UNIVERSITAT DE BARCELONA

Essays on Violence, Health, and Family Dynamics

Marie Beigelman

2024

PhD in Economics

Essays on Violence, Health, and Family Dynamics

Marie Beigelman

PhD in Economics | Marie Beigelman

UNIVERSITAT DE BARCELONA

LIBERTAS PERFVNDET

OMNIA LVCE

PhD in Economics

Thesis title:

Essays on Violence, Health, and Family Dynamics

PhD candidate:
Marie Beigelman

Advisor:
Judit Vall Castelló

Date:
April 2024

UNIVERSITAT DE BARCELONA

Car les parents pas faciles
Font des enfants tellement difficiles

- Oxmo Puccino

Acknowledgements

The journey to completing this thesis was a rocky one. Yet, every year that got me closer to defending got somehow easier, as I created bonds with mentors, colleagues, and friends who made this PhD a fascinating and rich experience.

My first acknowledgment goes to my supervisor Judit Vall Castelló. I wouldn't have hoped for such a skilled, humane, and approachable supervisor. I feel priviledged to have worked and grow as a researcher under your supervision. The support that you offered went way beyond an intellectual one, and you are a role-model on many aspects.

I also warmly thank my mentors, who - even though nothing bound them to do so - showed outstanding intellectual generosity and invaluable feedback in the past years, and even more so during the job market : Roberto Galbiati, Noam Yuchtman, Nathan Nunn, and Thomas Piketty. Your work has been a source of inspiration on so many levels, and I cannot thank you enough for your benevolence and continuous support in the past years. I am truly honored to count on you as mentors. I would also like to thank Camille Landais and again Nathan Nunn, for their warm welcome during my visitings at the LSE and UBC. These experiences marked a turn in my PhD, and I am indebted to you for making it possible.

I am also truly thankful for Noam Yuchtman, Ana Tur-Prat and Alfonso Herranz for participating in my defense committee. I couldn't hope for better researchers in Political Economy, Economic History and Development to reflect the pluri-disciplinarity of my thesis. I am grateful to share this day with you, and am honored to have such a jury.

I am also indebted to many researchers who took the time to give me feedback, suggested new ideas, and gave me the opportunity to share my work in various seminars. In particular, Siwan Anderson, Tim Besley, Alberto Bisin, Michael Callen, Denis Cogneau, Neil Cummins, Yannick Dupraz, Leopoldo Fergusson, Claudio Ferraz, Patrick François, Leigh Gardner, Cecilia Garcia Peñalosa, Libertad Gonzáles, Xavier Jaravel, Camille Landais, Jessica Leight, Sara Lowes, Gianmarco León-Ciliotta, Moses Shayo, Eric Schneider and Felipe Valencia Caicedo. I am also indebted to Historians Myriam Cottias and Cecile Vidal from the EHESS for their invaluable expertise in the History of Slavery and their precious guidance.

The next acknowledgment goes to Jordi Roca - the pillar of our school. Your amazing support and efficiency saved many of us during the PhD, and I hope you know how much this thesis owes to you ! I was generally lucky to be surrounded

by such amazing colleagues at UB: Elisabet Viladecans, Guillem Riambau, Jenifer Ruiz Valenzuela, Andreu Arenas, Rosa Sanchis-Garner, Francesc Amat... thank you for your support and for baring seeing me present the same thing over and over. Mari, Aly, Ghizlen, Dani, Filippo, Marta, Lili, Abhinav, Candan, Kaan, Ha, Claudia, Alessio, and all the others. You are the best possible group of friends to be with. I thank Lydia, for being such an amazing friend, mentor and pushing me to be ambitious. This thesis owes so much to you. And of course Sarah, for sharing many laughs and your huge heart. The job market year would have been very different without you both.

I am also grateful for the institutions that have supported my research along the way: the UB, AMSE, the STEG-CEPR, and the AES. I cannot thank enough Dimitri Garnier from the Departmental Archives of Guadeloupe, as well as the French Overseas National Archives in Aix, for allowing me to access precious archives and thus supporting my research in a substantial way.

My time as a PhD student cannot be distinguished from my time in Barcelona, which I was so lucky to call home for the past 5 years. I have loved every bit of my life in this city, and the friends I made here: Clem, Marce, Laura, Nico, Xavi, Elliot, Harold, Johanna, Clara, Marie, Raquel and the Casa del moviment.

I now want to thank my friends and my family for reminding me of what is important, and for their unwavering love and support. My mother, for her softness, strength, humor and her decisive influence in my decision to work on Slavery. My father, for showing me that laughing does help the medicine go down. Juliette and Simon, for your sensitivity, your passion, the laughs. Mimette, for all the joy you bring and your strength. Martine, for being my guardian angel always. My final acknowledgement goes to Arnault, I lack words to express how important you are, and how much this thesis owes to you.

Contents

1 General Introduction

Violence is a pervasive societal challenge with profound economic implications. The experience of violence can have durable impacts on the life-outcomes of victims: their health (Campbell, 2002), labor market opportunities (Sabia et al., 2013; Bindler and Ketel, 2022), educational attainment (Koppensteiner and Menezes, 2021) - but also on their descendants (Currie et al., 2022) and their peers (Carrell and Hoekstra, 2010), thereby contributing to the transmission of inequality and large costs for society (Hoeffler, 2017; Peterson et al., 2024).

Violence's toll goes beyond its immediate, negative effects on victims. Episodes of collective violence (wars, conflicts, colonization) can also have durable effects on long-run development and societies (Blattman, 2009; Grosjean, 2014; Tur-Prat and Caicedo, 2020; Alacevich and Zejcirovic, 2020; Besley and Reynal-Querol, 2014).

The experience of violence is, in addition, not random. Economically, socially and politically disadvantaged communities are much more likely to experience violence, whether it be collective - through political repression (Alacevich and Zejcirovic, 2020), wars, labor coercion... - or inter-personal[1] (e.g gender-based violence or racial violence). The experience of violence is, therefore, highly intertwined with the transmission of inequality.

The general goal of my thesis is to contribute to a better understanding of the causes and consequences of violence against minorities, with a particular focus on the expression of violence in the familiar setting. I use both historical and contemporary data and adopt a multifaceted approach: both in terms of the events of violence studied (institutional or interpersonal), the targeted minority (racial or gender), and the context (historical or contemporary).

The first section focuses on the lingering effects of *institutional* violence in a historical setting: New World Slavery. I focus on the French Caribbean islands of

[1] Interpersonal violence involves the intentional use of physical force or power against other persons by an individual or small group of individuals. Interpersonal violence may be physical, sexual, or psychological (also called emotional violence), and it may involve deprivation and neglect.

Guadeloupe and Martinique, and study the determinants of violence during Slavery (Chapter 2) and its lingering effects on the immediate descendants of former slaves (Chapter 3). The work presented in this section is the fruit of a large scale digitization effort of handwritten administrative archives (nearly 50,000 pages) on enslaved individuals and their descendants. To collect this new data, I built an optical character recognition pipeline tailored for French Handwritten Archives, which I present in Chapter 4.

Given the persistent effects of violence, investigating the efficiency of current policies in identifying and protecting victims of violence is of primary importance. The second section turns a contemporary context, and focuses on state's responses to *systemic* violence against women and children during the COVID-19 pandemic (Chapter 5).

Both sections focus on two potential drivers of violence in the familiar setting: intergenerational transmission of trauma, and exogenous stressful shocks. These two sections are also complementary in their approach to the problem of violence. While the first section focuses on the *mechanisms* through which violence has lingering effects, the second section focuses on the efficiency of institutional responses to intra-familial abuse. The health dimension is also highly present in all chapters: either as consequence of violence, using mortality as outcome of interest in Chapter 2 and 3 [2]; or as a cause of violence - the COVID-19 health shock - and institution of interest (the health sector) in Chapter 5.

Part One - Violence during French Caribbean Slavery

This first section focuses on the most violent form of racial oppression in modern history: Slavery, which was instrumental in European powers' expansion into the New World. It was also highly heterogeneous, both across and within countries. Enslavement conditions varied across plantations, with the worst and most violent found on sugarcane plantations (Fogel and Engerman, 1974; Higman, 1977; Smith, 1982; Follett, 2005). Yet, economists' approach to coercive systems has remained largely monolithic[3]

[2] I use mortality as a direct measure of violence during Slavery in Chapter 2, and child mortality as a measure of poor quality of childhood environment in Chapter 3

[3] Two exceptions in the case of coercive labor are Naidu and Yuchtman (2013); Dippel et al. (2020).

The purpose of the work presented in this section is two fold: to document the determinants of the experience of slavery (Chapter 2), and to explore how these different experiences might matter for the descendants of formerly enslaved (Chapter 3). I focus on French Carribean Slavery, in the islands of Guadeloupe and Martinique. These territories, which are now part of France, are amongst the poorest regions in Europe. While the memory of slavery and its legacy is currently at the center of French political debates, this context has remained largely under-studied by Economists and Economic Historians.

Chapter 2 explores the determinants of coercion intensity during Slavery, on the islands of Guadeloupe and Martinique. I focus on the *economic drivers* of enslavement conditions: both structural (depending on crop type), and cyclical (depending on competition in the sugar market). I find that the introduction of beet sugar in mainland France led sugarcane planters to escalate production through increased coercion, leading to an increase in mortality in the more sugarcane intensive counties. This descriptive evidence supports using crop types as reliable indicators of enslavement conditions. I thereby provide the first quantitative evidence on how changes in economic incentives can affect enslavement conditions.

Building on the findings of Chapter 2, Chapter 3 investigates the lingering effects of exposure to violence during slavery on immediate descendants after abolition. I compile a novel individual level dataset on *all* formerly enslaved families who had a child between 1852 and 1856, and explore the impact of parents' coercion history on child mortality, and family structure[4]. To determine where parents' were formerly enslaved, I rely on a unique institutional feature of the French West Indies: surnames given to the newly freed population at abolition were highly unique, distinct from slave owners', and specific to a former place of enslavement. This allows me to trace the coercion exposure of formerly enslaved individuals, even if they relocate. My key result is that children whose father was extremely coerced are 40% more likely to die before the age of five. I find suggestive evidence that this could stem from extremely coerced men being more violent. I thereby document a novel mechanism by which inequality resulting from slavery could result in lasting effects: trauma-related violence and child neglect.

This historical investigation is also relevant from a contemporary policy perspective. First, by further documenting the initial disadvantages caused by slavery, and drivers of these disadvantages, thereby contributing to the current debates on repa-

[4] I present the Optical Caracter Recognition pipeline used to digitize this novel dataset in Chapter 4

ration policies for descendants of former slaves (Darity and Frank, 2003; Piketty, 2020; Darity et al., 2022). Second, this research underscores the large effects of violence on the descendants' of survivors, emphasizing the need to consider the latter as victims themselves.

Part Two - Institutional Responses to Intrafamilial Violence during Pandemic Times

The next section of my thesis takes on a policy-related approach in the case of violence against women and children, which is a worldwide concern and a major public health problem[5].

The COVID-19 pandemic led to a worldwide surge in violence against women reports, prompting the United Nations to refer to a "shadow pandemic". Women's vulnerability to domestic abuse was indeed heightened by economic vulnerability and male unemployment (Anderberg et al., 2016), stress (Card and Dahl, 2011), and increased exposure to violent partners during lockdowns. Lockdown policies also disrupted the traditional channels of abuse reporting reporting: emergency shelters for victims were closed or imposed limited capacity; health centers were congested; and while police stations remained open, going to the police to seek help might have posed an even greater challenge for victims, while being on the watch of their violent partner.

The chapter of this section (co-authored with Judit Vall, Anna Garcia Altes, Neus Carrillero) focuses on one of the hardest hit Spanish regions by the pandemic - Catalonia - and in which the health sector is the most important channel of detection of male based violence. We exploit rich individual level data on cases of abuse detected by healthcare workers before and during the pandemic, and are able to identify changes in detection of abuse against children. Despite a large increase in help-seeking behavior of victims through hotline calls, we document a large and persistent decrease in care during and after lockdown, driven by the persistent pressure of the pandemic on the health beyond confinement measures. This is especially true for the most vulnerable groups of victims (children, lower income).

[5] 30% of women and young girls worldwide suffered from physical or sexual abuse from men at some point in their life, with numerous and long-lasting consequences on survivors' lives (World Health Organization, 2018)

Overall contribution

My thesis contributes to the literature on violence against minorities and its societal impacts in several ways. First, by showing that collective, institutionalized violence against minorities can translate into large negative effects on survivors' descendants through trauma and child neglect. This mechanism, which has long been discussed by psychologists in the case of slavery (Clark and Clark, 1950; Halloran, 2018), had yet to be explored empirically. Acemoglu et al. (2001), Dell (2010), Dell and Olken (2019) emphasize the role of persistent extractive institutions in driving long-term outcomes. In the context of slavery, Acharya et al. (2018); Althoff and Reichardt (2023) document how oppressive institutions in the US south contributed to sustain systemic discrimination today, while Nunn and Wantchekon (2011) and Bertocchi and Dimico (2014) show that persistent changes in norms induced by coercion exposure may also play a crucial role. The findings in this chapter highlight a significant effect of intergenerational transmission for inequality, underscoring the importance of further investment in this field. The second general contribution of this thesis lies in documenting important heterogeneity in violence exposure within targeted minorities. In Chapter 5, I find a much stronger effect of pandemic on the loss in detection of abuse - and therefore care provided to victims - among the most vulnerable groups: particularly children from low-income households. In Chapter 2, I find sizable differences in coercion intensity depending on planters' economic incentives, which I find reflected in substantial inequality among the next generation - through higher mortality rate of children of extremely coerced fathers. In both chapters, I document that children are highly vulnerable to violence against their parents: during slavery (Chapter 3), or against their mother during the contemporary context of the pandemic (Chapter 5).

An important and more specific contribution of this thesis lies in a better understanding of *what* New World Slavery was and its legacy. Chapter 2 contributes to the literature on the determinants of enslavement conditions and coercion intensity. I confirm the established narrative that enslavement conditions during slavery were strongly affected by planters' economic incentives, which varied with crop type (Fogel and Engerman, 1974; Higman, 1977; Smith, 1982). I show that the coercion gap actually *widened* following a plausibly exogenous shock on the sugar market, and the subsequent decrease in sugarcane prices. In a closely related context, Dippel et al. (2020) study the effect of the world sugarcane crisis on coercion levels

in the British West Indies, using wages and coercion-related incarcerations. This chapter uses a more direct and extreme measure of coercion intensity: mortality. This thesis also contributes to a burgeoning literature on the impact of coercion on individual families (Althoff and Reichardt, 2023) which allows for measuring coercion's effect even when families relocate. Althoff and Reichardt (2023) show that exposure to discriminatory policies, rather than slavery *per se*, could in fact drive the large Black-White income gap in the US. I find that *enslavement conditions per se* can have substantial intergenerational effects. I thereby contribute to the sociological and historical debates on the social consequences of slavery (Patterson, 1982), more specifically on families (Du Bois, 1908; Higman, 1975, 1977; Schneider, 2017; Trevon D. Logan, 2018), and provide the first empirical evidence on the interplay between enslavement conditions, family structure and child neglect. I also show that heterogeneity in the experience of slavery matters highly for its legacy on descendants.

The data collection efforts put into this thesis constitute valuable public goods. First, the novel data on enslaved population and their descendants contains precious information to approach the determinants of the experience of slavery, and its legacy. I will continue to enrich this dataset, and make it available to researchers and individuals. The OCR pipeline built to collect these valuable archives will also be released in open source. It contributes to scholars' effort in providing open source solutions for un-digitized text processing (Shen et al., 2021) in several ways. First, by creating an open-source Handwritten Data processing model, tailored from French Handwritten text. Second, by manually constructing a labeled dataset of handwritten French data, which can serve as crucial resource for other researchers looking to fine-tune their models. Third, by writing reproducible research notebook allowing for other researchers to reproduce our steps and train a TrOCR model for another language.

2 When Competition Hardens Coercion

2.1 Introduction

Slavery was instrumental in European powers' expansion into the New World. It was also highly heterogeneous, both across and within countries. Enslavement conditions varied across plantations, with the worst and most violent found on sugarcane plantations (Fogel and Engerman, 1974; Higman, 1977; Smith, 1982; Follett, 2005). Yet, economists' approach to slavery and coercive systems in general has remained largely monolithic [1].

The objective of this chapter is to explore the determinants of coercion intensity during French Caribbean Slavery, on the islands of Guadeloupe and Martinique. I focus on the *economic drivers* of enslavement conditions: both structural (depending on crop type), and cyclical (depending on competition in the sugar market). To do so, I exploit unique archival records on birth and death records of the enslaved population in the decade prior to the abolition of Slavery in 1848. I explore the relationship between coercion intensity, crop type, and competition on the sugar market.

10 years before slavery was abolished, I find that mortality in sugarcane intensive counties was 1 percentage point higher than in coffee intensive counties, respective to 1820 population (3% against 2.2%). These estimates are comparable to previously estimated mortality rates in the British Caribbean (Higman, 1975) and the French Caribbean (Cottias, 1989). I also find that the mortality gap between sugarcane and coffee intensive counties widened during the decade preceding abolition, because of a plausibly exogenous shock on the sugarcane market. The introduction of beetsugar in mainland France pushed sugarcane planters to escalate production

[1] Two exceptions related to coercive labor - but not slavery - are Naidu and Yuchtman (2013); Dippel et al. (2020)

through higher coercion up to the abolition, despite an overall decreasing trend in sugarcane prices. My findings indicate that changes in the competition on the sugarcane market - measured by the increase in beetsugar production - were associated with a differential +30% increase in deaths in counties with a 10 percentage point higher share of sugarcane from 1840 to 1847. I find that this increase in mortality occurred for both men and women, and was mostly pronounced for the working age group (between 20 to 40 years old).

Pushing some slaves to exhaustion - and inevitably death - could have been a profit maximizing decision in a context where sugarcane planters' dominion on the French sugar market was coming to an end. Indeed, words of abolition started to reach the islands starting the 1830's, and sugarcane prices were collapsing worldwide while beet sugar production was gaining terrain (Ward, 1978). In this context, a short-term strategy which consisted in producing as much as possible despite lowering world sugar prices likely allowed them to limit revenue losses. Aggregate statistics do suggest an overall increase in revenues from sugarcane up to the abolition of Slavery.

This chapter contributes to the literature on the determinants of enslavement conditions and coercion intensity. I confirm the established narrative that enslavement conditions during slavery were strongly affected by planters' economic incentives, which varied with crop type (Fogel and Engerman, 1974; Higman, 1977; Smith, 1982). I show that the coercion gap actually *widened* following a plausibly exogenous shock on the sugar market, despite the world trend in decreasing sugarcane prices. In a closely related context, Dippel et al. (2020) study the effect of the world sugarcane crisis on coercion levels in the British West Indies, using wages and coercion-related incarcerations. This chapter uses a more direct and extreme measure of coercion intensity: mortality.

I also contribute to our understanding of the economics of coercive systems (Naidu and Yuchtman, 2013). Producers' ability to resort to violence to prevent workers from moving or increase productivity has important consequences on their reactions to economic shocks. Studying the determinants of the economic divergence between Eastern and Western Europe following the Black Plague, Pamuk (2007) argues that an important factor lied in the coercive powers that feudal lords held over peasants: much stronger in the East, compared to the West. While the Plague led to a massive drop in labor supply, causing Western lords to increase wages to attract labor force, Eastern lords - whose power on the peasantry was

stronger - were able to intensify coercion to force peasants to work their lands. In a more related context - the British West Indies - Dippel et al. (2020) found that the world sugar crisis had a differential impact on workers' conditions (wages and coercion-related incarcerations) depending on the strength of the planter elite. In areas that were most dependent on sugar, and where sugar planters were the strongest, coercive power of planters did not weaken - allowing them to use coercive measures to keep workers on plantations.

2.2 Historical Background

This section provides historical context on the slave economy and the enslavement conditions depending on the type of crop grown.

2.2.1 The Sugar Economy

Guadeloupe and Martinique are two former French colonies that are now among the poorest provinces (*départements*) in France[2]. Each island is composed of approximately 30 counties, and is a typical example of extractive colony set up by European countries in the Caribbean. African slaves were imported massively between the 17th and 19th century to serve as labor in plantation systems, and made up 80% of the total population at abolition (Engerman and Higman, 2003).

Slavery was abolished in 1848 in a context of political crisis, with the re-establishment of the Republic that followed the revolution against the July Monarchy. By that time, the islands' population amounted to nearly 200,000 inhabitants.

The economy of the French West Indies centered on the production of exportable crops: sugarcane, coffee, cocoa, indigo. In the 19th century, sugarcane accounted for 50% of cultivated land, against 15% for coffee and 35% for other crops (cocoa, indigo and subsistence agriculture for local consumption). The distribution of enslaved workers across different crops remained relatively stable over time (Figure 2.A6), with 60% attached to sugarcane, 30% to secondary crops, and 10% attached to city dwellers (Schnakenbourg, 1977). The economic unit of production was the plantation. Its activity was ruled by the "Exclusivity Regime": planters were only allowed to export production to the mainland, and in return were granted preferential tariffs and the guarantee that their production would be purchased (Fallope, 1983).

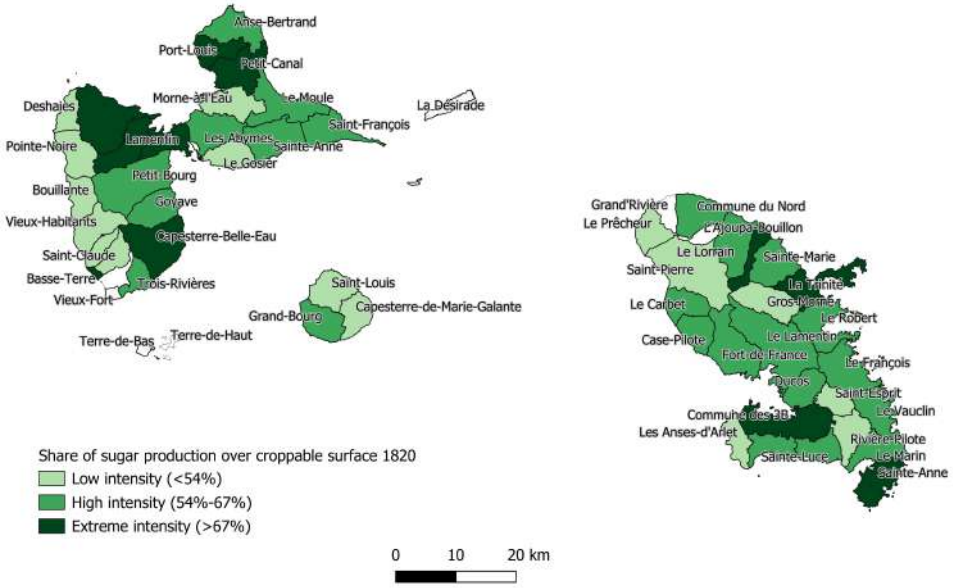
[2] INSEE

Sugarcane prices were determined competitively on the international market (Figure 2.A5), with France's sugarcane production represented a small fraction of the global output. Communication and relations between each plantations were minimal, with little within-island exchanges of enslaved workers even after the ban on slave trade in 1831 (Schnakenbourg, 1968).

Sugarcane plantations in the French West Indies were small in international standards, with an average of 50 enslaved workers in 1820 (against 15 in coffee plantations). Counties held around 15 distinct sugarcane plantations, and 40 coffee plantations (see Table 2.A1 for descriptive statistics on counties). These entities operated with considerable administrative and economic autonomy, with very little intervention from the authorities. The larger plantations had their own hospitals, prisons, and chapels.

2.2.2 Determinants of Sugarcane Intensity

Figure 2.1: Share of Sugarcane Over Cultivated Surface in 1820, in Guadeloupe (left) and Martinique (right)

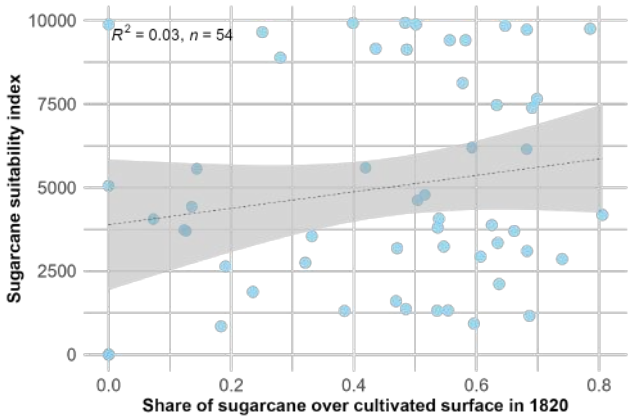


Note: Share of sugarcane over cultivated surface as per 1820 census data, **Source:** Ministère de la Marine et des Colonies

The relative homogeneity of the islands' soil and climatic conditions made it possible to produce sugarcane in nearly all counties[3].

[3]See Figure 2.A3 for a distribution of the sugarcane suitability, as calculated based on FAO's

Figure 2.2: Share of sugarcane over cultivated surface in 1820, and sugarcane suitability



Note: Own sugarcane SI calculated based on FAO's suitability indexes for sugarcane, for high inputs (large plantations). The black line represents the linear fit between the two variables, for a R^2 of 0.03.

This was also the case for subsistence agriculture, cocoa and indigo production. While initial suitability for sugar likely influenced the establishment of the first plantations, the high profitability of sugarcane from the 17th century onwards prompted planters to cultivate it extensively. By 1820, sugarcane had become the primary crop in nearly all counties, with a positive but weak correlation between the share of sugarcane cultivated area and initial sugarcane suitability (Figure 2.2).

Counties with different shares of sugarcane over cultivated surface could, therefore, be similar in terms of geographic characteristics or disease environment. When dividing counties along the median value of the share of sugarcane over cultivated surface (52%, see Table 2.A2), I find no statistically significant differences in the two types of counties in terms of soil characteristics, secondary crop suitability, or population density.

