Sensing of Environment*, *68*(1), 77–88.

Engerman, S. L., & Sokoloff, K. L. (2005a). *Colonialism, Inequality, and Long-Run Paths of Development.* Cambridge, Mass.

Engerman, S. L., & Sokoloff, K. L. (2005b). The Evolution of Suffrage Institutions in the New World. *The Journal of Economic History*, *65*(4), 891–921.

Engerman, S. L., & Sokoloff, K. L. (2012a). Colonisation and Development. *Economic History of Developing Regions*, *27*(S1), S28-S40.

Engerman, S. L., & Sokoloff, K. L. (2012b). *Economic development in the Americas since 1500 : endowments and institutions*. Cambridge ; New York: Cambridge University Press.

Enikolopov, R., Petrova, M., & Zhuravskaya, E. (2011). Media and Political Persuasion: Evidence from Russia. *The American Economic Review*, *101*(7), 3253–3285.

Erikson, R. S., & Titiunik, R. (2015). Using regression discontinuity to uncover the personal incumbency advantage. *Quarterly Journal of Political Science*, *10*(1), 101–119.

Esposito, E., Rotesi, T., Saia, A., & Thoenig, M. (2023). Reconciliation Narratives: The Birth of a Nation after the US Civil War. *The American Economic Review,*, *113*(6), 1461–1504.

Esteban, J., & Ray, D. (2008). On the Salience of Ethnic Conflict. *The American Economic Review*, *98*(5), 2185–2202.

Esteban, J., & Ray, D. (2011a). Linking Conflict to Inequality and Polarization. *The American Economic Review*, *101*(4), 1345–1374.

Esteban, J., & Ray, D. (2011b). A Model of Ethnic Conflict. *Journal of the European Economic Association*, *9*(3), 496–521.

European Commission. (2003). *Global Land Cover 2000 database.* Joint Research Centre.

Evans, J. S., & Murphy, M. A. (2021). *spatialEco.* Retrieved from https://github.com/jeffreyevans/spatialEco

Fajardo-Steinhäuser, M. (2024). Peace Dividends: The Economic Effects of Colombia's Peace Agreement. *Working Paper*. Retrieved from http://arxiv.org/abs/2301.01843

Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, Insurgency, and Civil War. *The American Political Science Review*, *97*(1), 75–90.

Fernandez, I., Peña, I., Angueira, P., Gil, U., Ordiales, J. L., & Landa, I. (2012). Building Penetration Losses and Noise Levels Characterization in Medium Wave. *IEEE Antennas and Wireless Propagation Letters*, *11*, 1056–1059.

Fetzer, T. (2020). Can Workfare Programs Moderate Conflict? Evidence from India. *Journal of the European Economic Association*, *18*(6), 3337–3375. doi: 10.1093/jeea/jvz062

Figueiras, J., & Frattasi, S. (2010). *Mobile positioning and tracking: from conventional to cooperative techniques*. Chichester: Wiley.

Fisman, R., & Miguel, E. (2007). Corruption, Norms, and Legal Enforcement: Evidence from Diplomatic Parking Tickets. *Journal of Political Economy*, *115*(6), 1020–1048.

Fouka, V., & Voth, H.-J. (2023). Collective Remembrance and Private Choice: German–Greek Conflict and Behavior in Times of Crisis. *American Political Science Review*, *117*(3), 851–870. doi: 10.1017/S0003055422001095

Frajman, E. (2014). Broadcasting populist leadership: Hugo Chávez and Aló Presidente. *Journal of Latin American Studies*, *46*(3), 501–526. doi: 10.1017/S0022216X14000716

Franklin, M. N. (2004). *Voter turnout and the dynamics of electoral competition in established democracies since 1945*. Cambridge, UK ; New York: Cambridge University Press.

Funke, M., Schularick, M., & Trebesch, C. (2023). Populist Leaders and the Economy. *The American Economic Review*, *113*(12), 3249–3288. Retrieved from http://hdl.handle.net/10419/226836

Gagliarducci, S., Onorato, M. G., Sobbrio, F., & Tabellini, G. (2020). War of the Waves: Radio and Resistance during World War II. *American Economic Journal: Applied Economics*, *12*(4), 1–38.

Garcia-Serra, M. J. (2001, 5). The "Enabling Law": The Demise of the Separation of Powers in Hugo Chavez's Venezuela. *The University of Miami Inter-American Law Review*, *32*(2), 265–293. Retrieved from http://www.jstor.org/stable/40176554

Gelb, A., Ramachandran, V., Shah, M. K., & Turner, G. (2011). What matters to African firms? The relevance of perceptions data. *Policy Research Working Paper*, 197–225.