2.2.3 Enslavement conditions

> *"The planter is the sole arbiter of the law [...]. He judges and condemns as a sovereign feudal lord; he administers justice on his lands, and his judgments remain without appeal. Refusal to work, breaking doors, theft – everything falls under his jurisdiction [...]. What are we to think of a social state in which a man of human inclinations takes the whip himself*

suitability indexes (SI) for sugarcane, for high inputs (large plantations). Said indexes are available by grid cells that are typically larger than the counties. I re-calculate the county sugarcane suitability values as follow: for each county, I take the average of all FAO-grid values, weighted by the share of the county surface covered by the grid.

> *and strikes a woman until she bears twenty bleeding wounds on her body!! If the righteous can stoop to such a level, imagine what the wicked will devise."* (Victor Schoelcher, 1842, p40, p68)

Testimonies of the time provide an appalling picture of living and working conditions on plantations. Victor Schoelcher, an advocate for abolition in the early 19th century, described the enslaved population as malnourished, overburdened, inadequately dressed, and neglected[4].

Crop and enslavement conditions. Using anecdotal evidence and cross country comparison of mortality patterns (Smith, 1982), historians argued that enslavement conditions might have been worse on sugarcane plantations compared to coffee or cotton. Work on sugarcane plantations was indeed more physically demanding than for secondary crops, particularly during the harvest[5]. The manufacturing process following the harvest also posed risks with numerous accidents: crushed limbs during the cane breaking process, or fatal burns during the cooking of cane juice. As planters compelled workers to grow a significant portion of their own food, this heavier workload decreased the time that enslaved workers could dedicate to cultivate their gardens, which impacted nutritional intakes (Schnakenbourg, 1977).

Both harvesting and manufacturing on sugarcane plantation required a large labor force and substantial capital investments. In contrast, secondary crop production did not involve extensive manufacturing, and the harvesting process was less labor-intensive. As a result, sugarcane plantations were larger, averaging around 70 slaves on the eve of abolition, compared to an average of 30 slaves on coffee plantations. The ratio of enslaved individuals to non-slaves was also higher on sugarcane plantations. Sugar plantation owners could therefore be more prone to resort to physical abuse to prevent potential collective action. Larger plantation size and proximity to livestock also made sugarcane plantations a more disease prone environment.

[4] In addition to the workload, enslaved individuals were responsible for producing a significant portion of their own food by cultivating small garden plots. The constant fear of physical punishment ensured that slaves continued to work despite exhaustion. Planters held absolute power, and crimes against enslaved workers were more than often unpunished. Planters often justified torture or death of slaves by accusing them of poisoning cattle (Victor Schoelcher, 1842)

[5] Sugarcane involved a lengthy harvesting season followed by various manufacturing steps to transform the cane into raw sugar. The more demanding step consisted in land preparation before planting the new canes. Unlike coffee or cocoa production, which required planting only when the trees were no longer productive, sugarcane had to be replanted every year

Table 2.1: Descriptive statistics on enslaved workers, end of 18th century

	Secondary crop		**Sugarcane**		**Difference**	
	mean	*(std.)*	*mean*	*(std.)*	*mean*	*p-value*
"Value" (Fr.)	1545.32	(12.60)	1536.87	(13.23)	8.45	0.66
Age (mean)	25.02	(0.29)	26.09	(0.32)	-1.07	0.05
Women (%)	0.51	(0.01)	0.51	(0.01)	0.00	0.98
Black (%)	0.77	(0.01)	0.81	(0.01)	-0.04	0.01
Mulato (%)	0.23	(0.01)	0.19	(0.01)	0.04	0.01
Health issues (%)	0.08	(0.00)	0.06	(0.00)	0.02	0.01
N	3772	3772	3597	3597		

Notes: Own calculations using the Manioc database, indexing over 3,200 notarial acts from the 18th century and documenting over 14,700 individuals with the status of slaves in colonial society in Martinique. These statistics were calculated over a sample of 7,500 enslaved workers' records, for which the type of the plantation could be established.

Enslaved workers characteristics. The need for a stronger workforce may have driven sugarcane planters to acquire the "strongest" and more expensive enslaved individuals. Previous studies of plantation notary records conducted by historians show no systematic differences in selection on sex ratio, ethnicity or age composition across on plantations type (Debien, 1974; Vanony-Frisch, 1987). I further test selection into crop types using the Manioc database, which contains notarial acts from plantations in Martinique during the 17th century.

Table 2.1 contains descriptive statistics on enslaved workers for which the type of the plantation could be established (N=7,500). In line with Vanony-Frisch (1987), I find little differences both in terms of population composition (sex-ratio, age), and assigned "value" - which was a close measurement of the workers' supposed health and productivity.

2.3 When Competition Hardens Coercion?

The existing gap in coercion intensity between sugar and non-sugar plantations likely widened in the decades preceding the abolition of slavery because of a new competitor on the sugar market.

Sugarcane production increases with competition In 1815, The Napoleonic Wars ignited the decrease in world sugarcane prices (Blattman et al., 2007; Ben-

sassi et al., 2017; Dippel et al., 2020). Faced with embargoes from foreign powers, Napoleon advocated for the adoption of beet sugar to fulfill domestic demand, as colonial imports became impractical. Beet sugar offered clear competitive edges over sugarcane: it was cost-effective, more efficiently cultivated, and allowed for local production. While political turmoils from the Napoleonic wars slowed down the initial expansion of this emergent industry, beet sugar production started to increase from the 1830's - raising concerns among sugarcane colonists.

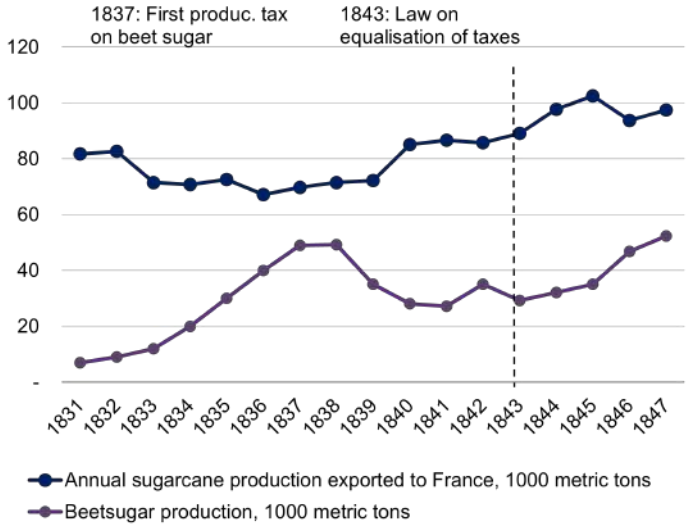
Sugarcane planters' initial reaction to this new competitor was to aggressively lobby for the introduction of specific taxes on beet sugar products. Their efforts paid: in 1837 the introduction of a specific tax on beetsugar led to the closure of several recently established factories in mainland France, causing beetsugar production to decrease by half in a few years time (Laloux, 2019)(Figure 2.3). However, colonists' economic arguments that sugarcane had to be preferred to beetsugar became increasingly difficult to sustain. In 1843, the preferential import tax system that had benefited sugarcane planters came to an end. This prompted a significant surge in beet sugar production, which accounted for 50% of sugar imports by 1847, against 5% in 1837[6].

Recognizing that lobbying alone could no longer secure their sugar monopoly, and in a context where words of abolition started to reach the island,sugarcane planters turned to a second strategy: maximizing sugarcane production, despite an overall decreasing trend in sugarcane prices (Figure 2.A5). This downward slopping relationship between aggregate production and sugarcane prices can be explained by two main factors. First, the exclusivity regime between France and its colonies guaranteed that colonists' production would be bought by France. Second, sugarcane prices were set in an international market, with French production representing a very little share of total outputs(Villeneuve, 1960).

Sugarcane production increases with coercion The increase in sugarcane production might be explained by three factors: changes in productivity, an expansion of the enslaved labor force, or an increase in the workload of enslaved workers. However, processes used in sugarcane production, such as machinery and fertilizers, remained unchanged throughout the 18th and 19th centuries (Schnakenbourg,

[6]Their lobbying efforts faced strong opposition from private investors who had injected significant capital in beet sugar production. Additionally, cultivating sugar beets was found to provide highly nutritious livestock feed, creating promising economic opportunities for the development of French livestock farming

Figure 2.3: Yearly Sugarcane (Blue) and Beetsugar production (Purple)



Notes: This figure displays the yearly aggregate exports in sugarcane, and beetsugar production in France (Source: *Tableaux Statistiques du Ministère de la Marine)*

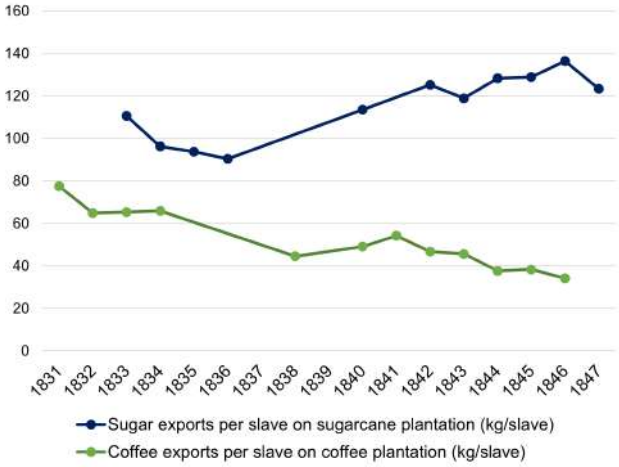
1977, 1987), until the establishment of the first sugarcane factories in the 1860s, after slavery was abolished. In addition, the ban on slave trade in 1831 led to an overall decrease in the enslaved population (Figure 2.A4). Consequently, changes in productivity or an increase in the enslaved population on sugarcane plantations are unlikely drivers of the increase in sugarcane.

Using aggregate statistics on the enslaved population attached to sugarcane or coffee from Schnakenbourg (1977), I find a +12% increase in the workload of enslaved workers on sugarcane, compared to those on coffee plantations (Figure 2.4).

Given the already harsh working conditions of enslaved workers on sugarcane, an increase in the workload would likely be met through higher coercion and overall worsening of conditions on sugarcane plantations through: (i) worse nutritional intake because of budget cuts, or because enslaved individuals have less time to grow food; (ii) exhaustion and degradation of health; (iii) physical abuse to maintain the work pace and prevent rebellious actions; (iv) accidents in the manufacturing process, which were much more frequent on sugarcane plantations and more likely to happen when individuals were exhausted.

Said increase in coercion likely came at the price of enslaved workers' lives. Pushing some of the slaves to exhaustion might, however, have been a profit maximizing decision, in a context where sugarcane planters' dominion on the French sugar market was coming to an end. First, words of abolition had started to reach

Figure 2.4: Workload per Enslaved Worker from 1831 to 1847, by Crop Type: on Sugarcane Plantations (Blue) and Coffee Plantations (Green)



Notes: This figure displays the yearly aggregate exports in sugarcane (resp. coffee) in kg divided by the total number of enslaved worker attached to sugarcane (resp. coffee). This back of the envelope calculation provides suggestive evidence of the evolution of the workload per enslaved worker on plantations, depending on crop types. Yearly exports and enslaved workers size were taken from Schnakenbourg (1977)'s collection of official statistics (*Tableaux Statistiques du Ministère de la Marine*)

the islands starting the 1830's. Second, sugarcane prices were collapsing worldwide while beet sugar production was gaining terrain (Ward, 1978). In this context, a short-term strategy which consisted in producing as much as possible likely allowed them to limit revenue losses or, for some, increase revenues (Figure 2.A7).

The next section explores empirically whether the mortality gap increased between sugarcane and coffee plantations in the decade preceding abolition, and the extent to which this could be attributed to changes in the pressure on the sugar market.

2.4 Data

2.4.1 Enslaved workers' deaths and births

In 1837, the French government introduced the mandatory registration of death and birth of enslaved workers[7]. These civil records provide a unique opportunity to ap-

[7]Planters had to pay taxes according to the number of their enslaved workers, but could easily hide this information from the government. This mandatory registration was done to keep track of

proach their living and working conditions through the lens of demographic events. They were available for approximately 40 counties, with some gaps in the covered years[8]. I construct a county level dataset on the number of enslaved deaths and births by county and year from 1837 to 1848 (252 observations)[9]. To identify which subgroups were most affected by potential changes in mortality, I also collect individual level information on all deceased individuals on a subsample of 7 counties for the period 1843 1847[10].

2.4.2 County characteristics and pressure on the sugarcane market

For each county, I collect information on land use in 1820 from the Statistical Bulletin of the Ministry of the Colonies. As land use was likely endogenous to events affecting the sugar market, I use information as close as possible to the introduction of beet sugar in France : 1820 for Guadeloupe, 1818 for Martinique. I define the sugarcane intensity as the share of sugarcane over cultivated surface in 1820. I also collect terrain characteristics from the FAO (terrain elevation, soil quality and suitability).

Measures of sugarcane and coffee suitability from the FAO were available at a larger grid level than the average county size, and are not available for all counties. As such, I do not use sugarcane suitability as my main measure of sugarcane intensity. For robustness checks, and to explore the relationship between sugarcane intensity and sugarcane suitability, I re-calculate a county level measure of sugarcane intensity by taking the average values of the sugarcane suitability grids covering each county[11](see Figure 2.A1 for both the original FAO values and the re-calculated indexes).

My treatment measure of pressure on sugarcane revenues is the yearly beetsugar production in France. I also consider an alternative measure, which are sugarcane

this tax base, and as a way to introduce more control over plantations. Incentives to declare deaths of enslaved workers were high, as taxes were proportional to the size of enslavers' labor force.

[8]Which is why I do not use this mortality data to measure pre-abolition coercion exposure in Chapter 3.

[9]I exclude counties which had less than four available years between 1837 and 1848.

[10]These counties all belong to Guadeloupe, and available registries covered the 1843-1847 period: Lamentin, Petit Bourg, Le Gosier, Deshaies, Les Abymes, Anse Bertrand, Vieux Fort. The individual level dataset contains 4,000 death records (2,500 from sugarcane intensive counties, 1,500 from non sugarcane intensive counties.)

[11]For each county-calculated average, the weight of each grid overlapping the county is equal to the share of county surface covered by the grid.

consumption prices in mainland France. I interpret an increase in beetsugar prodution (resp. decrease in sugarcane prices) as intensified pressure on sugarcane revenues. Yearly production and price series are collected from the yearly statistical bulletins from the French National Statistics Institute (INSEE, *Résumés retrospectifs*), and found in Villeneuve (1960).

2.5 Results

2.5.1 Increase in the mortality gap in sugarcane intensive counties

I now explore how death patterns evolved in the decade prior to abolition, depending on the sugarcane intensity in the counties of enslavement. I divide counties along the median value of sugarcane intensity (52%, see Table 2.A1) and plot the average yearly deaths in Figure 2.5 between 1837 and 1847. I find a difference in death patterns between the two groups, with an important increase in average yearly deaths in sugarcane intensive counties. Figure 2.A8 plots the average death to birth ratios for the two types of counties, and displays similar patterns: an increase in the death to birth ratio in sugarcane intensive counties, and stability in non sugarcane intensive counties.

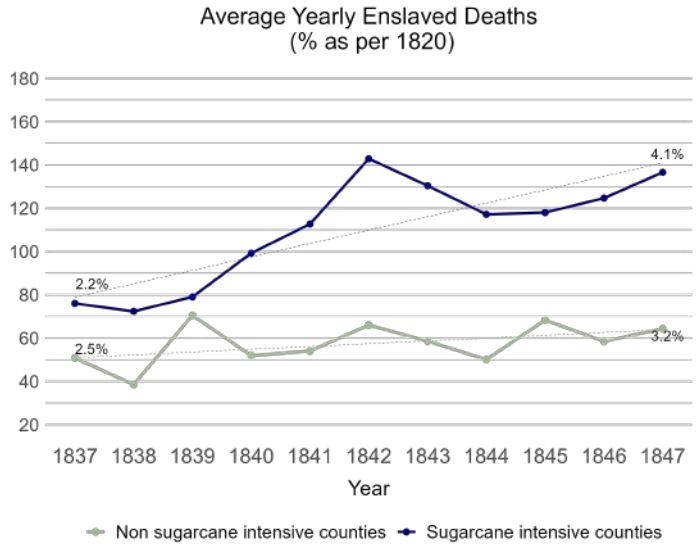
I formally assess whether enslaved population's living conditions were affected by sugarcane planters' response to decreasing revenues at the county level. I consider that a higher share of sugarcane production surface in a county in 1820 should increase workers' exposure to the competition shock (treatment exposure), which I expect should lead to an increase in the number of enslaved deaths.

I run the following log linear regressions at the county per year level:

$$LY_{c,y} = \alpha + \beta LSugarsurf_{c,1820} + \lambda LBeet_y LSugarsurf_{c,1820} + \eta_c + \gamma_y + \epsilon \quad (2.1)$$

Where LY refers to the log number of deaths or birth in county c and year y; $LSugarsurf$ is the log sugarcane production surface in 1820, which I consider as treatment exposure. I also use county sugarcane suitability and the number of sugarcane plantations in 1820 as alternative measures of exposure in robustness analyses. The continuous treatment $LBeet$ stands for the yearly log-beet sugar production in France in 1000 metric tons. In an alternative specification, I use $LPrice$ the yearly-log sugarcane consumption price. The coefficient of interest, λ, mea-

Figure 2.5: Average Number of Enslaved Deaths in Sugarcane Intensive Counties (Blue) and Non-Sugarcane Intensive Counties (Green), 1837-1847



Notes: This figure displays the average number of enslaved deaths registered in sugarcane intensive counties (resp. non sugarcane intensive counties), from 1837 to 1847. As statistics before 1834 are only available for Martinique, I restrict the descriptive figure to the 1837-1847 time period to ensure point averages are calculated over the same sample overtime. Sugarcane intensive counties are those whose share of sugarcane over cultivated surface is above the median (54%).

sures the association between average yearly deaths (or births) in more sugarcane intensive counties when pressure on the sugar market increases. η_c are county-fixed effects accounting for time-invariant characteristics likely to affect mortality (geographic characteristics, population density, etc.). In alternative specifications, I remove county fixed effects and include a time-invariant vector X_c of county-specific continuous variables[12]. Year fixed effects account for trends affecting mortality in all counties, as well as the yearly treatment variation (in price or production). Standard errors are clustered at the county level.

Main regression results are presented Table 2.2, specifications 1 and 3. I measure a positive association between enslaved yearly deaths and exposure to the pressure on the sugar market (interaction coefficient of 0.10). When comparing two counties with a ten percentage point difference in sugarcane intensity, I find that the increase in beetsugar production between 1840 and 1847 (+30%) was associated with a +30% increase in deaths from 1840 to 1847 in the more sugarcane intensive

[12]Total county size in 1820, slave population size in 1820, coffee production surface in 1820 as well as terrain characteristics (elevation, soil type)

county[13]. I find no effect on enslaved births, which suggests that changes in mortality patterns are not driven by population changes between and within counties. Results are robust to using different measures of exposure to the pressure on the sugarmarket (number of sugarcane plantations, sugarcane suitability in Table 2.3); using sugarcane prices as treatment (Table 2.3, specification 6); and robust in magnitude (although no longer significant) when using county time-invariant controls instead of county fixed effects (Table 2.2 specifications 2 and 4).

I therefore document a strong positive correlation between yearly pressure on the sugar market on the one hand, and increase in deaths in the more sugarcane intensive areas on the other hand.

2.5.2 Discussing potential drivers of mortality.

As treatment varies at the yearly level, I cannot rule out that other time varying factors unrelated to coercion might have increased the number of deaths in sugarcane intensive counties, compared to other. I now discuss such possible factors: natural disasters, epidemics, population changes, selection.

Natural Disasters. Two natural disasters occurred in 1843 and 1845 (forest fires, earthquake, hurricane) and caused heavy casualties in the city of Fort-de-France. I find that regression results are not sensitive to changing the timeframe to exclude the affected years (1843 onwards), or counties that were most affected by the disasters[14] (Table 2.3, specification 3 and 4). I searched for episodes of natural disaster, disease outbreak, and slave revolts that could have affected enslaved mortality in a sizeable way[15].

[13]Given that the interaction coefficient between log-beetsugar production and log-sugarcane intensity is 0.10, this implies that a one percentage point increase in beet production resulted in a 0.10% increase in enslaved deaths in a county with a one percentage point higher share of sugarcane, compared to another county. When considering two counties with a 10-percentage-point difference, a one-percentage point increase in beet production would lead to a differential increase in the number of deaths of 1%. Now, taking into account a 30-percentage-point increase in beetsugar production over the time period, this corresponds to a 30% increase in the number of enslaved deaths between 1840 and 1847.

[14]For Martinique, this includes the counties of Le Carbet, Case-Pilote, Lamentin and Ducos. For Guadeloupe, this includes the counties of Le Lamentin, Les Abymes, Le Gosier.

[15]I searched for natural disasters and disease related words in historical sources, using key words such as “disaster”, “catastrophe”, “disease”, “yellow fever”, “mortality” (Alfred, 1935). I read through the official statistical bulletins of the Ministry of the Colonies (“Tableaux Statistiques”) from 1837 to 1848, in which major events were typically recorded. I did not find traces of such events in the Colonial Ministry’s official statistics.

Table 2.2: Log-linear association between beet sugar production (1000 metric tons) and enslaved deaths and births between 1837 and 1847

	Enslaved deaths		Enslaved births	
	(1)	(2)	(3)	(4)
Beetroot prod. x Sugar surf. in 1820 (log)	0.10**	0.09	−0.004	−0.02
	(0.05)	(0.09)	(0.04)	(0.06)
Sugar surf. in 1820 (log)	−0.83	−1.05	0.34	0.24
	(0.64)	(1.10)	(0.46)	(0.80)
Cultivated surf. in 1820 (log)		0.65***		0.37**
		(0.20)		(0.19)
Population in 1820 (log)		−0.20		0.16
		(0.21)		(0.19)
Coffee surf. in 1820 (log)		−0.01		0.003
		(0.03)		(0.02)
Year FE	Yes	Yes	Yes	Yes
County FE	No	Yes	No	Yes
Terrain and soil charac.	Yes	No	Yes	No
Mean	79	79	77	77
N	252	243	252	243
R^2	0.91	0.73	0.91	0.70
Adj. R^2	0.89	0.70	0.89	0.67
F stat	36.88***	25.50***	37.62***	22.30***

Notes: This table displays the log-log linear estimated effect of changes in the yearly production of beet sugar (in 1000 tons), on the number of enslaved death by county, depending on the sugarcane intensity in the county of enslavement (measured by the log sugarcane surface in 1820). For λ a coefficient, the estimated effect of a one percentage increase in the covariate can be interpreted as a $\lambda\%$ increase in the number of enslaved deaths. $^{***}p < .01$; $^{**}p < .05$; $^{*}p < .1$

Changes in Population. The null association between changes in sugarcane prices and enslaved births suggests that an increase in the number of enslaved individuals on sugarcane plantations is not a likely explanation for the increased in enslaved deaths. An increase in the number of non-slave workers, such as indentured workers, would not be reflected in the enslaved population civil registries, and cannot explain my results. Internal migrations of enslaved workers could have occurred towards sugarcane plantations in need of labor. If that were the case, I would expect to see a decrease in the number of deaths and births in coffee intensive counties, of similar magnitude than what is observed in sugarcane intensive counties, which is not what I observe both descriptively (Figure 2.5) and in regression results (Table 2.A4). A final possibility is that some planters managed to resort to the illegal

Table 2.3: Log-linear association between beet sugar production (1000 metric tons) and enslaved deaths and births between 1837 and 1847: robustness checks using alternative measures of treatment exposure

	Enslaved deaths		Enslaved births	
	(1)	(2)	(3)	(4)
Beetroot prod. x Sugar suit. (log)	0.30**		0.003	
	(0.23)		(0.13)	
Beetroot prod. x Sugar plant. in 1820 (log)		0.21***		−0.01
		(0.09)		(0.07)
Sugarcane suitability (log)	−2.29		1.14	
	(2.99)		(1.62)	
Sugarcane plantations in 1820 (log)		−1.52		1.10
		(1.16)		(0.91)
Year FE	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Mean	79	79	77	77
N	243	238	243	238
R^2	0.91	0.91	0.91	0.91
Adj. R^2	0.89	0.89	0.89	0.89
F stat	38.02***	36.55***	37.62***	36.52***

Notes: This table displays the log-log linear estimated effect of changes in the yearly production of beet sugar (in 1000 tons), on the number of enslaved death by county, depending on the sugarcane intensity in the county of enslavement. I use two measures of sugarcane intensity: the log sugarcane suitability; the log number of sugarcane plantations in 1820. For λ a coefficient, the estimated effect of a one percentage increase in the covariate can be interpreted as a $\lambda\%$ increase in the number of enslaved deaths. $^{***}p < .01$; $^{**}p < .05$; $^{*}p < .1$

slave trade route. I expect that the death of illegally purchased workers would not be reported by planters and would therefore not be reflected in the enslaved civil registries. I still check whether results are sensitive to excluding counties close to ports, where planters were more likely to have access to the illegal trade (Fallope, 1983) in Table 2.A3, specification 4. Results are highly robust to this robustness check, suggesting that illegal slave trade is unlikely to affect my results.

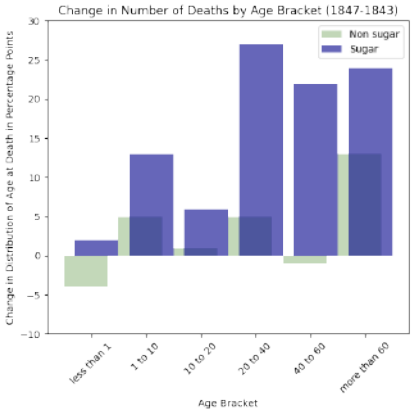
Selection. Another potential explanation could be selection. Namely, that enslaved individuals would be negatively selected relative to their health in sugarcane intensive areas, respective to the others. If that were the case, events affecting the entire islands might have had a stronger impact on them regardless of coercion conditions. I found, however, no selection bias on variables likely to reflect health

status or conditions ("value" and share of sick workers) (see Table 2.1). Finally, since labor conditions were more demanding on sugar plantations, I expect that any selection bias on enslaved workers' health should be positive. Therefore, the increase in yearly deaths that I document are likely to be lower bounds of the actual worsening of enslavement conditions in more sugarcane intensive counties relative to others.