Gelbach, J. B. (2016). When Do Covariates Matter? And Which Ones, and How Much? *Journal of Labor Economics*, *34*(2), 509–543.

Gentzkow, M. (2006). Television and Voter Turnout. *The Quarterly Journal of Economics*, *121*(3), 931–972.

Gerber, A. S., Karlan, D., & Bergan, D. (2009). Does the media matter? A field experiment measuring the effect of newspapers on voting behavior and political opinions. *American Economic Journal: Applied Economics*, *1*(2), 35–52. doi: 10.1257/app.1.2.35

Giuliano, P., & Nunn, N. (2021). Understanding Cultural Persistence and Change. *The*

Review of Economic Studies, *88*(4), 1541–1581.

Górecki, M. A. (2013). Election Closeness, Habit Formation and Voter Turnout: Evidence from Sixteen Swedish Elections. *Political Studies*, *61*(SUPPL.1), 234–248. doi: 10.1111/1467-9248.12017

Gott, R. (2005). *Hugo Chávez and the Bolivarian Revolution*. London : New York: Verso.

Grimmer, J., Hersh, E., Feinstein, B., & Carpenter, D. (2011). Are Close Elections Random? *Working Paper*.

Güell, M., Mora, J. V., & Telmer, C. I. (2013). The informational content of surnames, the evolution of intergenerational mobility, and assortative mating. *The Review of Economic Studies*, *82*(2), 693–735.

Guiso, L., Sapienza, P., & Zingales, L. (2009). Cultural Biases in Economic Exchange? *The Quarterly Journal of Economics*, *124*(3), 1095–1131. doi: 10.1162/qjec.2009.124.3.1095

Gupta, S., & Kopper, S. (2021). *Power calculations.* Retrieved from https://www.povertyactionlab.org/resource/power-calculations

Guriev, S., Melnikov, N., & Zhuravskaya, E. (2021). 3G Internet and Confidence in Government. *The Quarterly Journal of Economics*, 2533–2613. doi: 10.1093/qje/qjaa040.Advance

Guriev, S., & Papaioannou, E. (2022). The Political Economy of Populism: An Empirical Investigation. *Journal of Economic Literature*, *60*(3), 753–832. doi: 10.1057/s41294-018-0059-3

Hahn, J., Todd, P., & der Klaauw, W. (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica*, *69*(1), 201–209.

Henderson, J. V., Storeygard, A., & Weil, D. N. (2011). A Bright Idea for Measuring Economic Growth. *The American Economic Review*, *101*(3), 194–199.

Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring Economic Growth from Outer Space. *The American Economic Review*, *102*(2), 994–1028.

Herre, B. (2022). The world has recently become less democratic. *Our World in Data*. Retrieved from https://ourworldindata.org/less-democratic

Herrera Salas, J. M. (2005). Ethnicity and Revolution: The Political Economy of Racism in Venezuela. *Latin American Perspectives*, *32*(2), 72–91.

Herrmann, O., Jong-A-Pin, R., & Schoonbeek, L. (2019). A prospect-theory model of voter turnout. *Journal of Economic Behavior & Organization*, *168*, 362–373.

Hicks, R., & Tingley, D. (2011). Causal mediation analysis. *Stata Journal*, *11*(4), 605–619. doi: 10.1177/1536867x1201100407

Hodge, V. J., & Austin, J. (2003). A comparison of standard spell checking algorithms and a novel binary neural approach. *IEEE Transactions on Knowledge and Data Engineering*, *15*(5), 1073–1081. doi: 10.1109/TKDE.2003.1232265

Hoeg, W., & Lauterbach, T. (2003). *Digital audio broadcasting: principles and applications of digital radio* (2nd ed.). Chichester: Wiley.

Hollingsworth, A., & Wing, C. (2022). Tactics for Design and Inference in Synthetic Control Studies: An Applied Example Using High-Dimensional Data. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3592088

Hsiang, S. M. (2010). Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(35), 15367–15372. doi: 10.1073/pnas.1009510107

Hsieh, C.-T., Miguel, E., Ortega, D., & Rodriguez, F. (2011). The Price of Political Opposition: Evidence from Venezuela's "Maisanta". *American Economic Journal: Applied Economics*, *3*(2), 196–214.

Hsu, F. C., Baugh, K. E., Ghosh, T., Zhizhin, M., & Elvidge, C. D. (2015). DMSP-OLS radiance calibrated nighttime lights time series with intercalibration. *Remote Sensing*, *7*(2), 1855–1876.

Hunt, K., Cesky, T., Jeacock, T., Mägele, M., O'Leary, T., & Petke, G. (1995). *The CEPT T-DAB Planning Meeting Wiesbaden, July 1995* (Tech. Rep.). CEPT. Retrieved from https://tech.ebu.ch/docs/techreview/trev_267-hunt.pdf

Hyytinen, A., Meriläinen, J., Saarimaa, T., Toivanen, O., & Tukiainen, J. (2018). When does regression discontinuity design work? Evidence from random election outcomes. *Quantitative Economics*, *9*(2), 1019–1051.

Imai, K., Keele, L., & Tingley, D. (2010). A General Approach to Causal Mediation Analysis. *Psychological Methods*, *15*(4), 309–334. doi: 10.1037/a0020761

Imai, K., Keele, L., & Yamamoto, T. (2010). Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science*, *25*(1), 51–71. doi: 10.1214/10-STS321

Imbens, G. W., & Doudchenko, N. (2016). Balancing, Regression, Difference-in-Differences and Synthetic Control Methods: a synthesis. *NBER Working Paper Series*. Retrieved from http://www.nber.org/papers/w22791

IPUMS. (2020). *Minnesota Population Center. Integrated Public Use Microdata Series, International: Version 7.3 [dataset].* Minneapolis, MN.. doi: https://doi.org/10.18128/D020.V7.3

Ishibashi, J. (2003). Hacia una apertura del debate sobre el racismo en Venezuela. In D. Mato (Ed.), *Politicas de identidades y diferencias sociales en tiempos de globalización* (pp. 33–61). Caracas, Venezuela: FACES-UCV.

Jowett, G. S., & O'Donnell, V. (2015). *Propaganda & Persuasion* (6th ed.). SAGE Publications.

Jun, S. J., & Lee, S. (2023, 2). Identifying the Effect of Persuasion. *Journal of Political Economy*, *131*(8), 2032–2058. Retrieved from https://doi.org/10.1086/724114

doi: 10.1086/724114

Kamenica, E. (2019). Bayesian Persuasion and Information Design. *Annual Review of Economics*, *11*, 249–272. doi: 10.1146/annurev-economics-080218-025739

Kamenica, E., & Gentzkow, M. (2011). Bayesian Persuasion. *The American Economic Review*, *101*(6), 2590–2615.

Klašnja, M., & Titiunik, R. (2017). The Incumbency Curse: Weak Parties, Term Limits, and Unfulfilled Accountability. *The American Political Science Review*, *111*(1), 129–148.

Knight, B., & Tribin, A. (2019). The limits of propaganda: Evidence from Chavez's Venezuela. *Journal of the European Economic Association*, *17*(2), 567–605.

Knight, B., & Tribin, A. (2022). Opposition Media, State Censorship, and Political Accountability: Evidence from Chavez's Venezuela. *The World Bank Economic Review*, *36*(2), 455–487.

Kornblith, M. (2001). Confiabilidad y transparencia de las elecciones en Venezuela : examen de los comicios del 30 de julio de 2000. In J. V. Carrasquero, T. Maingón, & F. Welsch (Eds.), *Venezuela en transición: elecciones y democracia 1998-2000.* Red Universitaria de Estudios Políticos de Venezuela-RedPol: CDB Publicaciones.

Kristensen, R. A. (2015). La Santa Muerte in Mexico City: The Cult and its Ambiguities. *Journal of Latin American Studies*, *47*(3), 543–566.

Lee, D. S. (2008). Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics*, *142*(2), 675–697.

Lee, D. S., & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, *48*(2), 281–355.

León, G. (2012). Civil Conflict and Human Capital Accumulation: The Long-term Effects of Political Violence in Perú. *The Journal of Human Resources*, *47*(4), 991–1022. doi: 10.3368/jhr.47.4.991

Levitsky, S., & Ziblatt, D. (2018). *How Democracies Die: what history tells us about our future*. London: Viking.

Lewis, P., Barr, C., Clarke, S., Voce, A., Levett, C., & Gutiérrez, P. (2019, 3). *Revealed: the rise and rise of populist rhetoric.* Retrieved from https://www.theguardian.com/world/ng-interactive/2019/mar/06/revealed-the-rise-and-rise-of-populist-rhetoric

Li, K. T. (2020). Statistical Inference for Average Treatment Effects Estimated by Synthetic Control Methods. *Journal of the American Statistical Association*, *115*(532), 2068–2083.