2.5.3 Evidence that increased slave mortality stems from coercion

I now provide further suggestive evidence that the increase in death stems from an increase in coercion on sugarcane plantations, using the individual level sample of deceased enslaved individuals in a subset of 7 counties, from 1843 to 1847.

Figure 2.6: Changes in Average Number of Deaths, between 1843 and 1847

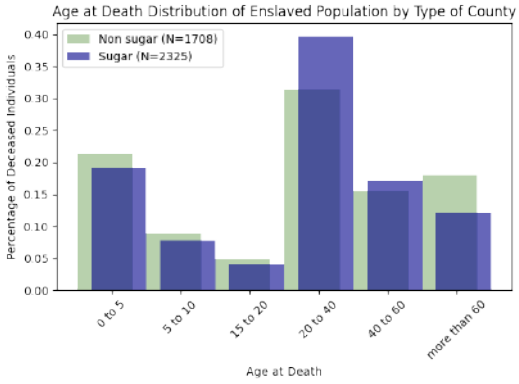


Notes: This figure displays the difference between the number of deceased workers in 1847 and 1843, for each age group and type of county (sugarcane intensive and non sugarcane intensive). Sugarcane intensive counties are those whose share of sugarcane over cultivated surface is above the median (52%).**Source:** Own computation and data collection from Civil Records of the Enslaved Population in 7 counties.

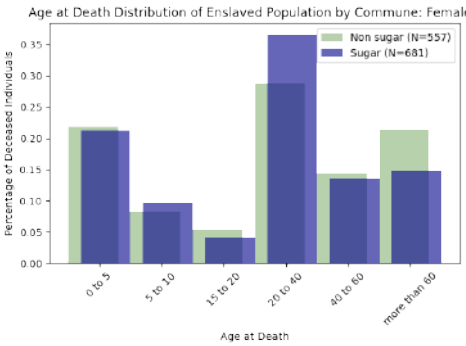
Working age individuals are dying more

I first examine differences in the age at death distribution depending on the sugarcane intensity in the county of enslavement. I find an increase in the number of deaths for all age-groups in sugarcane intensive counties (Figure 2.6), suggesting an overall degradation of living conditions determined by planters. I find, however,

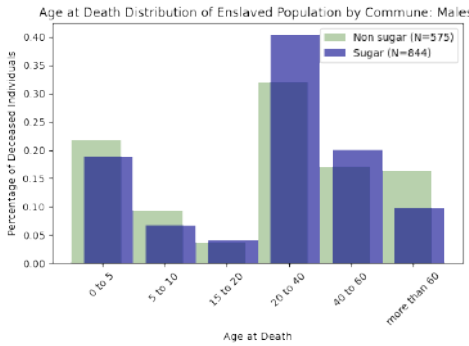
Figure 2.7: Distribution of Age at Death, Total Population and By Gender, between 1843 and 1847



(a) All



(b) Girls and Women



(c) Boys and Men

Notes: This figure displays the average distribution of age at death among deceased enslaved workers, between 1843 and 1847. Sugarcane intensive counties (in blue) are those whose share of sugarcane over cultivated surface is above the median in 1820 (52%). **Source:** Own computation and data collection from Civil Records of the Enslaved Population in 7 counties.

that this degradation of conditions was more pronounced for the working age category (Figure 2.7). I find a larger share of individuals dying at working age - 20 to 40 years old - in sugarcane intensive counties (38%) compared to non sugarcane intensive counties (29%). I note that said differences appear for both genders, suggesting that men and women were equally exposed to the intensification of coercion.

Seasonality of deaths

I now examine the seasonality of deaths, focusing on different types of counties and age groups. I first focus on under 5 years old mortality, which is known to be much more sensitive to disease and climate related environment than mortality at later ages. Régent (2010) documents that peaks of child mortality occurred during the rainy season, between July and January (which was also the time with the lowest number of births). I also find that child mortality increases throughout July to January, with peaks registered between November to January, suggesting a relatively similar disease environment in both types of counties. Turning to the working-age population (ages 15 to 50), I observe the highest mortality rates in sugarcane intensive counties during the months of May and June, which corresponds to the peak of the harvesting season, indicating a potential association between adult mortality rates and working conditions. Both the distribution of age at death and its seasonal patterns suggest that shifts in mortality within sugarcane-intensive regions are linked to on-plantation conditions, particularly working conditions, rather than external factors such as natural disasters or diseases.

2.6 Conclusion

This chapter exploits unique archival data on enslaved individuals deaths and births to document the determinants of violence during French Caribbean Slavery. I find that enslavement conditions responded to planters' economic incentives: both structural - depending on crop type - and cyclical, depending on the pressure in the sugar market and the political landscape.

While descriptive, the empirical evidence presented in this Chapter suggests large within-slavery heterogeneity in the experience of slavery, which is likely to have an important impact on the trajectories of formerly enslaved individuals after the abolition of slavery (Chapter 2).

To further our understanding of the determinants of the experience of slavery,

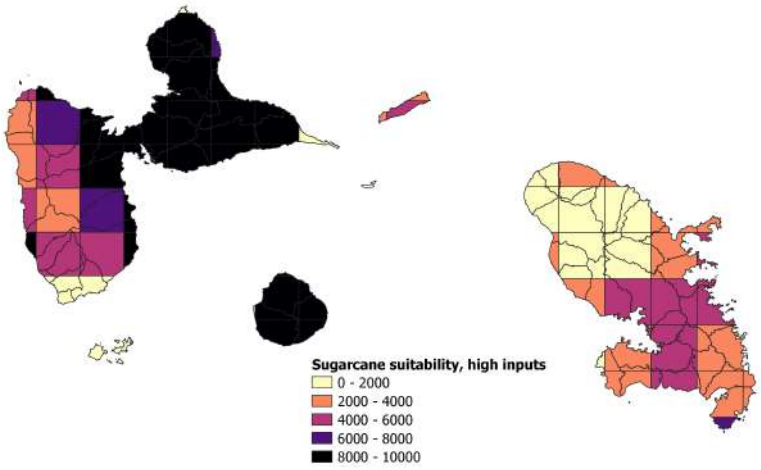
I am currently digitizing the full set of available birth and death registries on the enslaved population. This will allow me to study the individual role of planters' on violence, and its possible consequences for family formation, fertility, or rebellious actions from enslaved workers[16].

[16]Enslaved individuals that escaped plantation were called "Marrons".

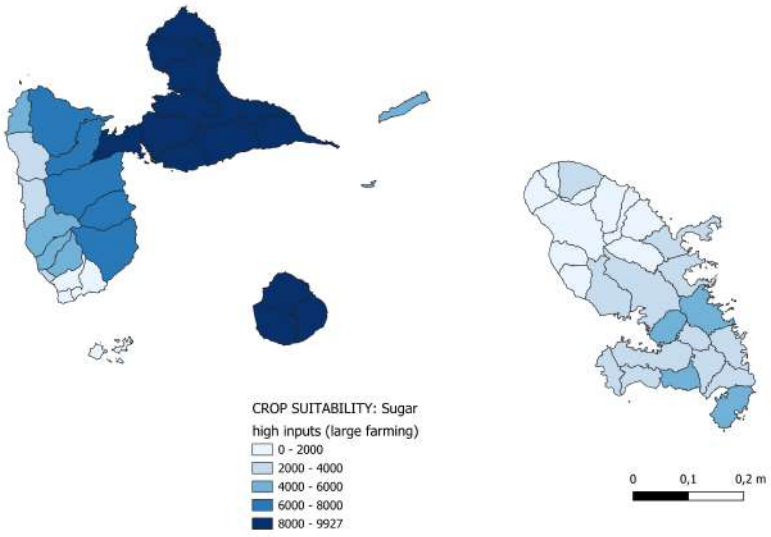
2.A Appendix, Chapter 2

2.A.1 Descriptive Statistics

Figure 2.A1: Sugar Suitability by County, original FAO and re-calculated indices



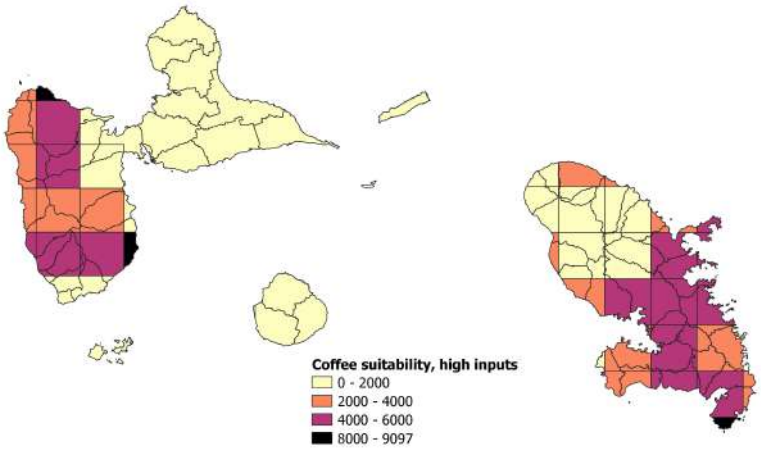
(a) Original values (high inputs)



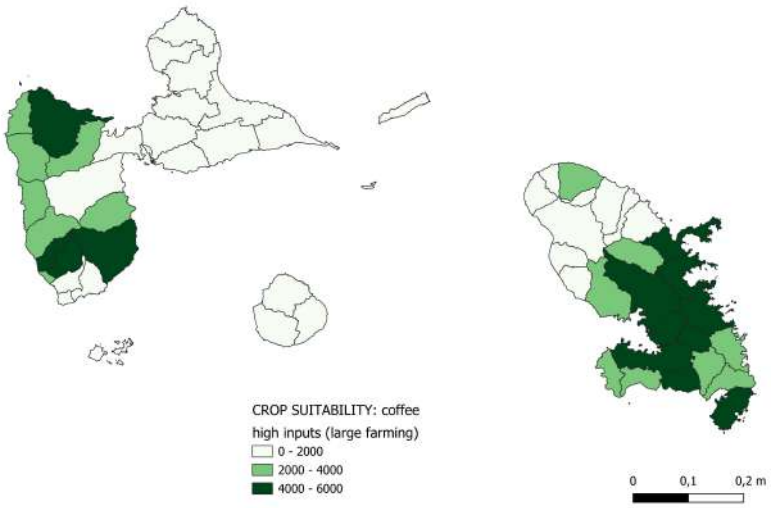
(b) Re-calculated values (average by county)

Source: FAO. The re-calculated values are averages of the grid-FAO values in each county (where the weight associated with a grid-value is equal to the share of the county surface covered by the grid)

Figure 2.A2: Coffee Suitability by County, High and Low Inputs



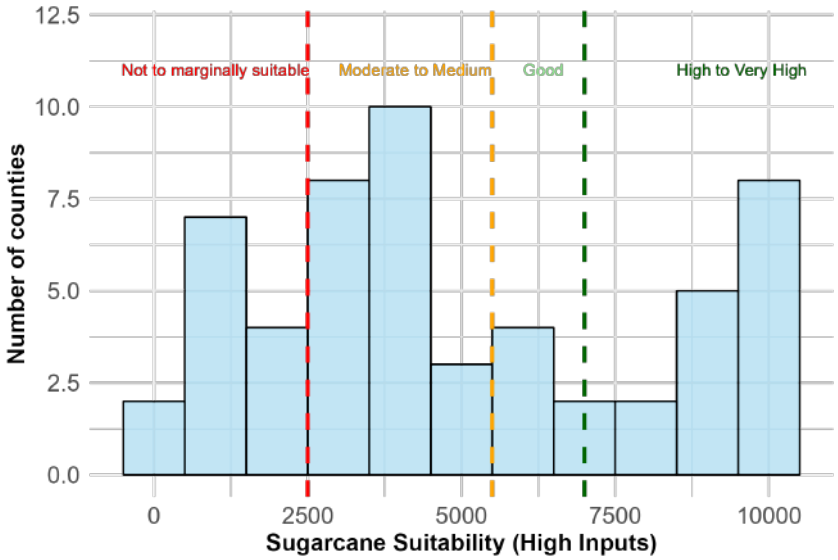
(a) Original values (high inputs)



(b) Re-calculated values (average by county)

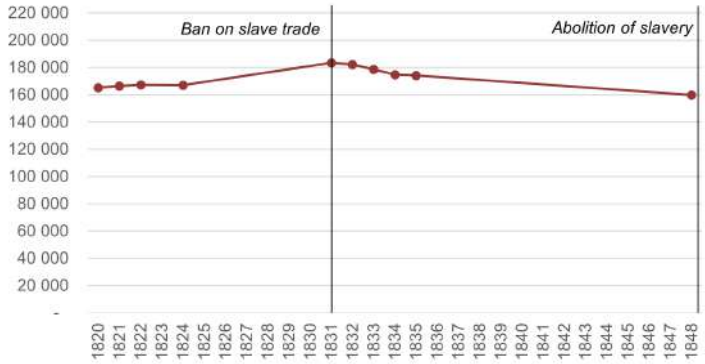
Source: FAO. The re-calculated values are averages of the grid-FAO values in each county (where the weight associated with a grid-value is equal to the share of the county surface covered by the grid)

Figure 2.A3: Distribution of sugarcane suitability index (SI) and suitability classes



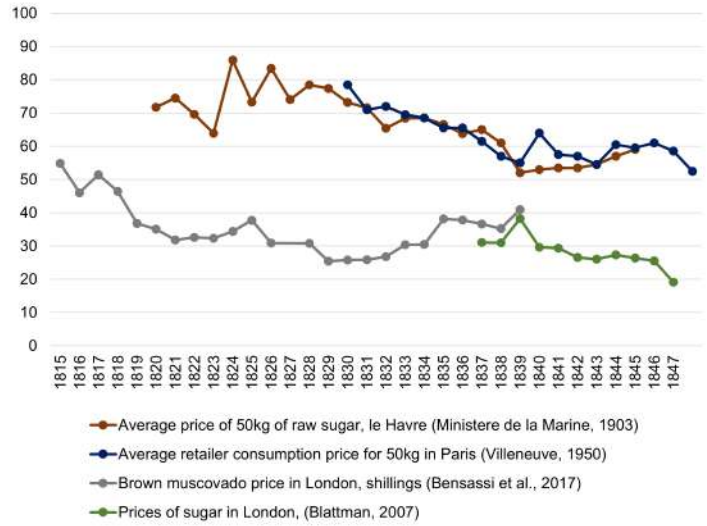
Note: Sugarcane suitability is calculated based on FAO's suitability indexes for sugarcane, for high inputs (large plantations). Said indexes are available by grid cells that are typically larger than the counties. I re-calculate the county sugarcane suitability values as follow: for each county, I take the average of all FAO-grid values, weighted by the share of the county surface covered by the grid. The lines delimit the different suitability classes used by the FAO: (1) Very high, when SI > 8500; (2) High, when 7000 < SI $\leq$ 8500; (3) Good, when 5500 < SI $\leq$ 7000; (4) Medium, when 4000 < SI $\leq$ 5500; (5) Moderate, when 2500 < SI $\leq$ 4000; (6) Marginal, when 1000 < SI $\leq$ 2500; (7) Very marginal, when 0 < SI $\leq$ 1000; and (8) Not suitable, when SI = 0.

Figure 2.A4: Enslaved Population in Martinique and Guadeloupe, from 1820 to 1848



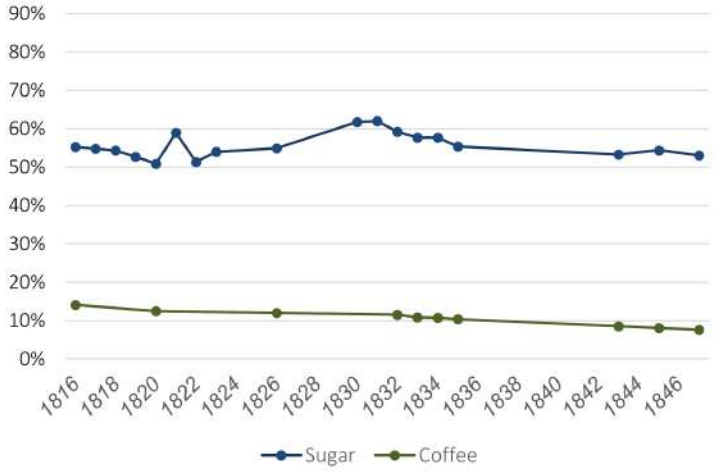
Notes: This figure displays the total enslaved population between 1820 and 1848 in both Guadeloupe and Martinique at different points in time, taken from Schnakenbourg (1977). **Primary sources**: *Tableaux de population et de cultures des Colonies Françaises, Ministère de la Marine et des colonies*

Figure 2.A5: Yearly price series for sugarcane in France and Great Britain



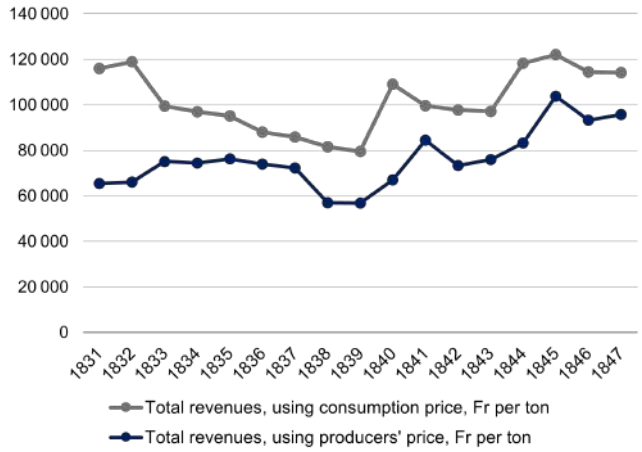
Notes: This figure displays yearly series of sugarcane prices from different sources in France and Great Britain

Figure 2.A6: Share of Sugarcane and Coffee Over Cultivated Surface



Notes: This figure displays the aggregate share of sugarcane and coffee over cultivated surface in Guadeloupe and Martinique, taken from Schnakenbourg (1977). **Primary sources**: *Tableaux de population et de cultures des Colonies Françaises, Ministère de la Marine et des colonies*

Figure 2.A7: Yearly revenues based on consumption prices and production prices



Notes: This figure displays the yearly product of total sugarcane production and price of sugar per ton, using either the consumption price (which would vary with import taxes), and producers price (which did not vary with import taxes). **Source**: consumption price series are taken from Villeneuve (Villeneuve, 1960), total production and producers price series are taken from official statistics on exports (*Annuaire officiel de la Guadeloupe, 1903*)

Table 2.A1: Descriptive Statistics on Counties in 1820

	Min	Median	Average	Std.dev	Max
Median Slope	2.5	4.8	4.9	1.3	7.0
Soil suitability	1.0	3.0	2.9	0.9	4.0
Sugarcane suitability	0.0	3,962.3	4,936.6	3,206.3	9,927.0
Coffee suitability	186.0	3,176.5	3,609.8	2,042.7	7,303.4
Total surface, hect.	13	3,081	3,280.3	1,904.2	8,480
Cultivated surface, hect.	13	949	1,055.8	742.4	3,320
Sugarcane surf.	0	401	565.5	484.9	1,980
Sugarcane over total surf.	0.0	0.12	0.24	0.13	0.43
Sugarcane over cultivated surf.	0.0	0.51	0.52	0.21	0.83
Sugarcane plantations, N	0	17	18.6	14.5	64
Coffee plantations*, N	0	36.5	32.5	24.3	84
Slave population	337	2,793	2,988.7	2,137.8	13,170
Total population	447	3,472	3,781.4	2,929.6	18,394
Slaves per hect. (density)	0.3	0.8	6.1	37.6	274.5

Notes: This table displays key summary statistics on the distribution of crops by county in 1820, over total surface and cultivated surface. Number of coffee plantations are calculated over Guadeloupe only, as information was not available for Martinique. **Primary sources**: *Tableaux de population et de cultures des Colonies Françaises, Ministère de la Marine et des colonies*

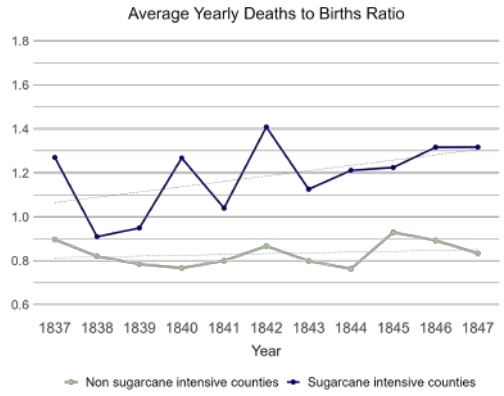
Table 2.A2: Descriptive Statistics on Counties in 1820, by category of county

Sugarcane over cultivated surf.	Intensive *(>52%)*		Non intensive *(52%)*		Difference in Means	
	Mean	*(St. Dev.)*	Mean	*(St. Dev.)*	Diff	*P-value*
Median Slope	4.53	(1.13)	5.18	(1.31)	-0.65	0.058
Soil suitability	2.94	(0.88)	2.77	(0.94)	0.17	0.499
Sugarcane suitability	5115.63	(3027.53)	4887.52	(3417.72)	228.11	0.798
Coffee suitability	3243.09	(1972.04)	3804.55	(2031.4)	-561.46	0.413
Total surface, hect.	3994.15	(1805.33)	2592.85	(1765.63)	1401.30	0.006
Cultivated surface, hect.	1335.04	(730.08)	786.81	(660.9)	548.22	0.006
Sugarcane surf.	844.27	(456.83)	297.04	(341.99)	547.23	0.000
Sugarcane over tot. surf.	0.22	(0.08)	0.10	(0.08)	0.12	0.000
Sugarcane over cult. surf.	0.64	(0.07)	0.29	(0.18)	0.34	0.000
Sugarcane plantations, N	24.50	(14.29)	12.52	(12.31)	11.98	0.002
Slave population	3497.85	(1670.26)	2498.44	(2439.05)	999.40	0.087
Total population	4250.23	(2214.48)	3329.89	(3466.72)	920.34	0.254
Pop density	0.93	(0.44)	11.09	(52.64)	-10.15	0.326

Notes: This table displays descriptive statistics on cultivated crops, topographic characteristics and population, in sugarcane and non sugarcane intensive counties. Sugarcane intensive counties are those whose share of sugarcane over cultivated surface in 1820 is above 52% (median value). The last two columns of the table display the difference in means, and the associated p-value calculated for each variable using t-tests. **Primary sources**: *Tableaux de population et de cultures des Colonies Françaises, Ministère de la Marine et des colonies* and *FAO*

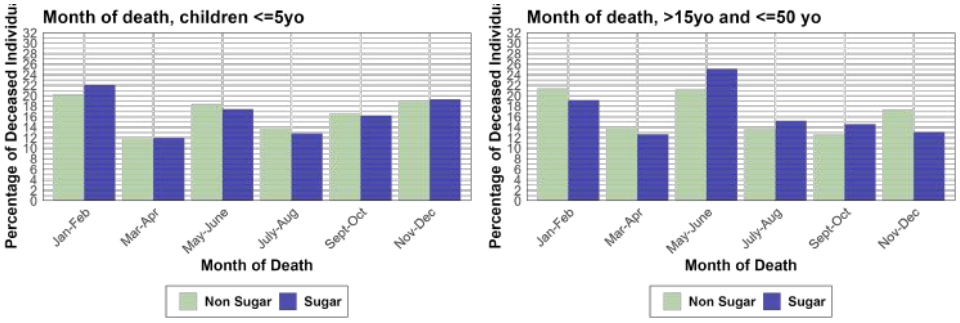
2.A.2 Additional Results

Figure 2.A8: Average Death to Birth Ratio in Sugarcane Intensive Counties (Blue) and Non-Sugarcane Intensive Counties (Green), 1837-1847



Notes: This figure displays the average death to birth ratio registered in sugarcane intensive counties (resp. non sugarcane intensive counties), from 1837 to 1847. Sugarcane intensive counties are those whose share of sugarcane over cultivated surface is above the median (52%).

Figure 2.A9: Distribution of Month of Death, between 1843 and 1847



(a) < 5 years old (b) > 15 yo to 50 yo

Notes: This figure displays the seasonality of month of death, by age and crop type. **Source:** Own computation and data collection from Civil Records of the Enslaved Population.