Li, K. T., & Bell, D. R. (2017). Estimation of average treatment effects with panel data: Asymptotic theory and implementation. *Journal of Econometrics*, *197*(1), 65–75.

Linden, A. (2019). ESIZEREG: Stata module for computing the effect size based on a linear regression coefficient. *Working Paper*. Retrieved from https://ideas.repec.org/

c/boc/bocode/s458607.html

Lombardi, J. V. (1971). *The decline and abolition of Negro slavery in Venezuela, 1820-1854*. Westport, Conn: Greenwood.

Maggio, D., Karra, M., & Canning, D. (2021). Family Planning Services and Children's Human Capital: Experimental Evidence from Urban Malawi. *Working Paper*.

Mahoney, J. (2010). *Colonialism and postcolonial development: Spanish America in comparative perspective*. Cambridge: Cambridge University Press.

Mallén, A. L., & García Guadilla, M.-P. (2017). *Venezuela's polarized politics: the paradox of direct democracy under Chávez*. Boulder, Colorado.

Manski, C. F. (2007). *Identification for prediction and decision*. Cambridge, Mass.: Harvard University Press.

Marcano, C., & Barrera Tyszka, A. (2007). *Hugo Chávez*. New York: Random House.

Matsusaka, J. G. (1993). Election closeness and voter turnout: Evidence from California ballot propositions. *Public Choice*, *76*(4), 313–334. doi: 10.1007/BF01053302

Matsusaka, J. G., & Palda, F. (1993). The Downsian voter meets the ecological fallacy. *Public Choice*, *77*, 855–878.

maxentlabs. (2022). *Ocrpy: OCR, Archive, Index and Search any documents with ease.* GitHub. Retrieved from https://github.com/maxent-ai/ocrpy

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*, *99*(2), 210–221.

Mensah, J. T., Tafere, K., & Abay, K. A. (2021). Saving Lives Through Technology: Mobile Phones and Infant Mortality. *Working Paper*.

Michalopoulos, S., & Papaioannou, E. (2013). Pre-Colonial Ethnic Institutions and Contemporary African Development. *Econometrica*, *81*(1), 113–152. doi: 10.3982/ecta9613

Michalopoulos, S., & Papaioannou, E. (2014). National Institutions and Subnational Development in Africa. *The Quarterly Journal of Economics*, *129*(1), 151–214.

Michalopoulos, S., & Papaioannou, E. (2016). The Long-Run Effects of the Scramble for Africa. *The American Economic Review*, *106*(7), 1802–1848.

Miguel, E., & Roland, G. (2011). The long-run impact of bombing Vietnam. *Journal of Development Economics*, *96*(1), 1–15.

Miguel, E., Satyanath, S., & Sergenti, E. (2004). Economic Shocks and Civil Conflict: An Instrumental Variables Approach. *Journal of Political Economy*, *112*(4), 725–753.

Mitra, A., & Ray, D. (2014). Implications of an Economic Theory of Conflict: Hindu-Muslim Violence in India. *Journal of Political Economy*, *122*(4), 719–765.

Molina V, J. E. (2001). Comportamiento electoral en Venezuela 1998-2000: cambio y continuidad. In J. V. Carrasquero, T. Maingón, & F. Welsch (Eds.), *Venezuela en transición: elecciones y democracia 1998-2000* (1st ed.). Red Universitaria de Estudios Políticos de

Venezuela-RedPol: CDB Publicaciones.

Mueller, H., Piemontese, L., & Tapsoba, A. (2017). Recovery from Conflict: Lessons of Success. *Policy Research Working Paper*. doi: 10.1596/1813-9450-7970

Mueller, H., & Tobias, J. (2016). *The cost of violence: Estimating the economic impact of conflict* (Tech. Rep.). London: International Growth Centre. Retrieved from https://documents1.worldbank.org/curated/en/519741603804458786/pdf/Conflict-and-Poverty.pdf

Nunn, N. (2008). The Long-term Effects of Africa's Slave Trades. *The Quarterly Journal of Economics*, *123*(1), 139–176.

Nunn, N., & Puga, D. (2012). Ruggedness: The Blessing of Bad Geography in Africa. *The Review of Economics and Statistics*, *94*(1), 20–36.