Table 2.A3: Log-linear association between beet sugar production (1000 metric tons) and enslaved deaths and births between 1837 and 1847: additional robustness checks

	Main (1)	1840-1847 (2)	1837-1842 (3)	Excl. cities (4)	Placebo (5)	Price (6)
Beetroot prod. x Sugarcane surface in 1820 (log)	0.10**	0.11***	0.12	0.09**		
	(0.05)	(0.05)	(0.09)	(0.05)		
Beetroot prod. x Coffee surface in 1820 (log)					−0.04	
					(0.04)	
Sugarcane price x Sugarcane surface in 1820 (log)						-0.33*
						(0.18)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
N	252	200	132	217	252	252
R^2	0.91	0.92	0.94	0.92	0.91	0.91
Adj. R^2	0.89	0.90	0.91	0.90	0.88	0.89
F stat.	36.88***	37.72***	27.66***	41.92***	35.79***	36.31*******

Notes: This table displays the log-log linear estimated effect of changes in the yearly production of beet sugar (in 1000 tons), on the number of enslaved death by county, depending on the sugarcane production surface in 1820 in the county of enslavement. Results presented in this table test the robustness of the results to: changing the time frames around natural disasters (specifications 2 and 3), excluding cities affected by natural disasters (specification 4), running a placebo test where exposure to changes in beetsugar production is measured using the coffee production surface in 1820, using the yearly sugarcane prices as treatment instead of yearly beetsugar production (specification 6). For λ a coefficient, the estimated effect of a one percentage increase in the covariate can be interpreted as a $\lambda\%$ increase in the number of enslaved deaths. ***p $<$ .01; **p $<$ .05; *p $<$ .1

Table 2.A4: Log-linear association between beet sugar production (1000 metric tons) and enslaved deaths and births between 1837 and 1847: placebo test

	Enslaved deaths		Enslaved births	
	(1)	(2)	(3)	(4)
Beetsugar prod. x Coffee surf. in 1820 (log)	−0.04	−0.05	−0.02	−0.02
	(0.04)	(0.08)	(0.04)	(0.07)
Coffee surf. in 1820 (log)	5.33***	0.58	3.64***	0.28
	(0.69)	(0.99)	(0.62)	(0.91)
Cultivated surf. in 1820 (log)		0.63***		0.37**
		(0.20)		(0.19)
Population in 1820 (log)		−0.17		0.17
		(0.21)		(0.19)
Year FE	Yes	Yes	Yes	Yes
County FE	Yes	No	Yes	No
Terrain and soil charac.	No	Yes	No	Yes
Mean	79	79	77	77
N	252	243	252	243
R^2	0.91	0.73	0.91	0.70
Adj. R^2	0.88	0.70	0.89	0.67
F stat.	35.79***	25.21***	37.70***	22.30***

Notes: This table displays the log-log linear estimated effect of changes in the yearly production of beet sugar (in 1000 tons), on the number of enslaved death by county, depending on the coffee intensity in the county of enslavement (measured by the log coffee surface in 1820). For λ a coefficient, the estimated effect of a one percentage increase in the covariate can be interpreted as a λ% increase in the number of enslaved deaths. $^{***}p < .01$; $^{**}p < .05$; $^{*}p < .1$

3 Impact of Enslavement Conditions on Families

3.1 Introduction

The enduring impact of slavery on long run inequality is a central policy concern, which fuelled debates regarding the implementation of reparations policies for formerly enslaved communities (Darity and Frank, 2003; Piketty, 2020; Darity et al., 2022). Economists are taking an increasing interest in this topic, and showed a strong negative association between slavery and long run economic outcomes (Nunn, 2008; Acemoglu et al., 2012; Bertocchi and Dimico, 2014; Laudares and Caicedo, 2023). Yet, our understanding of slavery's legacy suffers from a lack of evidence on several aspects.

Heterogeneity in the experience of slavery was substantial (Chapter 2). Yet, we have limited understanding on how this within-slavery variation might matter for the experience of freedom, and inequality within the formerly enslaved communities. As economists have mainly focused on the long run legacy of slavery, we are also lacking empirical evidence on the *initial disadvantages* caused by slavery. As said initial disadvantages were likely determinant for the trajectories of formerly enslaved individuals, documenting them is crucial to further our understanding of the mechanisms driving Slavery induced inequality. A final gap in the literature lies in the limited amount of evidence related to the *social* consequences of Slavery - particularly as it relates to families. The high rate of intra-familiar violence and absent fathers in Caribbean societies has led scholars to impute a potential link between Slavery on the one hand, and family environment in the other hand (Du Bois, 1908; Mulot, 2013; Mulot et al., 2023), with limited empirical evidence in that regard. While family environment could play a crucial role in perpetuating economic inequality (Black et al., 2005), identifying the causal impact of slavery on families is empirically challenging. Post-abolition institutionalized oppression, such as the

Jim Crow laws in the United States, or systemic racial discrimination, might equally participate to persistent economic inequality and confound the association between slavery and post-abolition outcomes (Althoff and Reichardt, 2023).

This chapter builds on the findings of Chapter 2 to study whether and how enslavement conditions during slavery[1] affected the quality of childhood environment and paternal presence[2] right after abolition in the islands of Guadeloupe and Martinique - where slavery was abolished in 1848. While nearly all families became free with little to no economic resources, differences in the experience of slavery could have affected their children in several ways. Indirectly, through parents' physical health and income, contributing to lower resources for their children. More directly, with psychological studies finding important effects of parents' past exposure to violence and trauma on their parenting practices and their children's environment (Kestenberg, 1980). According to Du Bois (1908), harsher conditions during slavery might have also led to more absent fathers and less stable family ties, which might also affect the resources dedicated to children.

Investigating the intergenerational consequences of violence during slavery poses significant empirical challenges. The first challenge relates to the proper identification of enslavement conditions. In my context, sugarcane was cultivated in nearly all counties but at varying intensities. The second challenge relates to the identification of individuals' history of coercion. The context of the French West Indies provides a unique opportunity in this regard, owing to the highly distinctive surnames assigned to formerly enslaved families. After abolition, the "Census of the Newly Freed" was conducted to assign civil identities to the 150,000 liberated individuals. Government officials guided the selection of surnames, ensuring they were unique, distinct from each other and from white settlers' surnames[3] (Durand, 2011). This provides a rare opportunity for precisely linking families to their former county of enslavement, and thereby coercion history. A final challenge faced by intergenerational studies in general lies in the selective measurement of transmission through fathers due to the inheritance of surnames. In the French West Indies, par-

[1] Throughout this chapter, I use the terms enslavement conditions and coercion intensity interchangeably to refer to the overall components of enslaved populations' living conditions during slavery. This encompasses direct violence, workload, and living standards.

[2] The choice of outcomes was driven by my wish to focus on family environment and the available information in data sources. This chapter is part of a broader research agenda, in which I explore the legacy of slavery across several generations and a wider-range of outcomes.

[3] Official guidelines recommended the use of anagrams, mythological names, etc.

ents within the same household often have distinct last names, which are both listed in civil registries[4]. This unique setting allows me to measure both maternal and paternal transmission.

To to so, I compile a novel individual level dataset on all formerly enslaved families who had a child between 1852 and 1856, which was the product of a large-scale digitization effort of nearly 50,000 pages of handwritten administrative data. To collect this archival data, I develop an open-source Optical Character Recognition Pipeline specifically tailored for handwritten documents, which I present in Chapter 4.

My baseline empirical strategy compares families depending on their former enslavement conditions, determined by the share of sugarcane over the cultivated surface in the surnames' county of origin. I regress family-level outcomes on coercion history while controlling for county of residence and year fixed effects. I am able to disentangle the specific effect of place-of-origin coercion from place-of-residence effects because most of formerly enslaved individuals moved within the islands after slavery ended. As mobility decisions can be endogenous to other determinants of family health and structure, I run alternative specifications to assess the sensitivity of baseline results to potential selection into mobility (focusing on movers only, excluding county fixed effects). Causal interpretation of coercion history's effect on post-abolition family outcomes lies in the assumption that enslaved workers were assigned to different crops upon purchase in a quasi-random way. I verify this assumption using both external sources (Vanony-Frisch, 1987) and information on 7,000 enslaved workers from notary records[5], and find no significant differences in socio-demographic characteristics and price upon purchase between enslaved workers of different crop types.

My primary finding is that, while single-mother households are common after abolition, they did not constitute the predominant family structure: 60% of children did have a known father. I also find a potential positive effect of worse enslavement conditions on fathers' presence. Although this effect is small in magnitude and no longer significant once place of residence effects are netted out from maternal

[4]With children typically inheriting their mothers' surname - unless parents were married, which only concerned 8% of individuals immediately after abolition.

[5]From the Manioc database.

coercion history[6], this finding calls for a more nuanced approach to the idea that violence during slavery should have led to more absentee fathers.

My second key finding indicates that among present fathers, the most extremely coerced ones had a negative impact on the quality of childhood environment for their immediate descendants. Their presence increased the risk of child mortality by 40%, compared to families with less coerced fathers *or* absent fathers. Differences in mothers' former coercion exposure had no effect, despite men and women being exposed to similar working conditions during (and after) slavery. This suggests that poorer health status of extremely coerced fathers - which should also be the case for extremely coerced mothers - cannot account for the deteriorating effect of these fathers. I also find that this effect holds when focusing on higher income fathers, low child mortality areas, and regardless of mothers' characteristics. A potential explanation for this gender differences in parental response to extreme coercion is that extremely coerced men might have been more violent: I find suggestive evidence that formerly enslaved men exposed to worse conditions committed more violent crimes and sexual assaults. I find that mothers might have responded to coercion's effect on fathers through matching decisions. In particular, extremely coerced mothers were 20% less likely to have a child with a man exposed to the same enslavement conditions during slavery, relative to less coerced mothers.

Considering that child mortality is an extreme measure of poor health, and that experiencing a child death could have important negative spillovers on family members, my findings point to substantial inequality among descendants of formerly enslaved individuals. I highlight that said inequality could be driven by the lingering effects of historical violence on fathers' negative contribution to the quality of childhood environment, possibly through violent behavior.

This chapter contributes to our understanding of how extractive institutions can negatively impact the economic outcomes of descendants. Acemoglu et al. (2001), Dell (2010), Dell and Olken (2019) emphasize the role of persistent extractive institutions in driving long-term outcomes. In the context of slavery, Acharya et al. (2018); Althoff and Reichardt (2023) document how oppressive institutions in the US south contributed to sustain systemic discrimination today, while Nunn and

[6]Using the alternative specification, which focuses on mothers who moved from their county of enslavement.

Wantchekon (2011) and Bertocchi and Dimico (2014) show that persistent changes in norms induced by coercion exposure may also play a crucial role. This chapter introduces a novel mechanism by which income inequality resulting from slavery can have a lasting impact: the intergenerational transmission of trauma and violence. While psychologists and historians have long been studying intergenerational trauma among descendants of formerly enslaved communities (Clark and Clark, 1950; Halloran, 2018), economists have yet to explore this area. The findings in this chapter highlight a significant effect of intergenerational transmission for inequality, underscoring the importance of further investment in this field.

While most of the literature on slavery's legacy has used country or regional level comparison (Nunn, 2008; Acemoglu et al., 2012; Bertocchi and Dimico, 2014; Acharya et al., 2018; Laudares and Caicedo, 2023), this chapter contributes to a burgeoning literature on the impact of coercion on individual families (Althoff and Reichardt, 2023) which allows for measuring coercion's effect even when families relocate. Althoff and Reichardt (2023) show that exposure to discriminatory policies, rather than slavery *per se*, could in fact drive the large Black-White income gap in the US. I find that *enslavement conditions per se* can have substantial intergenerational effects. I thereby contribute to the sociological and historical debates on the social consequences of slavery (Patterson, 1982), more specifically on families (Du Bois, 1908; Higman, 1975, 1977; Trevon D. Logan, 2018), and provide the first empirical evidence on the interplay between enslavement conditions, family structure and child neglect. I also show that heterogeneity in the experience of slavery matters highly for its legacy on descendants.

My findings thereby expand on our understanding of the effects of parental inputs at early ages of children (Heckman et al., 2013). Prior research has stressed that the lack of inputs from one parent (because of divorce, incarceration, death) can have important negative consequences for affected children (Gay Painter and David I. Levine, 2000; McLanahan et al., 2013; Kalil et al., 2016; Dupraz and Ferrara, 2023). I offer the first analysis of fathers' influence for early childhood development within a coercive context. My historical findings align with recent causal evidence presented by Norris et al. (2021), who found that the imprisonment of fathers can in fact have positive effects on children, particularly when there is a risk of abuse.

3.2 Measuring Past Enslavement Conditions

In Chapter 2, I documented the Slavery institution and Sugar economy of the French West Indies. Using newly collected data on enslaved mortality at the county, I found strong evidence that coercion intensity was higher in more sugarcane intensive areas and assessed the reliability of using the share of sugarcane over cultivated surface as measure of enslavement conditions. Building on these findings, I now present the methodology that I adopt to explore the lingering effects of coercion intensity on descendants after abolition.

3.2.1 Data collection

I study enslavement conditions' intergenerational impact on descendants after abolition between 1852 and 1856[7].

Data sources

To do so, I compile a novel dataset on *all* formerly enslaved individuals who had children after abolition (18,000 newborns) using two administrative archival sources.

Census of the Newly Freed, 1848: The "Census of the Newly Freed" was conducted between 1848 and 1852 to give civil identities to 150,000 freed individuals. Government officials recorded individuals' basic demographic characteristics, former place of enslavement, and assigned them surnames. Official guidelines prohibited the use of white settlers' surnames, and recommended mythological names, name transformation - anagrams, letter intercession – to ensure that surnames would be distinct from each other (Durand, 2011). As a result, family names given to newly freed individuals were highly distinguishable, providing a rare opportunity for tracking families using surnames as a quasi-perfect identifier of an individual's ancestry. Own data collection effort, transcriptions from genealogists ("Anchoukaj") and the Departmental Archives of Martinique and Guadeloupe, allowed for the collection of nearly 110,000 registries covering 70% of the newly freed population, with 45,000 distinct surnames. 90% of these surnames were given to less than 6 individuals[8] (Figure 3.1). I exploit this unique dataset to assign a given sur-

[7] Enslaved mortality data is available for 60% of the counties. As such, I do not use pre-1848 enslaved mortality as a direct measure of coercion exposure.

[8] Individuals sharing the same surnames typically referred to a mother and her children. Indeed, marriages were uncommon, and children usually bore the surname of the mother

name to a county of origin. For surnames given to several counties (5% of the family names), I consider the county with the majority of individuals bearing the surname as the county of origin. I then use the share of sugarcane over cultivated surface in the counties of origin to determine surnames' exposure to coercion intensity prior to 1848.

Birth and Death records, 1852-1856: I study coercion's impact on family environment along two dimensions: family structure (absent fathers), and child development (child mortality). To do so, I use never yet digitized civil records covering all birth and death events in the French West Indies population from 1852 to 1856. I digitized 30,000 birth records and 25,000 death records with information on date, place of registration, name and gender of newborn or deceased and parents (name, age, occupation, place of residence). Death records were exclusively used to identify children who died before 1857[9]. I identify birth records referring to formerly enslaved mothers, and match mothers' names with the census to retrieve its former enslavement conditions. I am able to achieve a 80% matching rate. Using only the matched birth records of formerly enslaved mothers, I then identify fathers' former coercion exposure using the same matching process [10]. I provide further details on the matching procedure in Appendix, section 3.A.1. The final birth sample comprises 18,000 births, born of 15,000 mothers, with 9,000 who had more than one child. These mothers represent 60% of the female population aged 17 to 40 years old in 1848.

Criminal records. I use additional data sources to investigate mechanisms and explore the link between historical coercion and post-abolition behaviors. Using the Overseas French National Archives search engine[11], I compiled a list of *all* French convicted criminals sent to the French colony of Guyana between 1850 and 1950, totaling 100,000 individuals. Within this comprehensive list, including convicts from both the French mainland and the colonies, I identified 3,500 surnames unique to the French West Indies. I employed this newly collected data to examine whether surnames originating from these areas were disproportionately represented

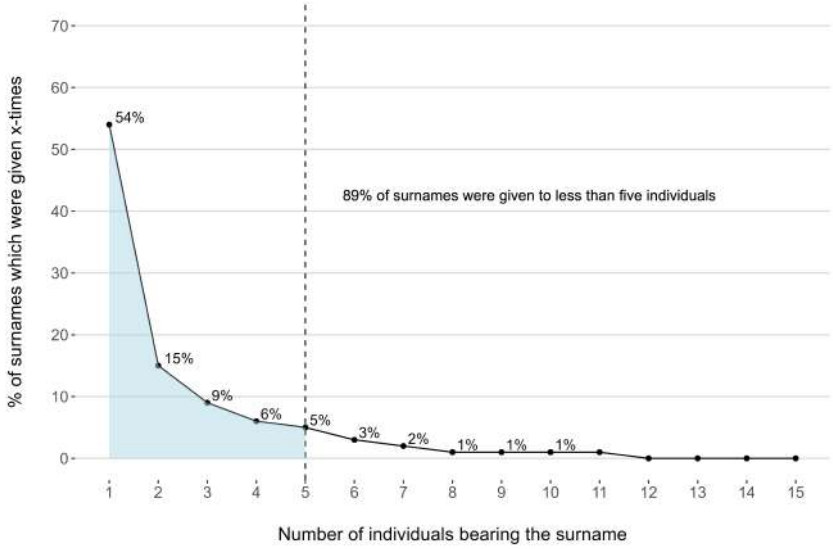
[9] In some counties, death records were not available - this was the case for Pointe a Pitre, the main city of Martinique

[10] 4,000 children had a formerly enslaved mother, and a father whose former exposure to coercion could not be determined. When analyzing the conjunct effect of mothers and fathers' coercion exposure, I remove these observations from the analysis.

[11] Accessed via "Registres des bagnards" with surname-based search available at IREL

among transported criminals. In addition to the surname dataset, I digitized the records of 700 individual convicts residing in the French West Indies (see Image 3.A3 for an example). These convicts belonged to the specific category of "réclusionnaires", sent to labor camps in Guyana with the possibility of returning after serving their sentences. Among this group, 300 were formerly enslaved on the islands, and had information on their county of birth. I use this subsample of detailed criminal records to explore differences in the type of crime commited.

Figure 3.1: Distribution of Surnames given to the Newly Freed Population in 1848, According to Number of Times they were Given



This figure displays the distribution of surnames given to the newly freed population, according to the number of times the surname was assigned to a newly freed individual in the census. We can read on the figure that 89% of these surnames were given to less than 6 individuals. Most of the time, these individuals were from the same family. This visual shows that formerly enslaved surnames are highly unique.

Data collection tool

The large number of handwritten pages that I processed called for an automated approach. Accessible and reproductible solutions involving Optical Character Recognition are often well suited for printed text, but fall short when dealing with handwritten documents - even more so for archival documents. Despite thriving research in document processing techniques published in open source, researchers often face high costs when dealing with archival documents. With the support of the STEG-CEPR PhD Grant Scheme, I have developed a comprehensive pipeline designed to

digitize scanned archives from the 19th century. I leverage on machine learning techniques, including Optical Character Recognition (OCR) and the use of a Large Language Model (LLM) to extract key information from OCRed text. The pipeline encompasses three stages:

- Parsing Layout: The initial stage involves layout parsing, wherein global images are segmented into smaller units, each containing a single line of text.
- Optical Character Recognition (OCR) Module: Following layout parsing, the OCR module takes center stage, transcribing individual images into text.
- Named Entity Recognition (NER) Module: In the final stage, the Named Entity Recognition module extracts key information from the transcribed archival content.

The OCR module is the most notable contribution of this pipeline: I developed a model of text recognition tailored for French handwritten text from archival records. It is a special case of the TrOCR (transformer OCR) (Li et al. 2022), which I adapted to French handwritten text using archival records from enslaved and formerly enslaved civil records. More details on each step of the pipeline and use cases are given in Chapter 4.

3.2.2 Main variables of interest

Child mortality. Children who did not survive prior to 1857 were identified by matching birth and death records using date of birth, age of the child at death, name of the child, and parents information (name, age, occupation). The recording of dates of birth and death allows for precise measurement of children's survival at the monthly level.

Presence of father in birth record. Birth records provide precise information regarding the identification of a child's parents, with statements such as "Mr. X acknowledges himself as the father." While maternal details are consistently recorded, paternity is unspecified in more than half of these registries. As official birth records serve as legal evidence of a child's lineage, the omission of paternal information implies that, legally, the child lacks a recognized father.

3.2.3 Empirical Strategy

I use parent's former place of enslavement, which I determine based on their surnames, as a plausibly exogenous variation in enslavement conditions to identify the effect of slavery on family structures and child mortality.

Treatment. For each parent, treatment is defined as a categorical variable, where exposure to the most extreme conditions will be distinguished from lower and high coercion exposure. This allows me to identify whether enslavement conditions have a non-linear effect, while still including children born from unknown fathers in the analysis.

- Lower Intensity: Sugarcane production surface over cultivated surface below the median (30 counties, for 7,325 mothers and 2,398 known fathers)
- High Intensity: Median to 85th percentile (18 counties, for 7,450 mothers and 2,159 known fathers)
- Extreme Intensity: 85th to 100th percentile; (12 counties, for 3,421 mothers and 1,349 known fathers)
- *Unknown Fathers (fathers only)*

In alternative specifications, I test the sensitivity of my results to using different thresholds for the extreme category (80th and 90th percentile, instead of 85th).

General Specification. I estimate the effect of enslavement conditions on absent father using a logistic regression, and estimate the risk of child death using a Cox proportional hazard model (Cox, 1972). In both cases I consider parents' former enslavement conditions as treatments and use two-way fixed effects to control for unobserved heterogeneity at the county of residence and year level. I consider the general specification:

$$Y_{i,m,f,c,y} = \alpha + \sum_{k=1}^{2} \beta_{m,k} T_{k,m} + \sum_{k=1}^{3} \beta_{f,k} T_{k,f} + \gamma_m X_m + \gamma_f X_f + \gamma_i Gender_i + \eta_c + \theta_y + \epsilon_{i,m,f,y,c} \quad (3.1)$$

Where $Y_{i,m,f,c,y}$ represents the outcome for the child i, born to mother m and father f, in year y, and county c. η_c are county of residence fixed effects, and θ_y

year-fixed effects. X_m, X_f are vectors of characteristics of mother m and father f (occupation, residence, age). The inclusion of these variables allows for netting out the enslavement condition effect from income or demographic characteristics. $T_{k,m}$ (resp. $T_{k,f}$) refers to the mother's (resp. fathers') coercion category, with $k = 0$ referring to the lower coercion group; $k \in \{1, 2\}$ referring to the higher and extreme coercion groups; and $k = 3$ referring to the absent fathers group. $\beta_{m,k}$ (resp. $\beta_{f,k}$) measures the effect of mothers' (resp. fathers) coercion history on $Y_{i,m,f,c,y}$, relative to the low coercion mothers (resp. fathers). Treatment categories are determined by matching parents' surnames to the Census of the Newly Freed. In all specifications, standard errors are clustered at the island and county level.

Movers' design. To net out the effect of coercion history (origins) from place-of-residence effects, I rely on the fact that more than half of formerly enslaved parents relocated from their county of enslavement(see Table 3.A1 and 3.A2). I adopt alternative specifications to assess the robustness of my results to potential selection into mobility: (i) focusing solely on individuals that moved from their former county of enslavement; (ii) excluding county fixed effects.

Exogeneity assumption. *Enslaved workers were not selected into different crop types* As per external sources, there was no apparent evidence of selection bias across crop types concerning ethnicity, gender, age composition, or price upon purchase (Fallope, 1983, 1987). I also find no significant differences in assigned values of slaves on either sugarcane or coffee plantations, suggesting little initial selection into crop types (Table 2.1). Finally, since labor conditions were more demanding on sugar plantations and sugar planters tended to be wealthier, I expect that any selection bias on enslaved workers' initial health endowments should be positive. A negative effect of enslavement conditions on child health should, therefore, be interpreted as a lower bound of the actual impact of coercion on the quality of childhood environment.

3.3 Effect of Enslavement Conditions on Families

3.3.1 Selection into Parenthood

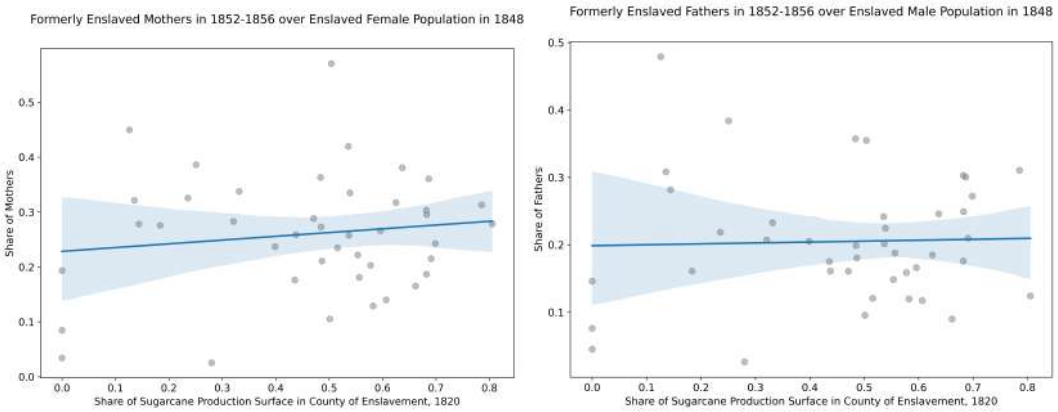
Post Abolition Mobility. Table 3.A1 and Table 3.A2 in Appendix contain descriptive statistics for parents who had a child between 1852 and 1856. Over half of the parents relocated from their county of enslavement after abolition. Across all coercion groups, movers appear to be more skilled then stayers, with a lower share of field workers among movers for both men and women (Table 3.A3), suggesting a potential positive selection of parents into mobility.

To understand where formerly enslaved individuals move, I measure the attractiveness of a county in Figure 3.A6 in Appendix, where I compare the share of emigrants to immigrants by county. Counties with average sugarcane intensity appeared to be the most appealing destinations compared to urban areas or those highly specialized in sugarcane production (with the highest sugarcane intensity). For urban areas, this could be explained by the fact that 70% of formerly enslaved individuals were field workers (as shown in Table 3.A1 and 3.A2), with therefore higher job prospects in rural areas. As opposed to the most sugarcane intensive counties, those who relied on diversified crops like subsistence agriculture, were probably less affected by the sugar crisis and the abolition of slavery - because less labor intensive. This could, in turn, attract more formerly enslaved workers after abolition.

Descriptive statistics on movers Table 3.A3 suggest a positive selection of movers on occupation, with 50% of fathers that moved being field workers against 70% for stayers. This positive selection does not appear to differ depending on coercion history.

Who becomes a parent. I find no selection into parenthood depending on former coercion exposure. Binned scatterplots relating the share of formerly enslaved individuals who became parents between 1852 and 1856, with former enslavement conditions (sugarcane intensity in county) (Figure 3.2) show a close to 0 linear relationship between the two, both for mothers and fathers. I find that both mothers and fathers are similar in terms of age in 1852, number of children between 1852 and 1856, and occupation. When regressing mothers' age at first birth (Table 3.A6) and number of children (3.A5) on their former enslavement condition, I find a null effect of former coercion intensity exposure on either fertility outcomes.

Figure 3.2: Selection into Parenthood Depending on Former Enslavement Conditions



Note: These figures are binned scatterplots relating the share of formerly enslaved individuals who became parents between 1852 and 1856, with the share of sugarcane over cultivated surface in their former county of enslavement. One dot corresponds to a unique county. The share of sugarcane reflects the intensity of coercion exposure. Each point represents a county. For a given county, the y-value is equal to the number of mothers (resp. fathers) who were enslaved in this county, divided by the total formerly enslaved female (resp. male) population aged 17-40 (resp. 17-65) enslaved in this county as per the 1848 Census. The blue line is a linear fit of the relationship between share of parents and sugarcane intensity, the shaded area represents 95 percent confidence bands.