Nunn, N., & Wantchekon, L. (2011). The Slave Trade and the Origins of Mistrust in Africa. *The American Economic Review*, *101*(7).

OCEI. (1993a). *Censo indígena de Venezuela, 1992*. Caracas: Oficina Central de Estadística e Informática, OCEI.

OCEI. (1993b). *El Censo 90 en ...* Caracas: Oficina Central de Estadística e Informática, OCEI.

OCEI. (1993c). *El Censo 90 en Venezuela*. Caracas: Oficina Central de Estadística e Informática, OCEI.

Olken, B. A. (2009). Do television and radio destroy social capital? Evidence from indonesian villages. *American Economic Journal: Applied Economics*, *1*(4), 1–33. doi: 10.1257/app.1.4.1

Ortega, M. A. (1992). *La esclavitud en el contexto agropecuario colonial: siglo XVIII*. Caracas: APICUM.

Özler, B. (2015). *Why is Difference-in-Difference Estimation Still so Popular in Experimental Analysis?* Retrieved from https://blogs.worldbank.org/impactevaluations/why-difference-difference-estimation-still-so-popular-experimental-analysis

Penfold Becerra, M. (2001). El colapso del sistema de partidos en Venezuela: explicación de una muerte anunciada. In J. V. Carrasquero, T. Maingón, & F. Welsch (Eds.), *Venezuela en transición: elecciones y democracia 1998-2000* (1st ed.). Red Universitaria de Estudios Políticos de Venezuela-RedPol: CDB Publicaciones.

Perlin, K. (1985). An image synthesizer. *Computer Graphics*, *19*(3), 287–296.

Pfeifer, U., Poersch, T., & Fuhr, N. (1996). Retrieval effectiveness of proper name search methods. *Information Processing & Management*, *32*(6), 667–679. Retrieved from https://www.sciencedirect.com/science/article/pii/S0306457396000428 doi: 10.1016/S0306-4573(96)00042-8

Plutzer, E. (2002). Becoming a Habitual Voter: Inertia, Resources, and Growth in Young

Adulthood. *The American Political Science Review*, *96*(1), 41–56.

Porto Tapiquén, C. E. (2020). *Shapefiles (*.shp) de Venezuela (Capas Base).* Retrieved from https://www.efrainmaps.es/descargas-gratuitas/venezuela/capas-base/

R Core Team. (2020). *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Raleigh, C., Kishi, R., & Linke, A. (2023). Political instability patterns are obscured by conflict dataset scope conditions, sources, and coding choices. *Humanities and Social Sciences Communications*, *10*(1), 74. Retrieved from https://doi.org/10.1057/s41599-023-01559-4 doi: 10.1057/s41599-023-01559-4

Ray, D., & Esteban, J. (2017). Conflict and Development. *Annual Review of Economics*, *9*(1), 263–293.

Reporters Without Borders. (2022). *2022 Round-Up: Journalists detained, killed, held hostage and missing* (Tech. Rep.). Retrieved from https://rsf.org/sites/default/files/medias/file/2022/12/RSF_Bilan2022_EN.pdf

Reuters Staff. (2012, 10). *Factbox: Venezuela's nationalizations under Chavez.* Retrieved from https://www.reuters.com/article/us-venezuela-election-nationalizations-idUSBRE89701X20121008

Riker, W. H., & Ordeshook, P. C. (1968). A Theory of the Calculus of Voting. *The American Political Science Review*, *62*(1), 25–42.

Riley, E. (2022). Resisting Social Pressure in the Household Using Mobile Money: Experimental Evidence on Microenterprise Investment in Uganda. *Working Paper*.

Rocha, R., Ferraz, C., & Soares, R. R. (2017). Human Capital Persistence and Development. *American Economic Journal: Applied Economics*, *9*(4), 105–136.

Rohner, D. (2024). *The peace formula: voice, work and warranties, not violence*. Cambridge, United Kingdom: Cambridge University Press.

Rohner, D., & Saia, A. (2019). Education and Conflict: Evidence from a Policy Experiment in Indonesia. *IDEAS Working Paper Series from RePEc*.

Rohner, D., & Thoenig, M. (2021). The Elusive Peace Dividend of Development Policy: From War Traps to Macro Complementarities. *Annual Review of Economics*, *13*(1), 111–131. doi: 10.1146/annurev-economics-073120-102652

Rohner, D., Thoenig, M., & Zilibotti, F. (2013). Seeds of distrust: conflict in Uganda. *Journal of Economic Growth*, *18*(3), 217–252. doi: 10.1007/s10887-013-9093-1

Rohner, D., & Zhuravskaya, E. (2023). *Nation Building Big Lessons from Successes*. CEPR.