3.3.2 Child Mortality

Parental exposure to coercion's effect on child mortality. I estimate children's risk of death using a Cox Proportional Hazard Model (CoxPH) (Cox, 1972), which is a specific class of survival models. It estimates age-specific hazard-functions (probability of failure at age t conditional on having survived at $t-dt$), while taking into account data censoring: that is, when the hazard (death) can only be observed up to a certain date, which is the case in my setting (I am able to observe children's death up to the 31st of December 1856).

The empirical specification used to estimate the effect of parents' former enslavement conditions on child mortality is the following:

$$\lambda_{i,m,f,y,c}(t) = \lambda_0(t) \cdot \exp\left(\sum_{k=1}^{2} \beta_{m,k} T_{k,m} + \sum_{k=1}^{3} \beta_{f,k} T_{k,f} + \gamma_m X_m + \gamma_f X_f + \gamma_i Gender_i + \eta_c + \theta_y + \epsilon_{i,m,f,y,c}\right) \quad (3.2)$$

Where $\lambda_{i,m,f,y,c}(t)$ is the hazard rate for individual i at time t, born in county c and year y of mother m and father f. $\lambda_0(t)$ is the baseline hazard rate, in comparison

Table 3.1: Risk of Child Death Before 1857, Depending on Coercion Exposure of Parents

	(1)	(2)	(3)	(4)
Mother: High coercion vs. Low	−0.07	−0.10	−0.11	0.09
	(0.07)	(0.08)	(0.09)	(0.15)
Mother: Extreme Coercion vs. Low	−0.03	−0.08	−0.05	−0.01
	(0.08)	(0.10)	(0.10)	(0.16)
Father: High Coercion vs. Low		−0.06	0.01	0.07
		(0.12)	(0.13)	(0.15)
Father: Extreme Coercion vs. Low		**0.28****	**0.33****	**0.33****
		(0.12)	(0.13)	(0.15)
Father: Absent vs. Low		0.03	0.04	
		(0.09)	(0.10)	
Island, County, Year FE	Yes	Yes	Yes	Yes
Mother controls and gender	No	No	Yes	Yes
Father controls	No	No	No	Yes
Mean	14.11	14.08	13.96	14.15
Observations	17,274	12,801	11,019	4,441

Notes: This table displays the Cox proportional estimated effect of parental exposure to coercion on children's risk of death before 1857 for all children born between 1852 and 1856. For λ a coefficient, the hazard ratio associated with a one-unit increase in the corresponding predictor variable increases by a factor of exp(λ), or (exp(λ)-1)% .***p $< .01$; **p $< .05$; *p $< .1$

to which the coefficients of the model are estimated in relative terms[12]. Variables included in the model are the same as the general specification 3.1: $\beta_{m,k}$ and $\beta_{k,k}$ represent the effects of mothers and fathers' coercion category respective to the lower category, on the hazard ratio.

Baseline estimates are presented Table 3.1. I find that mother's differential exposure to coercion has a null effect on children's odd-risk of death, which holds when controlling for age, gender of the birth, occupation and paternal exposure to coercion. I find, however, that the presence of *extremely coerced fathers* increases the probability of experiencing a child death before 5 years old by 35% respective to lower coercion fathers[13]. This is equivalent to a 5 percentage points (pp) increase

[12]The baseline hazard represents the hazard at time t for a reference group (e.g., individuals with all covariates set to zero). It is not directly estimated by the model, but represents the reference point against which all other hazard rates are compared. It allows to evaluate the effect of covariates on the relative hazard without having to specify a parametric form for the baseline hazard.

[13]For an estimated coefficient of 0.28, the effect of paternal exposure to extreme coercion corresponds to an increase of $exp(0.28) - 1$ in the probability of child death.

relative to the average child mortality (14%). This result is robust to controlling for parents' characteristics (age, occupation), excluding county fixed effects (Table 3.A9), focusing on movers only (Table 3.A10), or using different thresholds for the *extreme coercion* treatment category (Table 3.A11 and 3.A12).

Alternative Non-Parametric Estimation. An important assumption of the CoxPH model is that the effect of each covariate on the risk of death should be constant at each age. While children's health has been shown to be primarily determined by mothers' inputs in the first months of life (Lee, 2005; Zhao et al., 2017) broader inputs (environmental, paternal) start having a more prominent role when children start walking or stop being breastfed. To account for the possibility that paternal inputs might not have a constant effect overtime, and to visualise changes in survival probabilities by month of age, I use the non-parametric Kaplan-Meier probability function[14]. I estimate the survival probability separately for groups defined by parents' coercion exposure, while accounting for county clustering. Graphical visualizations are presented in Figure 3.3. Visuals show that children with more coerced mothers exhibit a lower, although statistically non-significant, survival probability. When turning to fathers, I find that children whose father was extremely coerced have a significantly lower survival probability relative to less coerced father.

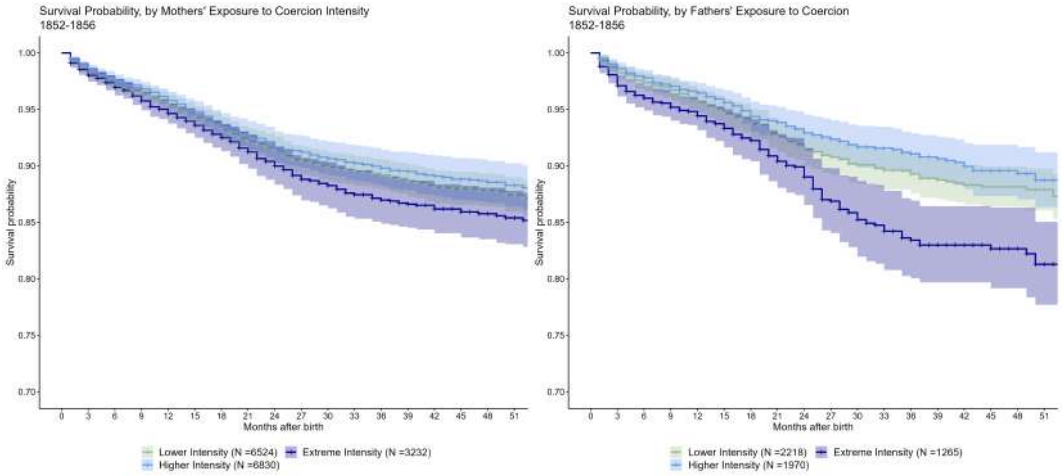
3.3.3 Family Structure

Absent Fathers I follow the general empirical specification detailed in Equation 3.1 and estimate the probability of fathers being absent from birth records using a logit specification[15]. Baseline results using the sample of all mothers are displayed in the first three specifications of Table 3.A8. I find a positive and significant effect (+20%) of worse enslavement conditions on fathers' presence. When turning to the alternative specification which focuses on mothers that moved from their county of enslavement, I find that the effect of maternal coercion is no longer significant and much smaller in magnitude (last three specifications in Table 3.A8). This suggests

[14]The Kaplan-Meier function calculates the product of the conditional probabilities of survival up to time t for all distinct event times t_i that are less than or equal to t. $S(t) = \prod_{i:t_i \leq t} \left(1 - \frac{d_i}{n_i}\right)$ Where: $S(t)$ is the survival probability at time t; t_i represents the distinct event times observed in the dataset; d_i is the number of events (e.g., deaths) that occur at time t_i, n_i is the number of individuals or subjects at risk of experiencing an event at time t_i. I estimate the Kaplan Meier estimator for each value of mothers' coercion exposure, and fathers' coercion exposure, separately.

[15]I use the general specification for binary outcomes, which is a generalized least squares model with binomial family and account for clustered standard errors at the island, county and year levels

Figure 3.3: Monthly-Survival Probability of Children, Depending on Mothers' (left) and Fathers' (right) Former Coercion Exposure



(a) Mothers (b) Fathers

Notes: These figures display the Kaplan-Meier non-parametric survival probabilities of children born between 1852 and 1856, which are estimated separately depending on the coercion history of the mothers (left) and the fathers (right). Shaded areas represent robust 95 percent confidence bands.

that geographical factors, more than coercion history *per se*, are likely to be driving the measured differences in fathers' presence (Miller, 2018).

Assortative Matching. I then focus on children whose both parents are known and estimate the effect of mothers' past enslavement conditions on the probability to match with a partner exposed to the same coercion intensity. I find that extremely coerced mothers have a 21% lower chance of partnering with a father exposed to the same enslavement conditions, even after controlling for county and year population composition in terms of coercion groups[16], or excluding county fixed-effects (Table 3.A16). This suggests that these partnership patterns cannot be accounted for by composition effects in the county of residence. Results might in fact reflect partnership decisions and a lower attractivity of extremely coerced men for women that had also been exposed to the worst conditions.

[16]Total number of mothers and fathers of each coercion group, in the county of birth and the year of birth.

3.4 Mechanisms

I now delve into the potential underlying mechanisms contributing to elevated child mortality within households where fathers experienced extreme coercion.

3.4.1 Health

The first possible explanation would be that fathers previously exposed to extreme coercion might be in worse physical health - which could translate into poorer economic resources and cause higher child mortality. Since extreme coercion affected both men and women (Section 2.4), I would also expect to see an effect of mothers' exposure to extreme coercion on child mortality[17] - which is not what I observe. Finally, the effect of paternal exposure to extreme coercion on child mortality is observed relatively to families with *absent* fathers. This suggests that extremely coerced fathers have a strong *deteriorating effect* on child mortality, rather than limited economic inputs.

3.4.2 Income

I further explore the role of economic inputs using information on parents' occupation. Controlling for the occupation of either parent maintains the value and significance of paternal coercion history's impact on child mortality (Table 3.1, last two specifications). This implies that the adverse effect of fathers' exposure to extreme coercion cannot be fully explained by the occupational composition of this group. To delve deeper into the potential influence of income on child mortality, I narrow my focus to families where the father holds a higher-income occupation (non-field workers, constituting 30% of all fathers). The results are outlined in Table 3.2, columns (1) and (2). Notably, the positive effect of paternal exposure to extreme coercion on child mortality persists, showing a magnitude similar to baseline estimates. Furthermore, analyses exploring heterogeneity in the additional impact of paternal exposure to extreme coercion reveal no discernible differences based on whether parents were field workers or not (refer to Figure 3.A7). These additional analyses suggest that income, through occupation, is not driving the large effect of paternal exposure to coercion on children's risk of death.

[17] The vast majority of women worked

3.4.3 Spatial Sorting

I now explore whether spatial sorting of families into high mortality areas could explain paternal exposure to extreme coercion's effect on child mortality. Previous results shown in Table 3.1 were robust (and very similar in magnitude) to focusing on movers or excluding county fixed effects (Tables 3.A10 and 3.A9), suggesting that place-of-residence specific factors are unlikely to be the main driver of my results. I further study whether spatial sorting could explain part of the measured effect, by focusing solely on families residing in low mortality counties - which I identify as those whose estimated fixed effect on chid mortality in the main specification (Table 3.1, column (2)) is negative. Results are displayed Table 3.2, with or without controlling for parental coercion history. I still find a positive, although no longer significant effect, of paternal exposure to extreme coercion on child mortality. A final interesting result lies in the fact that sugarcane intensity in the county of residence is *negatively* correlated with child mortality, suggesting that the more coercive counties during slavery might have also been the less disease prone environment (Table 3.A13).

3.4.4 Partnership

Maternal attributes. Another potential explanation could be that extremely coerced men systematically match with women that have less resources to care for children (either economic, social or emotional). Heterogeneity analyses display no differential impact of paternal exposure to extreme coercion depending on mothers' own coercion history (see Figure 3.A7), or when controlling for their sociodemographic characteristics. This suggests that the effect of paternal exposure on child mortality stems from fathers' themselves[18].

Union Stability. I explore the possibility that the adverse impact of extreme paternal coercion on child health could be linked to a removal of fathers' inputs into children's environment, either because of death or separation with the mother during children's early years of life. To do so, I focus on women who had several

[18] Another possibility could be that women might be systematically more vulnerable on unobserved characteristics. If this were solely driven by mothers, I would anticipate differences in infant survival across coercion groups to show from birth, when child survival primarily depends on maternal factors. However, visuals displayed Figure 3.3 show that the effect of extreme paternal coercion becomes more evident starting at the age of two and beyond, when children are exposed to a broader range of environmental factors, including potential interactions with their fathers.

Table 3.2: Risk of Child Death Before 1857, Robustness Checks on Subsamples

	Father non Field Worker		Low Mortality Areas	
	(1)	(2)	(3)	(4)
Father: High Coercion vs. Lower	−0.01	0.12	0.05	0.21
	(0.23)	(0.25)	(0.29)	(0.31)
Father: Extreme Coercion vs. Lower	0.42**	0.50***	0.34	0.40
	(0.22)	(0.24)	(0.34)	(0.38)
County and Year FE	Yes	Yes	Yes	Yes
Mother controls	No	Yes	No	Yes
Mean	14.46	14.32	14.4	14.41
Observations	1,611	1,322	1,946	1,640

Notes: This table displays the Cox proportional estimated effect of paternal exposure to coercion on children's risk of death on two subsamples of children: those whose father is known and not a field worker; those born in a low mortality county. before 1857, for children born in low mortality counties, which are identified as counties whose estimated fixed effect on child mortality in the main specification is negative (Anse bertrand, Le Carbet, Commune du nord, Le Marin, Le Moule, Petit canal, Le Prêcheur, Le Robert, Saint esprit, Sainte anne, Sainte luce, Sainte marie, Trinite, Trois rivieres, Le Vauclin, Vieux habitants). For λ a coefficient, the hazard ratio associated with a one-unit increase in the corresponding predictor variable increases by a factor of exp(λ), or (exp(λ)-1)% .***p $< .01$; **p $< .05$; *p $< .1$.

children between 1852 and 1856, and where the father of the first child is known (N = 643). I use this subsample to estimate the probability that fathers of subsequent children would be different than the first one or unknown, depending on the coercion exposure of the first father, and that of the mother. Baseline estimates are presented in appendix, Table 3.A14. I find that extreme coercion exposure of the first father has a large but non significant effect on the probability that next fathers would be different or missing, relative to less coerced fathers. This could reflect the fact that extremely coerced fathers tend to disappear more, which could cause an income (or emotional) shock that might partially contribute to the deterioration of the quality of childhood environment. Since only 8% of the mothers with several children have them with different fathers, this explanation can only provide a partial explanation for my results[19].

These findings indicate that the significant negative effect of the presence of extremely coerced fathers on child health cannot be accounted for by income, geographical location, maternal attributes, or fathers' disappearance. This suggests that extreme coercion exerts a discernible influence on what fathers pass on to their children. This leads to the crucial question of identifying the specific factors that contribute to this adverse influence.

3.4.5 More Coerced Men Could Be More Violent

I argue that fathers' limited contributions to child care or inattention to children are not the primary drivers of fathers' effects on child mortality. If this were the case, I would expect the impact on child health to be similar to that of absent fathers. However, children's survival probability is significantly lower when fathers are extremely coerced, indicating that these fathers actively worsen the child's environment. A possible explanation for the large detrimental effect of paternal exposure to extreme coercion could be that extremely coerced men are more prone to violence. I expect this mechanism to have large effects on child mortality, as previously shown in various contexts, whether developed or developing (Aizer, 2011; Bhalotra and Rawlings, 2011; Currie et al., 2022)[20]. To test this hypothesis, I ex-

[19] I also note that mothers' coercion exposure has a positive effect on the probability of fathers to be different, although extremely coerced mother actually appear *less* likely to have children with several fathers relative to high coercion mothers.

[20] Previous estimates of the impact of intimate partner violence on child mortality are large in magnitude, but vary greatly across studies. In their study on 32 developing counties, Rawlings and

amine descriptive evidence on the types of crime committed by prisoners who were formerly enslaved, and were sent to the penal colony of Guyana between 1850 and 1890 (400 observations). I plot the share of convicted prisoners per 1000 formerly enslaved individuals[21] according to the sugarcane intensity in their former county of enslavement in Figure 3.A8 for all crimes in Appendix and Figure 3.4 for violent crimes. Interestingly, while the relationship between overall crime rates and harsh enslavement conditions is negative, it becomes positive when I focus on violent crimes, and more specifically sexual crimes. I do a similar exercise using the list of all 100,000 French transported criminals between 1850 and 1900, and plot the share of surnames found among transported individuals, by county of origin of the surname in Figure 3.5. I also find a positive association between a surname' coercion history, and its prevalence among transported criminals. Although both analyses are descriptive and do not allow for a causal link, this descriptive evidence suggests that worse enslavement conditions might be at the origin of persistent violent behavior.

Figure 3.4: Prevalence of Formerly Enslaved Prisoners Sent to Guyana 1850-1890, by type of Crime and Former County of Enslavement

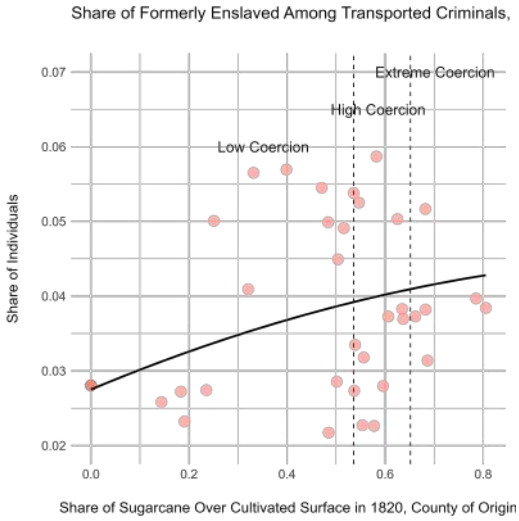
(a) Violent Crimes (b) Sexual Crimes

Notes: These figures are binned scatterplot with one bin corresponding to a former county of enslavement. They relate the number of prisoners sent to Guyana between 1850 and 1890 per 1,000 formerly enslaved individuals, related to the sugarcane intensity in the former county of enslavement. The first graph displays the number of prisoners per 1,000 inhabitants convicted for violent crimes (murder, sexual assault, physical assault); the second graph, for sexual crimes only. Black lines are linear fits between the two variables, weighted by the enslaved population in 1848.

Siddique find that mothers' experience of intimate partner violence at any point in the year following the birth is associated with a 0.7pp and 1pp increase in the probability for children to die before 2 years old, and 5 years old respectively - representing a 10 to 12% increase in the mean. In rural india, researchers (Koenig et al., 2010) found that mothers' recent experiences of domestic abuse were associated with a +68% increase in infant mortality.

[21] I use the population data from the 1848 Census to calculate crime rates by former place of enslavement

Figure 3.5: Share of surnames found among convicted criminals between 1850 and 1900, by surnames' county of origin



Notes: This figure displays the share of formerly enslaved surnames found among all french transported criminals between 1850 and 1950, by the county of origin of the surname. The line is a second order polynomial fit between the two variables

3.5 Discussion

3.5.1 Gender Differences in Response to Violence

I now explore why men seem to react more to former exposure to violence in their effect on children, relatively to women. One possible explanation lies in gender-differences in the consequences of parental trauma on children. Fathers have been shown to face greater challenges in shielding their parenting practices from trauma compared to mothers in studies on Holocaust survivors (Levav et al., 1998; Kestenberg, 1980; Harkness, 1993). Given that men can also react violently to stress (Card and Dahl, 2011), gender-differences in the expression of trauma could give rise to the sizable deterioration of children's health when more traumatized fathers are around. It is also possible that men might have been *more* exposed to violence when the coercion gap increased prior to abolition. While both genders undertook physically demanding tasks (Davis, 1983), the strongest workers, often assigned to strenuous work like cane harvesting, may have faced more violence, potentially explaining part of the observed gender differences in response to violence during slavery. Overall, my findings could reflect a mechanism of intergenerational transmission of trauma through fathers.

A complementary interpretation is that the most violent places shaped more unequal gender norms, favoring male-based violence in the familiar setting (Abramsky et al., 2011). A first hypothesis is that the sexual exploitation of women by white men, which was likely linked to the overall violence exerted against enslaved workers, might have shaped persistent norms valuing male domination over women even among formerly enslaved men after abolition (Mulot, 2013). Another channel through which worse enslavement condition might have affected gender norms is through its effect on masculine identity, leading to possible backlash after abolition (Akerlof and Kranton, 2000; Abramsky et al., 2011). Heterogeneity analysis on paternal exposure to extreme violence by the child's gender suggests a stronger - although non significantly - effect on girls compared to boys (see Figure 3.A7), possibly reflecting more gender-unequal norms.

3.5.2 Relative Effect of Coercion, Income and Geography

Previous regression analyses have consistently shown that paternal exposure to coercion has a negative impact on child health, even when we include occupation controls for both mothers and fathers. To put these results into perspective, I now examine the specific effect of occupation to compare it with the effect of parents' coercion history. Table 3.A15 in Appendix presents regression results of the survival model 3.2 with the specific effect of parents' occupation on risk of death. I find that mothers' occupation has minimal impact on child mortality, while lower income status of fathers (e.g., field workers) is positively associated with children's risk of death, although this relationship is not statistically significant. Notably, the magnitude of this effect (-0.12) is smaller than the estimated effect of previous regressions (0.30), suggesting a larger contribution of coercion exposure relative to income.

A similar exercise can be done regarding geographical characteristics. I find that sugarcane intensity in the county of *residence* is negatively associated child mortality, with an estimated coefficient of -0.5 (against 0.3 for paternal exposure to extreme coercion), suggesting a larger effect of residence related factors compared to coercion history (Table 3.A13).

3.6 Conclusion

This chapter provides the first empirical assessment of enslavement conditions' impact on post abolition family structure and children's health. The unique setting of this study, and novel datasets that I compile, allow me to precisely link formerly enslaved individuals to their enslavement conditions even if they relocate after abolition, and to effectively disentangle coercion's effect from place of residence or income-related factors.

After slavery ended, I find that the presence of fathers exposed to the worst enslavement conditions was associated with large adverse effects on child health. I find suggestive evidence that this could be driven by worse enslavement conditions leading to more violent men, to which mothers may have responded through strategic matching decisions with less coerced partners. As child mortality is not only an extreme marker of poor childhood environment but can also have significant negative spillover effects on family members, this results reflects substantial inequality among children of formerly enslaved parents. While prior research highlighted the importance of persistent extractive institutions and changes in norms to explain slavery-induced economic inequality, this chapter underscores the importance of a relatively less-studied mechanism: the intergenerational transmission of violence and child neglect.

Another noteworthy finding of this study is that, while single-mother households are common after abolition, they did not constitute the predominant family structure following abolition. My research also points to a potential positive effect of worse enslavement conditions on fathers' presence, which seems to be related to place-of-residence specific factors rather than coercion history *per se*. This finding calls for a nuanced approach to the previously made link between violence during slavery and absentee fathers.

This research has limitations which I aim to address as part of a broader research agenda on the social consequences of slavery. First, my analysis is limited to immediate descendants of the formerly enslaved. I am currently digitizing the complete set of civil records from 1848 to 1905 to gain a dynamic perspective on the importance of family-level transmission relative to place of residence determinants, in driving post abolition outcomes. Second, my examination of post-abolition inequality is restricted to formerly enslaved families. One of my upcoming projects will focus on identifying white families' surnames from pre-1848 civil records and studying interracial marriages, which was a key access point to capital after slavery

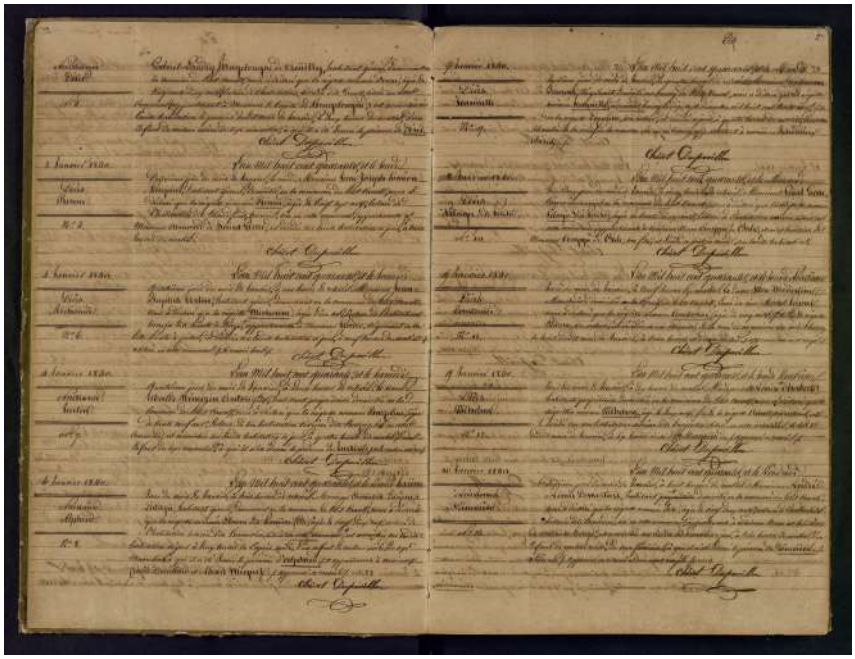
ended.

3.A Appendices, Chapter 3

3.A.1 Archives

Data sources

Figure 3.A1: Example of census of the Newly Freed record

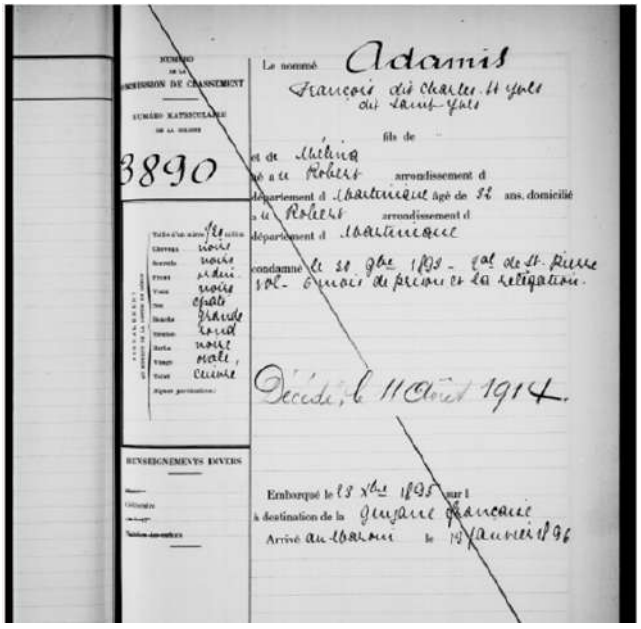


Source: French Overseas National Archives, *Registres des Nouveaux Libres*

Figure 3.A2: Example of post abolition civil registries

Source: French Overseas National Archives, *Registres d'Etat civil*

Figure 3.A3: Example of convicts' record



Source: French Overseas National Archives, *Registres des bagnards*

Matching procedure

For each birth record, I matched the surnames of both mothers and fathers with the Census of the Newly Freed to determine parents' former place of enslavement. To achieve this, I first isolated the surname from the parents' full names. In 20% of cases, the parents' full names only included common names (e.g., Marie Louise Elizabeth), making it difficult to pinpoint the surname. To ensure accurate matches, I excluded most common names from parents' full names using official statistics on names given at birth in the French West Indies between 1870 and 1910 (French National Statistics Institute, INSEE).

Next, I matched the remaining components with the surname database of the Census of the Newly Freed, employing Fuzzy matching techniques. I used the Levenshtein distance which quantifies the minimum number of single-character edits needed to transform one string into another. I considered as a potential match surnames where the similarity exceeded 90%, meaning that the share of single-character edits over the string's length did not surpass 10%. In cases of multiple matches for a given name, I selected the one with the highest similarity ratio. If several matches still existed, I chose the one who originated from the county closest to the parents' current residence. In total, I was able to match 50% of the births to an enslaved mother in the census. Considering that 30% of non-matched births likely originate from non-formerly enslaved mothers[22], this represents a matching rate of 80%.

The remainder of unmatched surnames could stem from spelling disparities between census and civil records that exceeded the 10% matching threshold, along with incomplete records in the Census of the Newly Freed. Given the impossibility to identify unmatched parents as either non-former slaves, unregistered slaves, or former slaves with incorrect surnames, I exclude these observations from my analyses.

[22] 30% of the total population were not enslaved in 1848

3.A.2 Additional Descriptive Statistics After Abolition

Table 3.A1: Descriptive Statistics on Mothers Who Had a Child between 1852 and 1856, According to Former Coercion Exposure

	Former Exposure to Coercion		
	Low	High	Extreme
N	7,325	7,450	3,421
Birth in plantations	0.55	0.60	0.63
Child deaths	0.10	0.08	0.11
Mother, Age	27.53	27.26	26.97
Mother, Moved	0.53	0.57	0.53
Mother, number of births	1.39	1.36	1.39
Mother, Field worker	0.71	0.70	0.74
Mother, Domestic worker	0.07	0.06	0.06
Mother, Trader	0.10	0.12	0.08
Mother, Skilled worker	0.02	0.02	0.01
Mother, No occupation	0.10	0.11	0.10
Absent Father	0.42	0.46	0.33
Characteristics of fathers with whom mothers had children			
Father, Age	34.40	34.58	34.42
Father, Moved	0.50	0.55	0.48
Father, Lower Coercion	0.27	0.10	0.15
Father, High Coercion	0.11	0.24	0.10
Father, Extreme Coercion	0.07	0.04	0.30
Father, *Unknown status**	0.13	0.16	0.12
Father, Field Worker	0.64	0.59	0.63
Father, Skilled worker	0.17	0.19	0.19
Father, Trader	0.04	0.05	0.04
Father, Employee	0.02	0.02	0.02
Father, Landlord	0.11	0.13	0.10

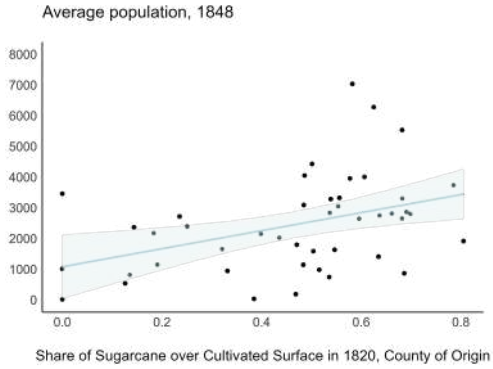
Notes: Average statistics computed on the sample of formerly enslaved mothers who had a child between 1852 and 1856. * Fathers of unknown status are those whose former enslavement status could not be determined based on their surnames: either because they were not slaves, or because they were slaves whose surname could not be linked to a former place of enslavement.

Table 3.A2: Descriptive Statistics on Births with Known Fathers, Depending on Fathers' Exposure to Coercion, 1852-1856

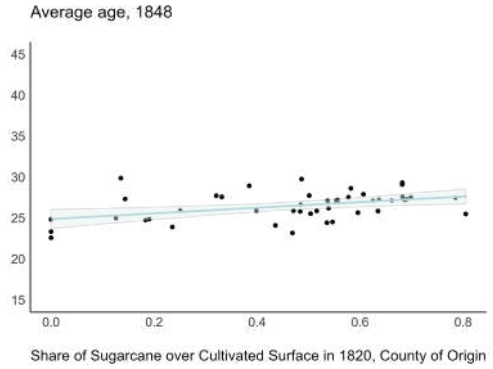
	Former Exposure to Coercion		
	Low	High	Extreme
N	2,398	2,159	1,349
Birth in plantations	0.58	0.62	0.71
Child deaths	0.10	0.08	0.14
Father, Age	33.81	34.11	33.99
Father, Moved	0.50	0.56	0.45
Father, Field Worker	0.71	0.69	0.69
Father, Skilled worker	0.14	0.15	0.19
Father, Trader	0.03	0.04	0.03
Father, Employee	0.01	0.01	0.01
Characteristics of mothers with whom known fathers have children			
Mother, Lower Coercion Intensity	0.62	0.28	0.27
Mother, High Coercion Intensity	0.22	0.61	0.18
Mother, Extreme Coercion Intensity	0.16	0.11	0.55
Mother, Age	28.19	28.14	27.44
Mother, Moved	0.48	0.53	0.44
Mother, Number of births	1.43	1.41	1.42
Mother, Field worker	0.79	0.77	0.80
Mother, Domestic worker	0.04	0.03	0.04
Mother, Trader	0.06	0.07	0.07
Mother, Skilled worker	0.02	0.01	0.01
Mother, No occupation	0.10	0.12	0.08

Notes: Average statistics computed on the sample of formerly enslaved fathers who had a child between 1852 and 1856.

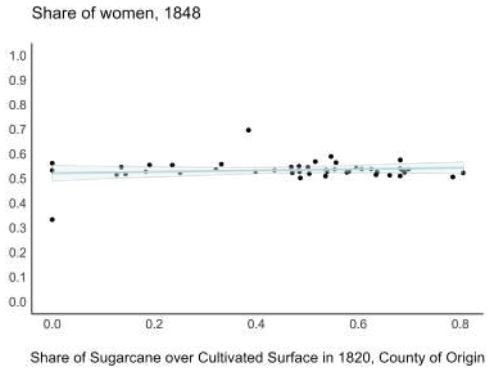
Figure 3.A4: Descriptive Statistics on the Newly Freed Population in 1848, by Share of Sugarcane over Cultivated Surface in Former County of Enslavement



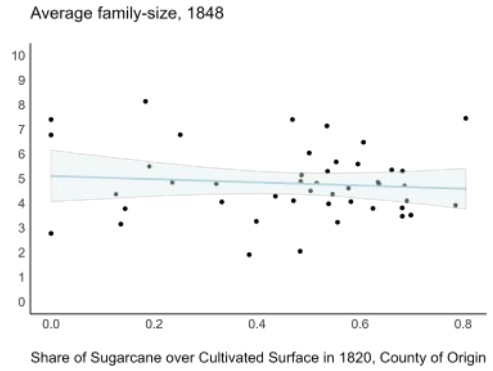
(a) Average population



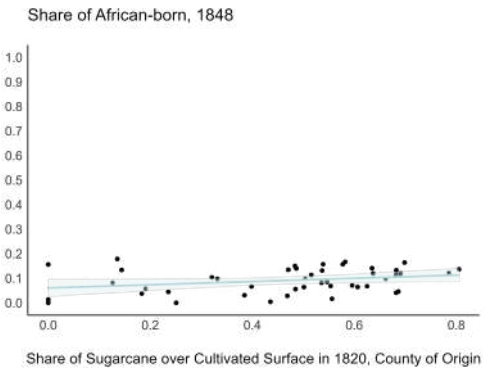
(b) Average age



(c) Share of women



(d) Average family size

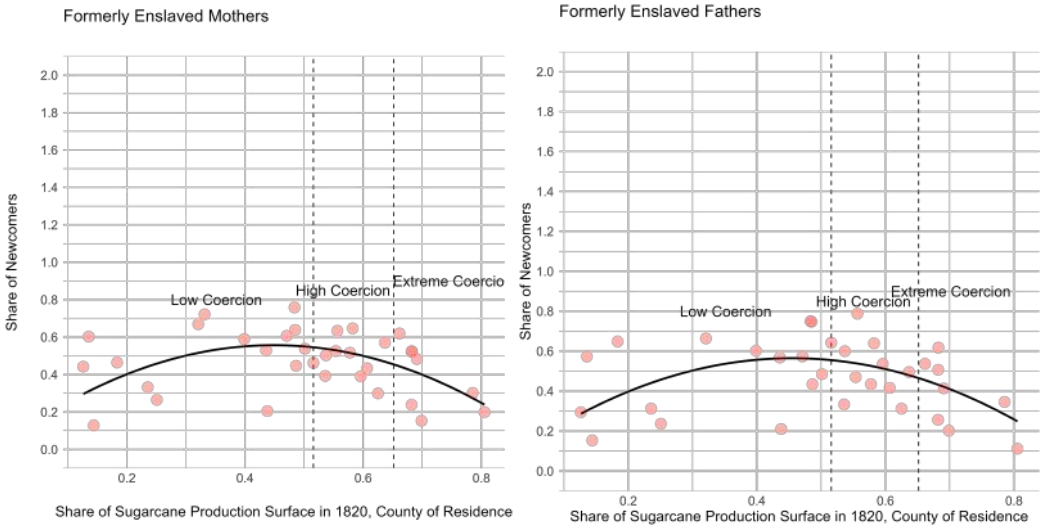


/2/2

(e) Share of African born

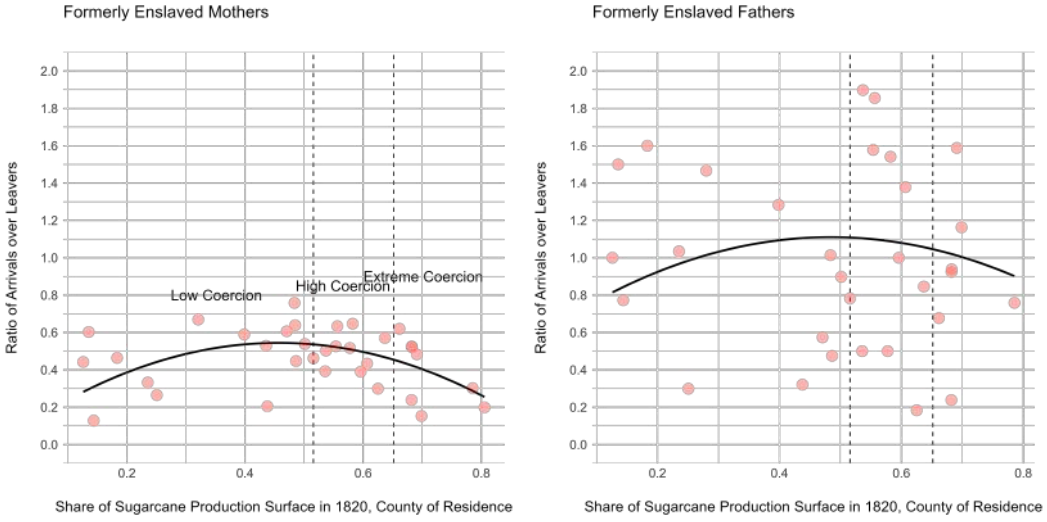
Note: The vertical doted lines delimit the enslavement conditions categories (lower, high, extreme)

Figure 3.A5: Share of Emigrants by Sugarcane Intensity in Receiving County



Note: These figures are binned scatterplot relating the share of parents in the county of residence that *emigrated* to this county, relative to the sugarcane intensity in the county of residence (share of sugarcane surface over total cultivated surface in 1820). One dot corresponds to a unique county of residence

Figure 3.A6: Ratio of Emigrants over Immigrants, by County of Residence: Mothers (left), Fathers (right)



Note: The sugarcane intensity in the residence county is the share of sugarcane surface over total cultivated surface in 1820.

Table 3.A3: Descriptive Statistics on Parents, according to Mobility and Coercion History

Coercion	Lower		High		Extreme	
	stayer	mover	stayer	mover	stayer	mover
Mothers						
N	3,299	3,727	2,944	4,038	1,584	1,790
Birth in plantations	0.61	0.49	0.72	0.49	0.79	0.49
Child deaths	0.10	0.10	0.09	0.09	0.12	0.10
Mother, age	27.57	27.54	27.28	27.16	26.61	27.34
Mother, number of births	1.53	1.26	1.44	1.30	1.50	1.30
Mother, Field worker	0.83	0.59	0.87	0.56	0.85	0.64
Mother, Domestic worker	0.05	0.08	0.04	0.07	0.05	0.07
Mother, Trader	0.06	0.14	0.06	0.18	0.06	0.10
Mother, Artisan/Skilled	0.02	0.03	0.01	0.02	0.003	0.02
Mother, No occupation	0.03	0.16	0.03	0.18	0.03	0.16
Unknown fathers	0.44	0.39	0.50	0.43	0.30	0.36
Fathers						
N	1,173	1,157	840	1,156	729	587
Birth in plantations	0.61	0.57	0.74	0.53	0.83	0.54
Child deaths	0.11	0.08	0.09	0.07	0.15	0.12
Father, age	33.51	34.06	33.44	34.55	34.00	34.07
Father, Field Worker	0.81	0.59	0.82	0.59	0.76	0.58
Father, Skilled worker	0.09	0.20	0.11	0.19	0.19	0.19
Father, Trader	0.02	0.05	0.02	0.05	0.01	0.06
Father, Employee	0.004	0.02	0.004	0.02	0.003	0.02

Notes: Average statistics computed on the sample of formerly enslaved parents, depending on their former coercion history and whether they moved from their county of enslavement. Descriptive statistics suggest that movers tend to be more skilled across all coercion group, with a lower share of field workers relatively to stayers who had a child between 1852 and 1856.

Table 3.A4: Distribution of Newborns in 1852-1856, According to the Status of their Father, By County of Residence

	Absent Fathers	Lower Coercion	High Coercion	Extreme Coercion
abymes	47.75	36.04	7.81	8.41
anse bertrand	33.49	47.27	5.46	13.78
capesterre belle eau	30.72	9.80	54.03	5.45
capesterre de mg	38.67	40.27	12.81	8.24
carbet	72.40	4.66	18.64	4.30
case pilote	80.32	1.06	16.49	2.13
commune des trois bourgs	70.80	5.40	21.60	2.20
commune du nord	53.35	4.30	38.81	3.54
commune du sud	71.35	21.15	4.62	2.88
ducos	54.37	33.24	9.58	2.82
françois	50.62	7.84	36.82	4.73
gosier	38.75	54.79	4.17	2.29
gourbeyre	30.84	31.31	18.22	19.63
grand bourg de mg	38.10	27.62	23.97	10.32
gros morne	62.76	23.98	9.69	3.57
lamentin	31.30	8.04	1.30	59.35
le lamentin	71.65	3.23	23.91	1.21
lorrain	64.93	13.72	14.60	6.75
marin	71.06	14.47	9.50	4.97
morne a l'eau	57.88	24.62	7.34	10.15
moule	37.11	20.88	34.73	7.28
petit canal	24.59	17.82	14.03	43.56
pointe a pitre	69.02	18.90	4.35	7.72
pointe noire	40.12	45.51	3.59	10.78
port louis	16.67	17.25	5.43	60.66
precheur	74.57	12.83	9.13	3.48
riviere pilote	67.38	15.41	11.97	5.25
robert	53.38	8.77	33.85	4.00
saint claude	25.83	63.74	6.87	3.55
saint esprit	72.45	12.69	13.36	1.50
saint francois	65.22	7.44	22.66	4.67
saint louis de mg	31.65	46.28	11.17	10.90
saint pierre	73.46	7.20	16.65	2.69
sainte anne	53.28	30.46	7.04	9.22
sainte luce	75.58	13.95	9.30	1.16
sainte marie	66.89	5.20	25.44	2.46
sainte rose	35.67	12.53	6.37	45.44
trinite	65.02	7.47	14.94	12.56
trois rivieres	36.97	29.70	20.61	12.73
vauclin	69.14	5.71	22.86	2.29
vieux habitants	54.62	29.23	11.92	4.23

3.A.3 Additional Results

Table 3.A5: Impact of Mothers' Enslavement Conditions on Number of Children

	(1)	(2)	(3)	(4)
High Coercion vs. Lower	−0.01	−0.01	0.004	0.001
	(0.01)	(0.01)	(0.01)	(0.01)
Extreme Coercion vs. Lower	0.02**	0.02*	0.002	0.003
	(0.01)	(0.01)	(0.01)	(0.01)
Year FE	Yes	Yes	Yes	Yes
County FE	No	No	Yes	Yes
Gender	No	Yes	Yes	Yes
Mother controls	No	No	No	Yes
Mean	1.21	1.21	1.21	1.23
Observations	14,618	14,618	14,618	12,827
Akaike information criterion	17,488.82	17,208.46	16,485.42	15,230.30

Notes: This table displays the OLS estimated effect of mothers' exposure to coercion on the number of children she has between 1852 and 1856. For λ a coefficient, an increase in one unit of a covariate is associated with an increase in λ unit in the outcome, everything else being equal. $^{***}p < .01$; $^{**}p < .05$; $^{*}p < .1$.

Table 3.A6: Impact of Mothers' Enslavement Conditions on Age at First Birth

	(1)	(2)	(3)	(4)
High Coercion vs. Lower	-0.34^{**}	-0.34^{**}	-0.27^{*}	-0.25
	(0.14)	(0.14)	(0.16)	(0.17)
Extreme Coercion vs. Lower	-0.55^{***}	-0.54^{***}	-0.09	-0.17
	(0.18)	(0.18)	(0.20)	(0.21)
Year FE	Yes	Yes	Yes	Yes
County FE	No	No	Yes	Yes
Gender	No	Yes	Yes	Yes
Mother controls	No	No	No	Yes
Mean	27.15	27.15	27.15	27.12
Observations	13,762	13,762	13,762	12,827
Akaike information criterion	94,165.38	94,162.44	94,046.31	87,570.14

Notes: This table displays the OLS estimated effect of mothers' exposure to coercion on her age at first birth, between 1852 and 1856. For λ a coefficient, an increase in one unit of a covariate is associated with an increase in λ unit in the outcome, everything else being equal. Differences in the number of observations relative to Table 3.A5 stems from the fact that some mothers have missing information on age. $^{***}p < .01$; $^{**}p < .05$; $^{*}p < .1$.

Table 3.A7: Probability for Parents to have Moved from Enslavement county, 1852-1856

	Mothers		Fathers	
	(1)	(2)	(3)	(4)
High Coercion vs. Lower	0.10^{**}	0.05	0.24^{***}	0.27^{***}
	(0.04)	(0.04)	(0.06)	(0.07)
Extreme Coercion vs. Lower	-0.02	-0.03	-0.24^{***}	-0.23^{***}
	(0.05)	(0.05)	(0.07)	(0.07)
County FE	No	No	No	No
Year FE	Yes	Yes	Yes	Yes
Parent controls	No	Yes	No	Yes
Mean	0.57	0.55	0.51	0.5
Observations	14,618	12,827	5,646	5,277
Akaike information criterion	19,891.70	16,415.88	7,747.68	6,949.98

Notes: This table displays the logit estimated effect of parental exposure to coercion on the probability to have emigrated from their county of enslavement. For these regressions, I do not include county of residence fixed effects, which is endogenous to mobility decision. For λ a coefficient, an increase in one unit of a covariate is associated in the odds of moving of $\exp(\lambda)$ times. $^{***}p < .01$; $^{**}p < .05$; $^{*}p < .1$

Table 3.A8: Probability of Absent Father in Birth Record Between 1852 and 1856

	Baseline Estimates *Full Sample*			Alternative Estimates *Mothers Moved*		
	(1)	(2)	(3)	(4)	(5)	(6)
High coercion vs. Low	−0.07	−0.07*	−0.05	−0.04	−0.05	−0.03
	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.06)
Extreme Coercion vs. Low	−0.12**	−0.13**	−0.09	−0.03	−0.05	−0.04
	(0.05)	(0.06)	(0.06)	(0.06)	(0.07)	(0.08)
Mother's age		−0.13***	−0.14***		−0.14***	−0.19***
		(0.03)	(0.04)		(0.05)	(0.05)
Male birth		−0.03***	−0.03***		−0.03***	−0.03***
		(0.003)	(0.003)		(0.004)	(0.004)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Mother controls	No	No	Yes	No	No	Yes
Mean	0.42	0.43	0.44	0.41	0.42	0.43
Observations	18,196	16,508	15,337	9,703	8,609	7,876
AIC	22,736	20,368	17,946	12,431	10,918	9,283

Notes: This table shows the logit estimated effect of mother's former enslavement conditions (taking as reference group those exposed to the lowest coercion intensity) on the probability that no father would be mentioned in the birth record of her child. The omission of paternal information implies that, legally, the child lacks a recognized father. Baseline estimates are calculated over the full sample of newborns (specifications 1 to 3). Alternative estimates are calculated over the restricted sample of newborns whose mother moved from their county of enslavement. For λ a coefficient, a one-unit increase in the covariate increases the odd of having an absent father by exp(λ)times. $^{***}p < .01$; $^{**}p < .05$; $^{*}p < .1$

Table 3.A9: Monthly Risk of Child Death Before 1857, Without County Fixed-Effects

	Without County Fixed Effects			
	(1)	(2)	(3)	(4)
Mother: High coercion vs. Low	−0.14	−0.10	−0.12	−0.02
	(0.06)	(0.07)	(0.07)	(0.13)
Mother: Extreme Coercion vs. Low	0.09	−0.01	−0.01	0.09
	(0.07)	(0.08)	(0.09)	(0.13)
Father: High Coercion vs. Low		−0.18	−0.15	−0.14
		(0.11)	(0.12)	(0.13)
Father: Extreme Coercion vs. Low		0.36***	0.39***	0.38***
		(0.11)	(0.12)	(0.12)
Father: Absent vs. Low		−0.01	−0.04	
		(0.08)	(0.09)	
Island FE	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Mother controls	No	No	Yes	Yes
Father controls	No	No	No	Yes
Mean	14.11	14.08	13.96	13.8
Observations	17,274	12,801	11,019	4,441

Notes: This table displays the Cox proportional hazard estimated effect of parental exposure to coercion on children's risk of death before 1857 for all children born between 1852 and 1856. This alternative specification does not include county fixed-effects. Results are very similar in magnitude to the baseline model displayed Table 3.1, suggesting that geographical factors at the county level cannot fully account for paternal coercion exposure's effect on child mortality. For λ a coefficient, the hazard ratio associated with a one-unit increase in the corresponding predictor variable increases by a factor of exp(λ), or (exp(λ)-1)% .***p $<$ .01; **p $<$ .05; *p $<$ .1

Table 3.A10: Monthly Risk of Child Death Before 1857, Movers only

	Fathers Who Moved				Mothers Who moved			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Reference category: Lower coercion exposure								
Mother: High coercion	0.19	0.17	0.21	0.20	−0.09	−0.11	−0.11	−0.11
	(0.16)	(0.16)	(0.17)	(0.18)	(0.08)	(0.10)	(0.10)	(0.11)
Mother: Extreme Coercion	0.13	0.09	0.17	0.27	−0.05	−0.05	−0.01	−0.01
	(0.19)	(0.19)	(0.20)	(0.21)	(0.10)	(0.12)	(0.13)	(0.13)
Father: High Coercion		−0.01	0.06	0.12		−0.01	0.04	0.09
		(0.16)	(0.17)	(0.18)		(0.17)	(0.18)	(0.19)
Father: Extreme Coercion		0.42***	0.47***	0.54***		0.37**	0.44***	0.51***
		(0.16)	(0.17)	(0.18)		(0.17)	(0.18)	(0.19)
Father: Absent						0.14	0.13	0.19*
						(0.13)	(0.14)	(0.15)
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mother controls	No	No	Yes	Yes	No	No	Yes	Yes
Father controls	No	No	No	Yes	No	No	No	Yes
Mean	14.26	14.27	13.97	14.12	14.15	14.18	13.99	14.06
Observations	2,919	2,886	2,562	2,345	9,504	6,575	5,913	5,437

Notes: This table displays the Cox proportional hazard estimated effect of parental exposure to coercion on children's risk of death before 1857 for children born between 1852 and 1856, whose father (resp. mother) moved from their county of enslavement. For λ a coefficient, the hazard ratio associated with a one-unit increase in the corresponding predictor variable increases by a factor of exp(λ), or (exp(λ)-1)% .***p $< .01$; **p $< .05$; *p $< .1$

Table 3.A11: Monthly Risk of Child Death Before 1857, Extreme coercion at 80th percentile