Romano, J. P., & Wolf, M. (2005a). Exact and Approximate Stepdown Methods for Multiple Hypothesis Testing. *Journal of the American Statistical Association*, *100*(469), 94–108.

Romano, J. P., & Wolf, M. (2005b). Stepwise Multiple Testing as Formalized Data Snooping. *Econometrica*, *73*(4), 1237–1282.

Shemyakina, O. (2011). The effect of armed conflict on accumulation of schooling: Results from Tajikistan. *Journal of Development Economics*, *95*(2), 186–200. doi: 10.1016/j.jdeveco.2010.05.002

Shen, Z., Zhang, R., Dell, M., Lee, B. C. G., Carlson, J., & Li, W. (2021). LayoutParser: A Unified Toolkit for Deep Learning Based Document Image Analysis. *arXiv preprint arXiv:2103.15348*.

Siddiqui, M. M. (1967). A Bivariate t Distribution. *The Annals of Mathematical Statistics*, *38*(1), 162–166.

Sonntag, H., & Maingón, T. (2001). Cambio político y resultados de las elecciones de 1998. In J. V. Carrasquero, T. Maingón, & F. Welsch (Eds.), *Venezuela en transición: elecciones y democracia 1998-2000* (1st ed.). Red Universitaria de Estudios Políticos de Venezuela-RedPol: CDB Publicaciones.

Thistlethwaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, *51*(6), 309–317.

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B, Methodological*, *58*(1), 267–288.

Tullock, G. (1968). *Towards a Mathematics of Politics*. University of Michigan Press.

Uppal, Y. (2009). The Disadvantaged Incumbents: Estimating Incumbency Effects in Indian State Legislatures. *Public Choice*, *138*(1/2), 9–27.

Valencia Caicedo, F. (2019). The Mission: Human Capital Transmission, Economic Persistence, and Culture in South America. *The Quarterly Journal of Economics*, *134*(1), 507–556.

Van Aelst, P., Sheafer, T., & Stanyer, J. (2012). The personalization of mediated political communication: A review of concepts, operationalizations and key findings. *Journalism*, *13*(2), 203–220.

Velásquez Malavé, G. D. (2009). *Una Historia en Frecuencia Modulada: Producción de una serie de micro radiofónicos sobre la aparición de la FM en Venezuela* (Tech. Rep.). Universidad Católica Andrés Bello. Retrieved from http://biblioteca2.ucab.edu.ve/anexos/biblioteca/marc/texto/AAR7136.pdf

Vila, P. (1965). *Geografía de Venezuela*. Caracas: Dirección de Cultura y Bellas Artes, Departamento de Publicaciones.

Voigtländer, N., & Voth, H.-J. (2012). Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany. *The Quarterly Journal of Economics*, *127*(3), 1339–1392.

Wang, T. (2021). Media, Pulpit, and Populist Persuasion: Evidence from Father Coughlin. *The American Economic Review*, *111*(9), 3064–3092.

Watts, D. (1987). *The West Indies: patterns of development, culture and environmental change since 1492*. Cambridge: Cambridge University Press.

Westfall, P. H., & Young, S. S. (1993). *Resampling-based multiple testing: examples and methods for P-value adjustment*. New York ; Chichester: Wiley.

Wilpert, G. (2006). *Changing Venezuela by taking power: the history and policies of the Chávez government*. London: Verso.

Wiltshire, J. (2025). allsynth: Synthetic control bias-corrections utilities for Stata. *Working Paper*.

Wright, W. R. (1990). *Café con leche : race, class, and national image in Venezuela*. Austin: University of Texas Press.

Yanagizawa-Drott, D. (2014). Propaganda and Conflict: Evidence from the Rwandan Genocide. *The Quarterly Journal of Economics*, *129*(4), 1947–1994. doi: 10.1093/qje/qju020.Advance

Zhuravskaya, E., Petrova, M., & Enikolopov, R. (2020). Political effects of the internet and social media. *Annual Review of Economics*, *12*, 415–438. doi: 10.1146/annurev-economics-081919-050239

Zobel, J., & Dart, P. W. (1995). Finding approximate matches in large lexicons. *Software: Practice and Experience*, *25*. Retrieved from https://api.semanticscholar.org/CorpusID:6776819