	Extreme Coercion at 80th percentile			
	(1)	(2)	(3)	(4)
Mother: High coercion vs. Low	−0.04	−0.09	−0.09	0.18
	(0.07)	(0.08)	(0.09)	(0.15)
Mother: Extreme Coercion vs. Low	−0.07	−0.10	−0.09	−0.08
	(0.07)	(0.09)	(0.09)	(0.15)
Father: High Coercion vs. Low		−0.06	−0.01	0.08
		(0.13)	(0.14)	(0.16)
Father: Extreme Coercion vs. Low		0.22**	0.28**	0.27*
		(0.11)	(0.12)	(0.14)
Father: Absent vs. Low		0.04	0.04	
		(0.09)	(0.10)	
County and Year FE	Yes	Yes	Yes	Yes
Mother controls	No	No	Yes	Yes
Father controls	No	No	No	Yes
Mean	14.11	14.08	13.96	13.8
Observations	17,274	12,801	11,019	4,441

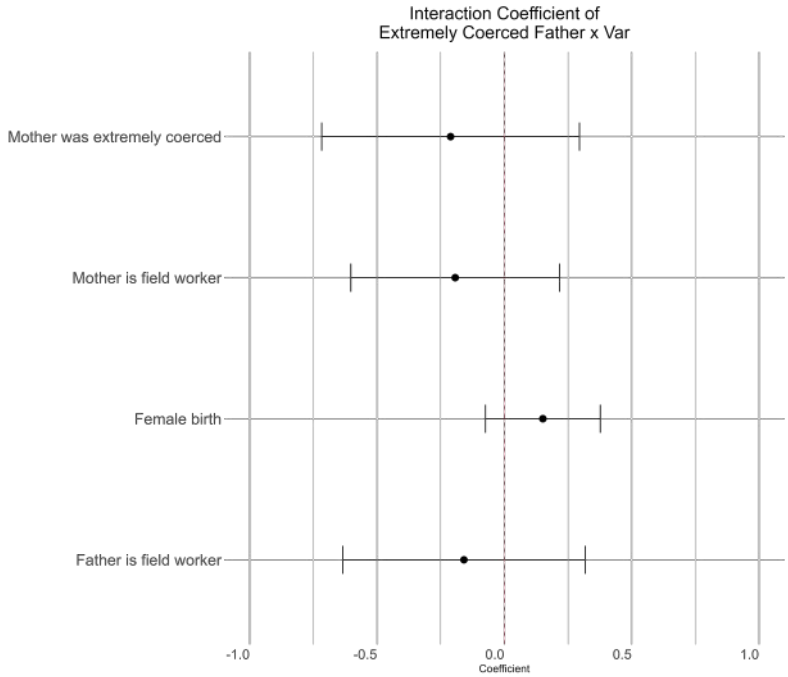
Notes: This table displays the Cox proportional hazard estimated effect of parental exposure to coercion on children's risk of death before 1857 for all children born between 1852 and 1856. This alternative specification uses a different "extreme coercion" category, using the 80th percentile of the share of sugarcane over cultivated surface in 1820 as threshold. Results are very similar in magnitude to the baseline model displayed Table 3.1, suggesting little sensitivity of the results to the definition of the coercion category. For λ a coefficient, the hazard ratio associated with a one-unit increase in the corresponding predictor variable increases by a factor of exp(λ), or (exp(λ)-1)% .***p $<$ .01; **p $<$ .05; *p $<$ .1

Table 3.A12: Monthly Risk of Child Death Before 1857, Extreme coercion at 90th percentile

	Extreme Coercion at 90th percentile			
	(1)	(2)	(3)	(4)
Mother: High coercion vs. Low	−0.08	−0.11	−0.11	0.08
	(0.06)	(0.08)	(0.08)	(0.14)
Mother: Extreme Coercion vs. Low	0.01	−0.04	−0.04	−0.04
	(0.09)	(0.11)	(0.12)	(0.18)
Father: High Coercion vs. Low		0.01	0.09	0.14
		(0.11)	(0.12)	(0.14)
Father: Extreme Coercion vs. Low		0.29**	0.31**	0.32*
		(0.14)	(0.15)	(0.17)
Father: Absent vs. Low		0.04	0.05	
		(0.09)	(0.10)	
Island, County and Year FE	Yes	Yes	Yes	Yes
Mother controls	No	No	Yes	Yes
Father controls	No	No	No	Yes
Mean	14.11	14.08	13.96	13.8
Observations	17,274	12,801	11,019	4,441

Notes: This table displays the Cox proportional hazard estimated effect of parental exposure to coercion on children's risk of death before 1857 for all children born between 1852 and 1856. This alternative specification uses a different "extreme coercion" category, using the 90th percentile of the share of sugarcane over cultivated surface in 1820 as threshold. Results are very similar in magnitude to the baseline model displayed Table 3.1, suggesting little sensitivity of the results to the definition of the coercion categories. For λ a coefficient, the hazard ratio associated with a one-unit increase in the corresponding predictor variable increases by a factor of exp(λ), or (exp(λ)-1)% .$^{***}p < .01$; $^{**}p < .05$; $^{*}p < .1$

Figure 3.A7: Additional Effect of Having an Extremely Coerced Father on Child Mortality



Notes: This figure displays the estimated Cox Proportional Hazard effect of having an extremely coerced father (respective to less coerced or absent) interacted with either of these variables: Mother was extremely coerced; Mother is field worker; female birth; Father is field worker. The average estimated effect of having an extremely coerced father is equal to 0.33, and all interaction coefficients are non statistically different from 0 and less than 0.2 in absolute value. This suggest that having an extremely coerced father always has a significant impact on children's risk of death, across all sub-groups. Confidence intervals are defined at the 95% level.

Table 3.A13: Risk of Child Death Before 1857, According to County of Residence Characteristics

	(1)	(2)	(3)	(4)
Total Surface	−0.000		−0.000	
	(0.000)		(0.000)	
Population in 1820	−0.000		0.000	
	(0.000)		(0.000)	
Median Slope	0.01		0.002	
	(0.03)		(0.03)	
Soil Suitability	0.18		0.19	
	(0.04)		(0.04)	
Sugarcane intensity		−0.52		−0.58
		(0.13)		(0.16)
Island FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Parents' coercion history	No	No	Yes	Yes
Mean	14.16	14.1	14.1	14.07
Observations	16,312	17,136	12,038	12,719

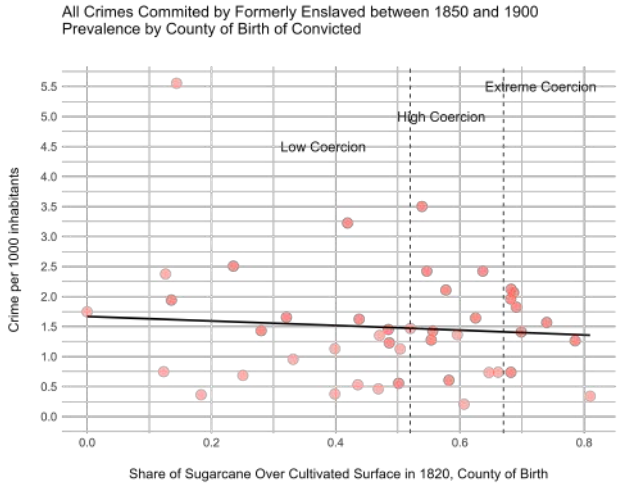
Notes: This table displays the estimated effect of sugarcane intensity in the county of birth, on children's risk of death before 1857. Sugarcane intensity is defined as the share of sugarcane over cultivated surface in 1820. Columns (2) and (4) contain estimated coefficients of the effect of the share of sugarcane over cultivated surface on child mortality. Columns (1) and (3) contain estimated coefficients of the effect of other local factors (population size, soil suitability, slope) on child mortality. These estimations are estimated separately, as sugarcane intensity is partially correlated with these other county level factors. Differences in the number of observations between columns (1) and (2) stems from the fact that topographic characteristics are not available for all counties. Reduction in the sample size for columns (3) and (4) stems from inclusion of parents' coercion history, which is not available for nearly 4,000 fathers. For λ a coefficient, the hazard ratio associated with a one-unit increase in the corresponding predictor variable increases by a factor of exp(λ), or (exp(λ)-1)% .***p $< .01$; **p $< .05$; *p $< .1$

Table 3.A14: Probability for Fathers to be Different, Conditional on First Father Being Known

	(1)	(2)	(3)	(4)
First father: High Coercion vs. Low	−0.29	−0.30	−0.33	−0.18
	(0.44)	(0.42)	(0.43)	(0.43)
First father: Extreme Coercion vs. Low	0.45	0.51	0.44	0.52
	(0.39)	(0.40)	(0.39)	(0.39)
Number of Births		0.71*	0.71*	0.69*
		(0.38)	(0.38)	(0.39)
Mother's age			0.02	0.02
			(0.03)	(0.03)
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Mother's occupation	No	No	No	No
Mother's coercion exposure	Yes	Yes	Yes	Yes
Mean	0.08	0.08	0.09	0.09
Observations	643	643	622	609
Akaike information criterion	401.61	400.36	400.09	394.55

Notes: The sample used to estimate these coefficients are mothers who had several children between 1852 and 1856, and where the father of the first child is known. Coefficients reflect the logit estimated effects of each covariate on the odds that the next father(s) would be different than the first father. For λ a coefficient, the hazard ratio associated with a one-unit increase in the corresponding predictor variable increases by a factor of exp(λ), or (exp(λ)-1)% .$^{***}p < .01$; $^{**}p < .05$; $^{*}p < .1$

Figure 3.A8: Prevalence of Formerly Enslaved Prisoners Sent to Guyana 1850-1890, by Former County of Enslavement



Notes: This figure is a binned scatterplot with one bin corresponding to a former county of enslavement. It relates the number of prisoners sent to Guyana between 1850 and 1890 per 1,000 formerly enslaved individuals, related to the sugarcane intensity in the former county of enslavement. The Black line is a linear fit between the two variables, weighted by the enslaved population in 1848.

Table 3.A15: Risk of Child Death before 1857, Depending on Whether Parents Are Field Workers

	(1)	(2)	(3)	(4)
Mother: Field Worker	0.05	0.08	0.08	0.001
	(0.07)	(0.12)	(0.12)	(0.16)
Father: Field Worker		0.12	0.12	0.21
		(0.10)	(0.10)	(0.14)
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Mother's Coercion	No	No	Yes	Yes
Father's Coercion	No	No	No	Yes
Mean	14.1	13.96	13.96	13.97
Observations	14,854	7,838	7,834	4,615

Notes: This table displays the Cox proportional hazard estimated effect of parental exposure to coercion on children's risk of death before 1857. The number of observations decreases from column (1) to (2-4) because analyses are restricted to present fathers with information on their occupation. Loss of observations from column (3) to (4) lies in the fact that some present fathers have missing information on their former coercion history. For λ a coefficient, the hazard ratio associated with a one-unit increase in the corresponding predictor variable increases by a factor of exp(λ), or (exp(λ)-1)% .***p $<$ .01; **p $<$ .05; *p $<$ .1

Table 3.A16: Impact of Mothers' Enslavement Conditions on Probability to Have a Child with Father exposed to the *Same* Enslavement Conditions

	(1)	(2)	(3)	(4)
Mother: High Coercion vs. Lower	0.16^{**}	-0.16^{*}	-0.04	-0.06
	(0.09)	(0.10)	(0.11)	(0.08)
Mother: Extreme Coercion vs. Lower	-0.43^{***}	-0.23^{**}	-0.20^{*}	-0.56^{***}
	(0.10)	(0.12)	(0.14)	(0.09)
Mothers' population size$_{c,y}$		0.04^{***}	0.04^{***}	0.01^{***}
		(0.002)	(0.002)	(0.001)
Fathers' population size$_{c,y}$		0.08^{***}	0.09^{***}	0.05^{***}
		(0.004)	(0.005)	(0.003)
County FE	Yes	Yes	Yes	No
Year FE	Yes	Yes	Yes	Yes
Mother and gender controls	No	No	Yes	Yes
Mean	0.6	0.6	0.61	0.61
Observations	5,642	5,642	4,724	4,724
Akaike information criterion	7,132.32	5,712.21	4,602.39	5,263.81

Notes: This table displays the logit estimated effect of maternal exposure to coercion on the probability to have a child with a known father exposed to the *same* enslavement conditions, controlling for the total number of mothers (resp. fathers) of each coercion group in the county and year of birth. For λ a coefficient, the odds ratio associated with a one-unit increase in the corresponding predictor variable increases by a factor of exp(λ), or (exp(λ)-1)% .***p $<$ .01; **p $<$ .05; *p $<$ .1

4 $TrOCR_{Fr}$: a Comprehensive Pipeline for Processing 19th Century French Archives

This chapter, co-written with Arnault Gombert[1], presents the archival processing pipeline which we built to collect the archival data used in Chapters 2 and 3. This pipeline leverages several machine learning techniques - Optical Character Recognition (OCR) and Large Language Models (LLM) - and makes several important contributions:

- Open-Source Handwritten Data Processing Model: We introduce an open-source model tailored for the processing of images containing handwritten French text. This model enables efficient extraction of handwritten content from scanned documents.
- Manually French Labeled Handwritten Dataset: To train our model effectively, we curated a meticulously labeled dataset of handwritten data. This dataset serves as a crucial resource for training and validating the model's accuracy.
- Reproducible Research Notebooks: In an effort to promote transparency and enable further research, we provide a set of comprehensive notebooks that allow researchers to reproduce our work. Additionally, these notebooks can be adapted to accommodate other languages, thereby extending their utility beyond the scope of French archives.

We thereby contribute to scholars' effort in providing free and reproducible document processing technologies such as the *layout-parser* developed by Melissa Dell and co-authors (Shen et al., 2021). While existing solutions perform well for printed

[1] Freelance data-scientist and machine-learning engineer

text, they typically fall short when dealing with handwritten and archival documents. We intend to release the pipeline in open-source and to provide instructions on how to adapt it to other languages.

4.1 Overview of Pipeline Components

In this section, we provide a concise overview of the main components of the pipeline:

1. Parsing Layout: The initial stage involves layout parsing, wherein global images are segmented into smaller units, each containing a single line of text.

2. Optical Character Recognition (OCR) Module: Following layout parsing, the OCR module takes center stage, transcribing individual images into text.

3. Named Entity Recognition (NER) Module: In the final stage, the Named Entity Recognition module extracts key information from the transcribed archival content.

Each of these pipeline components has been meticulously developed and optimized independently to ensure minimal interference with other elements. Our overarching goal is to foster modularity, allowing for the replacement or enhancement of individual components without disrupting the overall input/output pipeline.

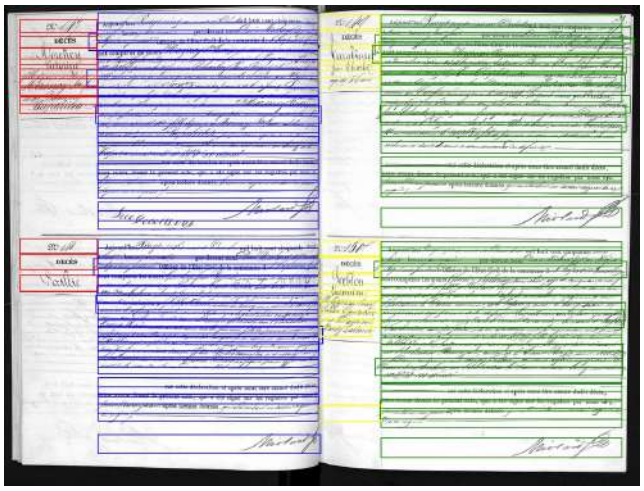
In the subsequent sections, we describe in more details each component.

4.2 Parsing Layout

Layout parsing is a document processing method that segments documents or images into distinct regions, enabling the analysis of their structure and arrangement. It plays a key role in tasks like OCR and information extraction by categorizing content elements based on spatial relationships and visual traits.

In this context, the LayoutParser library, introduced by Melissa Dell and her team (Shen et al., 2021), is a notable contribution. It is very useful for printed documents with clear visual delimitations between text components, but complex layout. Our specific research in digitizing 19th-century French archives presented unique challenges that called for a different approach. Our historical documents typically feature one or two pages of text with distinct margins and a single text column per page.

Figure 4.1: Example of layout-parsing on civil registries



Accurately identifying these columns is crucial for effective content consolidation, as content from different columns at the same vertical position may be unrelated.

We start from the methodology developed by (Grüning et al., 2019), which uses the ARU-Net neural network for detecting text lines in historical documents. The ARU-Net framework is available as an open-source solution for further research in this domain. We use the open-source implementation available at this GitHub repository, which aligns with our segmentation objectives.

Our main problematic lied in distinguishing margins and main texts accurately. We customize Grüning et al. (2019) as follow: we identified areas of high text density within the ARU-Net's XML outputs. Once we pinpointed the column boundaries, we focused on parsing each line consistently, ensuring lines of the same height were processed together to enhance the accuracy of our results.

As part of our ongoing research efforts, we intend to develop a tailored parser that is adaptable to various layouts, which are often found in historical archive documents (including tables). Our aim is to enhance its capacity to accurately identify the precise regions within these documents where textual content is located.

4.3 TrOCR French Handwritten Model

TrOCR (transformer OCR) is a technology and methodology used for Optical Character Recognition (OCR) tasks. It employs large transformer-based neural networks to extract text content from images, making it particularly effective for recognizing and transcribing text in scanned documents, images, or other visual media. We

start from the English large handwritten TrOCR model, originally introduced by (Li et al., 2022)[2]. A French version of the TrOCR handwritten model had not been released. To address this gap, we undertook the task of training a French model for proof-of-concept (PoC) purposes, using as baseline the English large handwritten TrOCR model.

4.3.1 Model Description

The TrOCR model is a sophisticated architecture, characterized by its encoder-decoder design. It incorporates an image Transformer serving as the encoder and a text Transformer functioning as the decoder. The initialization process for this model involves the use of two distinct pre-trained models: the image encoder is initialized from the weights of BEiT, whereas the text decoder is initialized from the weights of RoBERTa. In the TrOCR model, images are presented as a sequence of patches, each having a fixed size (16x16 resolution). These patches are linearly embedded to facilitate processing. Absolute position embeddings are introduced into the sequence before it undergoes transformations within the encoder layers of the Transformer. The Transformer text decoder operates in an autoregressive manner to generate textual tokens.

4.3.2 Intended Uses & Limitations

The TrOCR model is primarily designed for optical character recognition (OCR) tasks, specifically tailored for single text-line images. It exhibits optimal performance when applied within the context of these use cases. However, it's important to note that the model's capabilities are limited to single text-line OCR and may not be suitable for more complex image analysis tasks or multi-line text recognition.

4.3.3 Fine Tuning Process

The fine-tuning process for the TrOCR French Handwritten Model was conducted in two distinct phases, each employing specific datasets:

Dataset Generation. To adapt to the French vocabulary and incorporate names, surnames, occupations, cities, numbers, and text variations (informations that we needed to capture accuretly), we created a dataset containing 70,000 lines. This

[2]The English model was initially made available in the official repository as a TrOCR model fine-tuned on the IAM dataset.

dataset was generated using a combination of predefined lists and the Text Data Generator. We then conducted ten epochs of training exclusively on this dataset to specialize the model for French text patterns.

Fine-tuning with Handwritten Datasets. In the second phase, we fine-tuned the model over 20 epochs using two distinct handwritten datasets: a French Census dataset sourced from Constum et al. (2022), which we linked to the Hugging Face Hub to facilitate access. An additional dataset comprising 11,000 lines from French civil records, which we manually annotated.

4.3.4 Technical details

Parameters. In our experimentation, we employed heuristic parameters without undergoing a distinct hyperparameter tuning process. The key parameters used in training the model include:

Parameter	Value
Learning Rate	4e-5
Epochs	20
Mixed Precision	True
Max Sequence Length	64
Batch Size	128
Train/Dev Split	90/10

Metrics on development set. We evaluated the performance of the model on both the development (dev) and test sets, yielding the following results:

- Set Size: 700 examples from the French Census dataset and 1600 from our own dataset (historical civil records from the French West Indies, 19^{th} century)
- Character Error Rate (CER): 0.0575
- Word Error Rate (WER): 0.1651
- Loss: 0.5768

Metrics on test set.

- Set Size: 730 examples from the French Census dataset and 950 from our own dataset.

- Character Error Rate (CER): 0.09417
- Word Error Rate (WER): 0.23485
- Loss: 0.8700

Usage Instructions

The model is not yet published in public but will be released shortly. Below are the steps which will allow for using this model within the PyTorch framework:

```
from transformers import TrOCRProcessor, VisionEncoderDecoderModel,
    AutoTokenizer
from PIL import Image
import requests
from io import BytesIO

# Define the URL of the image to be processed
url = "mydesk/main/sample_imgs/5.jpg"

# Fetch the image data from the URL
response = requests.get(url)
img = Image.open(BytesIO(response.content))

# Initialize the TrOCR processor, model, and tokenizer
processor =
    TrOCRProcessor.from_pretrained('microsoft/trocr-large-handwritten')
model =
    VisionEncoderDecoderModel.from_pretrained('MarieBgl/trocr-large-handwritten-fr')
tokenizer =
    AutoTokenizer.from_pretrained('MarieBgl/trocr-large-handwritten-fr')

# Prepare the image data for input
pixel_values = processor(images=img, return_tensors="pt").pixel_values

# Generate textual output
generated_ids = model.generate(pixel_values)
generated_text = tokenizer.batch_decode(generated_ids,
    skip_special_tokens=True)[0]
```

By following these steps, users will be able to effectively use the TrOCR model to perform optical character recognition (OCR) on images using PyTorch: simply provide the image that needs to be processed, and the model will generate the corresponding textual content.

4.4 Extract entities from texts

The last step of the pipeline consists in extracting key information from archival records. Named Entity Recognition (NER) is a natural language processing technique that involves identifying and classifying specific named entities or entities of interest in text data. These entities can include names of individuals, organizations, locations, dates, monetary values, and more.

Large Language Models (LLM) are deep learning models that use transformer architectures, and have been pretrained on vast amounts of text data to understand and generate human-like language. These models can perform a wide range of natural language understanding and generation tasks, such as text completion, translation, and text summarization.

LLMs can efficiently perform NER tasks when provided with a schema parameter and a well-crafted prompt. By specifying a schema, it is possible to guide the model to recognize and extract specific types of entities, such as dates, names, or locations, from the text it generates. A well-constructed prompt, which is a textual input or question, can further instruct the model to focus on extracting particular information, making it a powerful tool for automating entity extraction tasks in various domains, from extracting financial data to answering specific questions about a given text.

Here is an example of a final output that we obtained for a birth certificate using LLM:


```
"name": "Petrin Isabelle Marie Constance",
"sex": "femme",
"birth date": "23 Fevrier",
"birth place": "Coma section des hauteurs",
"father's name": "Setrin Vital",
"father's age": "23",
"father's job": "cultivateur proprietaire",
"father's birth place": "Abymes",
```

```
"father's residence place": "section des hauteurs",
"mother's name": "Urie Marie Antoinette Dedoise",
"mother's age": "23",
"mother's job": "sans profession",
"mother's birth place": "None",
"mother's residence place": "Abymes"
```

5 Behind closed doors: detection of male-based violence in pandemic times

5.1 Introduction

After presenting my research on Slavery related violence, and the tool that I developed for processing French Handwritten Archives, this chapter presents another aspect of my research: on gender-based violence and institutional responses in contemporary settings.

Violence against women and children is a worldwide concern and a major public health problem[1]. According to the World Health Organization, 30% of women and young girls worldwide suffered from physical or sexual abuse from men at some point in their life, with numerous and long-lasting consequences on survivors' lives (World Health Organization, 2018).

Male-based violence (MBV) is also a largely under-reported event, which poses severe threats to the detection and therefore protection of abused victims by the state. The health sector holds a privileged position in that regard. First, as provider of comprehensive care to victims. Second, because victims might be more inclined to disclose their private situation to health workers, than they are to police officers. As a result, detection of abuse through the health care system is often much larger than detection of abuse through the police (World Health Organization, 2020). For children, schoolteachers are in best place to detect abuse and report them to social services or the police (Baron et al., 2020; Takaku and Yokoyama, 2021).

The outbreak of the COVID-19 pandemic pushed existing systems of protection to their limits, in a context where women's and children's risks of being abused

[1] In the remainder of this chapter, I will refer to violence against women and children, which is largely perpetrated by men, as male based violence (MBV).

rose significantly (Arenas-Arroyo et al., 2021; Miller et al., 2020; Beigelman and Vall Castelló, 2020; Leslie and Wilson, 2020; Asik and Nas Ozen, 2021; Berniell and Facchini, 2021; Bhalotra et al., 2024). This is particularly true for the health care system, which remained under heavy constraints even when lockdown policies were lifted.

Previous studies have mostly investigated the impact of the pandemic through the lens of hotline calls or reporting to the police, which makes it difficult to detect cases of abuse against children[2] and to approach effective care provided to victims. Previous evidence has shown that school closures during the pandemic imposed a heavy toll on the detection of abuse against children (Takaku and Yokoyama, 2021). We study another key channel of detection of abuse against both women and children which has been relatively left behind: the healthcare system.

We investigate the effect of the pandemic on detection of violence against women and children through the healthcare system in Catalonia, a Spanish region heavily hit by both the pandemic and stringent mobility restrictions. In this region, similarly to most other countries, public health providers play a key role in the detection and care provided to male-based violence victims: while around 4,800 victims were detected in 2019 through the police forces, this number reached 9,000 for those detected through the healthcare system. While previous studies have used indirect measures of detected abuse - such as entry through the pediatric emergency unit Bullinger et al. (2021) -, we are able to observe all cases of abuse detected by healthcare workers, as well as the characteristics of victims. We use an empirical framework similar to a difference in differences design, where 2020 is our treated year and 2019 our control year. We use as post indicator all weeks following March 15th, during which school closures, lockdown and state of emergency were declared (we refer to this bundle of events of March 15th as "pandemic outbreak"). Our identification assumption is that, in the absence of the pandemic outbreak, the seasonality in detection of abuse would have been similar in 2020 and 2019.

We find that the pandemic outbreak and emergency measures (introduction of lockdown and school closures) caused a strong and long-lasting decrease in the detection of male-based violence against women and children. This effect was particularly strong for the most vulnerable groups of victims (low-income households, and children aged 14 or less). We explore alternatives channels of detection and protection of victims (police and women's centers) and find that the loss of detection

[2] Cabrera-Hernández and Padilla-Romo (2020) use crime reports to detect murders of children, which allows them to capture only the most extreme forms of child abuse.

through the health system was not compensated by these other channels. Overall, we estimate that the start of the pandemic emergency was responsible for a drop in detection of at least 32%, which affected more than 4,000 victims, among which 800 children. However, we do report that part of this reduction in detection through the health care system was offset by an increase in help-seeking behavior of victims through the emergency hotline, which experienced strong increases in the number of calls related to extreme forms of abuse: physical and sexual abuse (+70%), and abuse against children and young adults (+134%).

We contribute to the existing literature on the unintended effect of the pandemic on male-based violence on several dimensions. First, by providing further evidence on the impact of the pandemic on violence against children, who represent the most vulnerable group of victims. Second, by looking at detection of victims through the healthcare system. While previous research have consistently shown that pandemic-related school closures led to a decrease in reported abuse against children (Baron et al., 2020; Takaku and Yokoyama, 2021), we provide the first evidence on all cases of detected abuse against children through the healthcare system[3]. By doing so, we also overcome data-quality concerns regarding call data and police reports. A final contribution is to link detection of male-based violence in the healthcare system with detection by the police and through emergency hotlines, which allows us to provide a more complete picture of the level of detection and protection of victims during the pandemic.

In the next section we provide an overview of the overall system of detection of male-based violence victims in Catalonia as well as a chronology of the pandemic and quarantine measures. Section 3 focuses on the effect of the pandemic outbreak on detection through the public healthcare system. Section 4 analyzes alternative channels of detection - hotline calls from male-based violence victims followed by police registers and data from women centers'. The combination of these different data sources allows us to identify the extent to which those channels might compensate the loss in detection from the health care system.

[3] Bullinger et al. (2021) use entry in pediatric emergencies as indirect measure of abuse

5.2 Context

5.2.1 Detection of male-based violence victims

The Spanish law specifically recognizes gender violence as a specific type of violence with devastating consequences, and its detection has been given increased attention in the political agenda in the past decades[4]. This is especially true for the region of Catalonia[5], which has been implementing a range of policies to eradicate male-based violence[6].

The public healthcare system is the main institution of detection of male-based violence victims (MBVV) in Catalonia. In 2019, approximately 9,000 victims were identified by healthcare professionals. The other main channels of detection are the Police forces and Women's Information and Attention centers (SIAD): 4,800 victims of male-based violence were detected through the Police, 5,500 women and 250 children through SIADs. Finally, the Catalan emergency hotline acts as an entry point for detection of MBV, but does not provide care or protection to MBV victims. Hotline counsellors can advise women to contact other services after assessing the severity of the situation. In 2019, the hotline received 9,000 calls and 15% of them were recommended for a follow up with the police. All legal residents in Catalonia are entitled to free public healthcare and are assigned a specific care center (CAP) as well as a general practitioner and a nurse. The public healthcare system is the main channel of primary care in Catalonia, and accounts for 80% of all primary care visits.

The Intervention Protocol for Approaching Male-based Violence (PAVIM) was implemented in 2008 to promote the prevention, detection, attention and recovering of male based violence victims (MBVV) by the public healthcare system. Since then, health professionals are trained to detect abuse and victims' care needs, and to redirect them to internal care services (for physical or psychological care) or social service units. When deemed necessary, extreme forms of MBV can also be redirected to the police – although this is a rare event.

Detected abuses are reported in a central database, making the information accessible to the GP or nurse attending the patient for her next visit to any public care

[4]Some of the Spanish policies regarding MBV are regarded as some of the most advanced in Europe. An example is the VioGèn system, which aims at monitoring the risk of women that already faced male-based violence (generally domestic), by exploiting real-time information from social services, health services and the police on either the victim or her potential aggressor.

[5]There are 17 autonomous communities (regions) in Spain

[6]Law 5/2008, "of the rights of women to eradicate male-based violence"

facility.

5.2.2 Emergency policies in Catalonia

National level quarantine was implemented on the 15th of March 2020, a few days after school closures were declared (11th of March). For two months, Catalonia remained under strict lockdown, while other regions in the country least hit by the pandemic progressively relaxed mobility restriction. Mobility was limited to essential needs and restrictions included the closure of bars, restaurants, night clubs, movie theatres and theatres, etc. Furthermore, while most countries maintained authorization to go outside for physical activity or stroll around even at the highest peak of the pandemic (particularly for families with children), in the Spanish case this was not allowed until two months later during the 1st phase of the deconfinement measures. Throughout this chapter, we refer to this bundle of emergency policies implemented during the week of March 15th as "pandemic outbreak".

On the 11th of May, progressive deconfinement measures were implemented in the least hit counties[7], and later extended to all counties by the 25th of May. All schools (from kindergarden to highshools) were able to reopen on the 1st of June and remained fully opened throughout the 2020/2021 the academic year[8].

5.3 The effect of the pandemic outbreak on detection through the healthcare system

5.3.1 Data

We exploit individual level data from the Catalan Health Care System on visits of patients identified as MBVV by health professionals, from January 2019 to December 2020.

Data is available at the entry level with information on the date of visit, type of health facility (primary care, hospital, emergency), the victim (age group, nationality, income group) and the type of violence detected (sexual, physical, psychological, confirmed/suspected). For patients that visited the health facilities several

[7]Increased possibilities to stroll around, opening of restaurants, bars and cultural places with restrictions on capacity.

[8]Conditions of opening remained however stricter for universities, whose capacity remained limited to 50% throughout the academic year

times, we keep as reference the information relative to the 1st visit during the year. We then aggregate the number of visits at the province x week level, by type of violence detected, health facility, and characteristics of the victim.

We use two main outcomes of interest, which we define at the province and weekly level between January 2019 and December 2020: the number of MBVV patients per 100,000 inhabitants (women and children), denoted Y^1; the number of patients younger than 25 years old per 100,000 inhabitants (of the same age), denoted Y^2: we refer to this subgroup as the "children and young adult sample". We include 18-25 years old individuals in this subgroup even if they are officially adults, as more than 50% of them were still living with their parents in Catalunya[9]. We specifically distinguish between young children and young adults in our heterogeneity analyses. Different subgroups of patients (according to demographic characteristics, severity of abuse) are also identified within the overall sample and the "children and young adults" sample to explore the heterogenous effects across those different subgroups[10]. There are 4 provinces in Catalonia and 42 comarcas (counties). While information on detection is available at the county level, we aggregate the number of patients per provinces rather than counties as 20% of the cells defined by counties x week contain 0's.

In 2019 healthcare workers identified 6,845 women aged 25 years old or more as MBVV. This is equivalent to a yearly detection of 118 patients per 100,000 inhabitants - or a monthly detection of 9.9 patients per 100,000 inhabitants. That same year 2,164 children and young adults aged less than 25 were detected as being victims of male-based violence. This represents 25% of all detected victims and is equivalent to a yearly detection of 151 victims per 100,000 children or young adults (or monthly detection of 12.6 victims per 100 000 children or young adults). On average, identified MBVV visited a health care facility 1.7 times in 2019, against 1.9 in 2020; 75% were women or girls; 83% were coming from low-income households; 71% were Spanish nationals.

Among the detected abuses, 65% were considered "confirmed" (as opposed to "suspected") by health workers. 65% were detected through the primary care sys-

[9] INE, 2021.

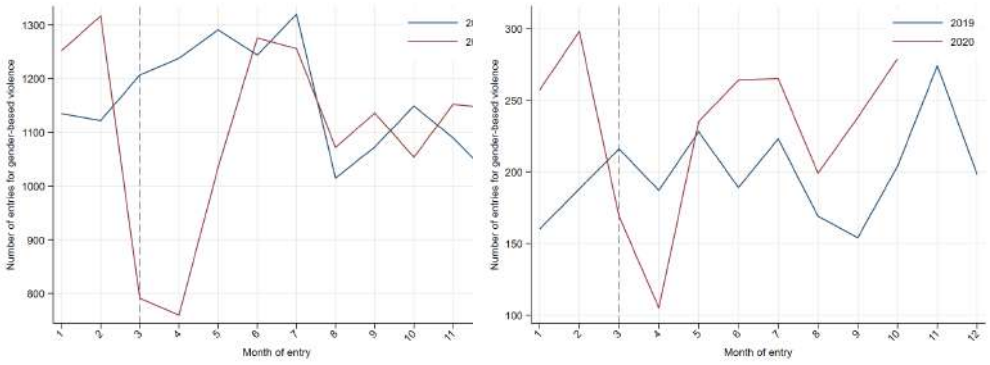
[10] For the overall sample, we study the additional effect of the pandemic outbreak on detection of : adults compared to children; cases of confirmed violence abuse compared to suspected abuse; foreign victims compared to native victims; victims of sexual abuse compared to other types of abuse; low-income victims compared to middle and high income victims. For the children sample, we study the additional effect of the pandemic outbreak on detection for: girls compared to boys; victims aged less than 14 years old compared to 14 years old or more; victims of sexual abuse compared to others; victims from low income families compared to others.

tem (regular visit to a GP), 30% through emergencies, 5% through mental hospitalization or ambulatory care.

Detection of abuse per 100,000 inhabitants in the pre-pandemic period was highly heterogenous across counties (see Maps 5.A2 and 5.A4 in appendix), and could be driven by a range of factors, including healthcare usage and professionals' training in detecting abuse[11]. To account for time-invariant characteristics likely to affect detection, we include province-level fixed effects in all econometric specifications.

Plots of the monthly number of patients detected through the healthcare system (adults and children) are displayed Figure 5.1. The month of the pandemic outbreak and emergency policies coincides with a clear drop in detected abuse through the emergency hotline in 2020 compared to 2019. We now formally investigate changes in detected abuses through the heathcare system, incurred by the pandemic outbreak.

Figure 5.1: Monthly number of victims of abuse detected by the Healthcare system, in 2019 and 2020



(a) Adults <25yo (b) Children and Young Adults

Source: Acquas

5.3.2 Empirical strategy

We exploit province and monthly (or weekly, depending on the specification) level information from the healthcare system on the number of patients detected as MBVV between January 2019 and December 2020 and use a difference-in-differences (DiD) framework to estimate the impact of the pandemic outbreak and emergency policies on the week of the 15th of March 2020 (11th week, or 3rd month) on detection of

[11] Another important factor is, of course, the prevalence of abuse

MBVV. We assume that the emergency measures was an exogenous, un-planned shock[12]: we consider 2020 as our "treated" year, and 2019 as our "control year"; our post-treatment period corresponds to the post-outbreak period from the 11th week or 3rd month (March), up to December. Our identification strategy relies on the assumption that, in the absence of the COVID-19 shock in Spain, seasonality in detection of male-based violence throughout 2020 would have been the same than in 2019 (common trend assumption).

Event-study estimates We first explore the dynamic effect of the outbreak and examine the plausibility of our common trend assumption by using the following event-study specification (EVS):

$$Y^{i}_{m,p,y} = \alpha + \sum_{j=-2}^{9} \eta_j \text{Month}_j + \sum_{j=-2}^{9} \beta_j \text{Month}_j D_{y=2020} + \lambda D_{y=2020} + \gamma_p + \gamma_y + \epsilon_{m,p,y} \tag{5.1}$$

Event study analyses are run at the monthly level rather than weekly level, to improve the precision of estimated coefficient and for graphical clarity[13]. The outcomes are: $Y^{1}_{m,p,y}$ the number of patients identified as MBVV by health professionals per 100,000 inhabitants in province p, month m, and year y; $Y^{2}_{w,p,y}$ the number of children or young adults identified as MBVV by health professionals per 100,000 children and young adults in province p, month m, and year y; $D_{y=2020}$ is a dummy which is equal to 1 for observations in 2020 – the "treated" year. For graphical clarity purposes, event-study analyses are done at the province and monthly level, with Month_0 being March, Month_1 April, etc. Estimates are therefore calculated for two lags before the implementation of lockdown (January and February), and 9 leads (April to December). β_j are our coefficient of interest: for a specific lead-/lag j around March, they capture the differential "effect" on detection of being in 2020, compared to being in 2019. We include province fixed effects to account for time-invariant characteristics that might explain differences in detection across provinces (health supply, demographic characteristics affecting health demand, previous training of health workers to detect MBV, etc.); year fixed-effects, which ac-

[12] While one can argue that confinement and school closures might have been anticipated in the light of the spread of the virus throughout the year and neighboring countries' experience, the specific date of implementation was sudden and largely unforeseen.

[13] Using weeks rather months decreases the precision of our estimates, as many province x week cells end up having null values.

count for yearly trends across all provinces (for instance, overall improvement in the detection of MBVV across provinces). Because we only have 4 provinces, we report wild-bootstrapped confidence intervals and p-values to account for clustering at the province-level.

Difference in differences We next use a DiD model to measure the average impact of the pandemic on detection at the province and weekly level:

$$Y^{i}_{w,p,y} = \alpha + \beta_1 D_{w\geq 11} + \beta_2 D_{y=2020} + \beta_{\text{did}} D_{w\geq 11} D_{y=2020} + \gamma_p + \gamma_w + \gamma_y + \epsilon_{w,p,y} \tag{5.2}$$

$D_{w\geq 11}$ is a dummy variable which is equal to 1 for the 11th and subsequent weeks of the year (equivalent to a "post" variable)[14]; $D_{y=2020}$ is a dummy equal to 1 for observations in 2020, and β_{did} is our coefficient of interest capturing the average effect of the pandemic outbreak on the number of MBVV detected per 100,000 inhabitants in each province from March 2020 to December 2020. We include province fixed effects, week fixed-effects, and year fixed-effects, as before. In alternative specifications, we also include province-level time-varying controls that might explain differences in detection of MBVV across provinces and over time: monthly police reports on sexual assaults and offenses per 100,000 inhabitants; quarterly unemployment rates; weekly number of deaths due to COVID-19 per 100,000 inhabitants. In all specifications, we use wild-bootstrapped confidence intervals and p-values to account for clustering of standard errors at the province-level.

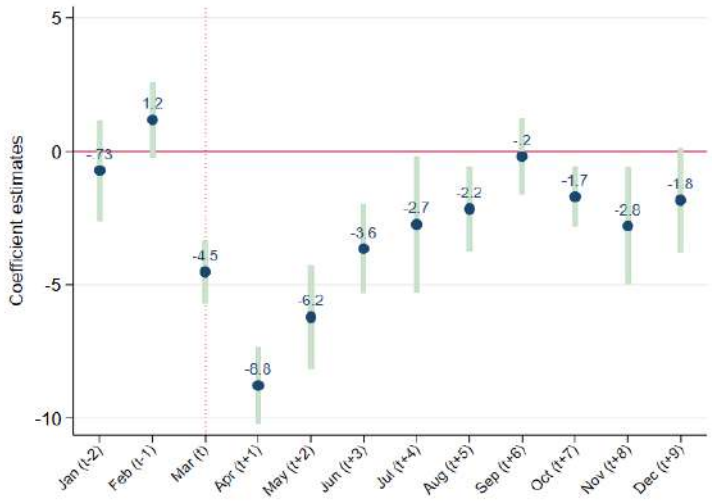
5.3.3 Results

Figure 5.2 displays event-study coefficients for the monthly number of MBVV per 100,000 inhabitants detected in 2020, compared to 2019 (Eq 1). Figure 5.A8 in appendix displays coefficient estimates for the children and young adult sample. In both graphs, point estimates associated with the lags are small in magnitude and non-statistically different from 0, which is suggestive of the existence of parallel trends.

We document a very large decrease in detected abuse against patients in the immediate aftermath of the pandemic outbreak. In the first fully treated month of April, we measure a decrease of -8.8 detected victims per 100,000 inhabitants in 2020 compared to 2019. This represents a drop in detection of 90%. For children

[14] Which corresponds to the week in which the first confinement was implemented in 2020

Figure 5.2: Effect of pandemic outbreak on detection of male-based violence against women and children, event study estimates



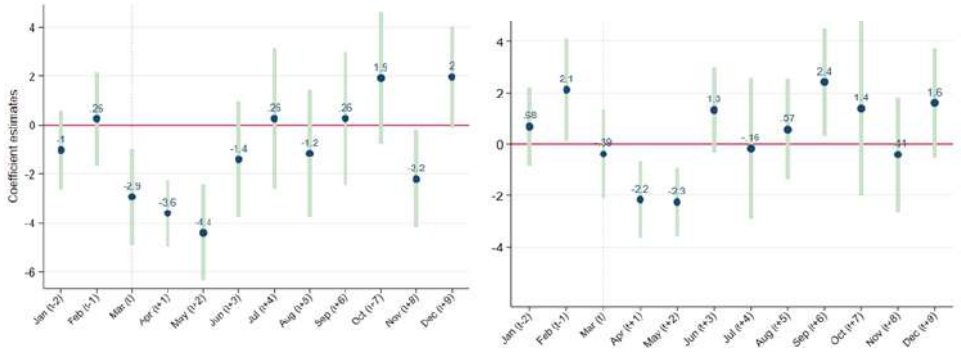
Notes: This graph displays monthly event study estimates. For a given month, the value of the coefficient corresponds to the difference in detection observed between the treated year (2020) during this month, and the control year for the same month (2019). The red dotted line indicates the month of implementation of the emergency policies (lockdown and school closures), and start of the congestion of the health system (15th of March).

and young adults, we document a decrease of -8.9 detected victims per 100,000 children and young adults, which represents a decrease of 70% compared to pre-pandemic levels[15]. Survey-based evidence of an increase in intrafamilial abuse following quarantine implementation in Spain (Arenas-Arroyo et al., 2021) suggests that the large decrease in detected cases from March to May 2020 unlikely stems from a decrease in violence. Rather, we argue that the pandemic outbreak had a large negative effect on detection of abuse through two channels: (1) a congestion effect, whereby primary health care facilities were overwhelmed with pandemic-related cases starting March 2020 (2) a health demand effect whereby women and children tended to visit care facilities only for urgent matters in the context of the quarantine and for fear of the virus.

From May to August, we measure a relative increase in detection of MBV compared to March 2020 as deconfinement measures were progressively set in place and the burden on reanimation units was relieved. We note, however, that detection levels remained systematically below 2019 levels even when strict social distancing measures were lifted. This suggests a "loss" in detection of MBV. Heterogeneity

[15] 12,6 MBVV per 100,000 children and young adults detected in 2019

Figure 5.3: Heterogeneity analyses of the impact of the pandemic outbreak on abuse against children and young adults



(a) Detection of abuse against low income children and young adults, relative to middle or higher income

(b) Detection of abuse against girls, relative to boys

Notes: This graph displays monthly event study estimates. For a given month, the value of the coefficient corresponds to the difference in detection observed between the treated year (2020) during this month, and the control year for the same month (2019). The red dotted line indicates the month of implementation of the emergency policies (lockdown and school closures), and start of the congestion of the health system (15th of March).

analyses of the differential effect of the pandemic by subgroups overtime are presented for all patients Figure 5.A9 and Figure 5.3 for children and young adults. We find that loss of detection was stronger for the least severe cases of violence (non-sexual violence vs. other cases of violence); for low-income patients (both women and children); and for girls. We also find that this "loss" in detection during the first few months was not caught up when confinement measures were gradually lifted from June.

Estimations of the average effect of the pandemic outbreak on detection between March and December 2020 compared to 2019 are presented Table 5.1. We find that the introduction of the emergency policies led to an average decrease of weekly-detection of -0.69 victims of MBV per 100,000 inhabitants (specification 1). This is equivalent to an accumulated decrease of -28,3 MBVV detected per 100,000 inhabitants between March and December 2020 compared to pre-pandemic levels (41 weeks); it represents a 32% decrease (36% when using a log-specification) compared to pre-pandemic levels. We find similar results when focusing on detection of violence against children and young adults (Specification 4 to 6).

Heterogeneity analyses of the average effect of the introduction of emergency policies and the pandemic outbreak on detection of MBVV are presented in the

appendix, Table 5.A4. In line with the event-study estimates, we find that the loss in detection was stronger for low-income patients, both in the overall sample as well as the children and young adult sample (Table 5.2). We also find that non-sexual MBV and more generally extreme forms of violence (entry through emergencies) suffered a lesser drop in detection. Focusing on violence against children and young adults, we find that the entire effect on detection for this subsample is driven by loss in detection against very young children (of 14 years old or less, Table 5.A4 specification 3). This is equivalent to a -47% change in weekly detection compared to pre-pandemic levels for this subgroup. Given the absence of alternative means of detection of violence against children in a context where schools were closed, this result is particularly worrying.

Our results point to a clear drop in detection and, therefore, of care provided to MBVV by the healthcare system. Given the evidence pointing towards an increase of violence against women and children during the pandemic, the drop in detection that we document is very unlikely driven by a decrease in MBV prevalence. We therefore argue that the pandemic outbreak in Catalunya caused a loss of care for at least: 3 245 male based violence victims (women and children), among which 822 aged less than 25 years old and 469 aged less than 14 years old (See Table 5.A5 for back of the envelope calculations). We provide further evidence to support this statement in the next section.

5.4 Alternative channels of detection

We now investigate whether the loss in detection through the healthcare system might have been compensated by other channels: emergency hotlines - that act as entry points for detection of MBV but do not provide care or protection to MBVV; police; women's center (SIAD's).

5.4.1 Emergency hotline

The Catalan emergency hotline for male-based violence was set in place in 2001 by the regional government. Hotline counsellors provide psychological support and advices. Counsellors do not transfer calls to other protective services (health, police, etc.), but do advise women on which institution they should turn to after assessing the severity of the situation. Data on the number of calls to the hotline is available at the daily and county level from 2013 to 2020, with information on the caller (age

Table 5.1: Impact of pandemic outbreak on weekly cases of abuse detected through the healthcare system, difference in differences estimates

Victims	Adult women			Children and young adults		
Outcome	*Per 100,000*		*Log*	*Per 100,000*		*Log*
	(1)	(2)	(3)	(4)	(5)	(6)
DiD: $D_{w\geq 11}D_{y>2019}$	-0.69*	-0.56**	-0.38***	-0.69**	-0.66*	-0.36***
	(0.33)	(0.19)	(0.13)	(0.23)	(0.27)	(0.07)
Post:$D_{w\geq 11}$	1.22***	0.86*	0.71***	0.22	0.18	0.07
	(0.14)	(0.39)	(0.15)	(0.14)	(0.24)	(0.06)
Treated: $D_{y>2019}$	0.32	0.31	0.19	0.22	0.18	0.07
	(0.40)	(0.52)	(0.36)	(0.26)	(0.30)	(0.23)
Monthly Police reports	0.03			-0.01		
	(0.17)			(0.15)		
Quaterly unemp. rate	-4.00			-0.06		
	(13.26)			(10.10)		
Weekly Covid-19 cases	-0.64^{+}			-0.59***		
	(0.33)			(0.17)		
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Week and year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	360	416	416	360	416	416
R^2	0.42	0.32	0.30	0.22	0.17	0.18

Notes: This table contains the difference in differences estimates of the effect of pandemic outbreak on the weekly cases of abuse detected through the healthcare system. For λ a coefficient, an increase in one unit in the covariate is associated with a λ unit increase in the weekly calls. ***p $< .01$; **p $< .05$; *p $< .1$

Table 5.2: Additional effect of pandemic outbreak on weekly detection of MBV against children and young adults per 100,000 children and young adults, heterogeneity analyses

D_g	Victim is a girl (1)	Abuse confirmed (2)	Victim is ≤14yo (3)	Entry by emergency (4)	Foreign Born (5)	Low income family (6)
DiD: $D_{w\geq 11}D_{y>2019}$	-0.20	-0.57*	-0.05	-0.48**	-0.43^{+}	-0.24
	(0.37)	(0.23)	(0.17)	(0.18)	(0.23)	(0.16)
Post: $D_{w\geq 11}$	0.70	0.59**	0.49*	0.33	0.03	0.14
	(0.44)	(0.20)	(0.25)	(0.23)	(0.28)	(0.27)
Treated: $D_{y>2019}$	-0.56	0.29	-0.20	0.55*	-0.13	-0.19
	(0.49)	(0.27)	(0.21)	(0.27)	(0.33)	(0.23)
DDD: $D_g D_{w\geq 11} D_{y>2019}$	-0.57	0.42	-0.49*	0.37	0.21	-0.18
	(0.54)	(0.27)	(0.25)	(0.26)	(0.28)	(0.27)
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
Week and year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	686	697	708	699	667	619
R^2	0.293	0.204	0.194	0.239	0.401	0.459

Notes: This table contains the triple differences estimates of the effect of the pandemic outbreak introduction on the weekly cases of abuse detected through the healthcare system, by subgroups, focusing only on the children and young adult sample. the DDD estimate for the 1st specification represents the additional effect on detection of *girl* victims, respective to boys. The total effect on the detection of abuse against girls is equal to $DD + DDD$. ***p < .01; **p < .05; *p < .1

group, employment status, educational level), the reason for calling, the relationship to the aggressor, the type of violence suffered, the number of children (<18 years old) exposed to the same violence at home.

One concern is that individuals might have reached out to the hotline because other channels of information and psychological supports where no longer accessible during quarantine, meaning that calls to the hotline during the pandemic might not reflect "real" abuses. We therefore exclude calls made for information purposes in which the caller was not reporting any abuse. Additionally, we also distinguish abuses according to their severity, and consider as severe: abuse of children and young adults; sexual or physical abuse; calls redirected to the police. We restrict our analyses to such calls, as said calls are much more likely to reflect cases of real abuse and help-seeking behavior of intra-familiar violence victims.

Table 5.A3 in the appendix displays descriptive statistics of calls in 2019 and 2020. Data for the year 2020 is only available from January to July 2020. Nevertheless, we already document a clear increase in reporting: from 9,000 reported abuse cases in 2019 to more than 14,000 in 2020. Among reported victims, 750 were children and young adults in 2019, and almost twice as much in 2020 (1,154) . The increase in calls in 2020 relatively to 2019 coincides with the introduction of emergency measures in March (Figure 5.A6). We also detect a sizable increase in calls for the most severe forms of abuse (physical or sexual abuse and calls redirected to the police), indicating an increase in help-seeking behavior for severe abuses through the hotline.

We compare changes in calls for severe cases of abuse in 2020 compared to previous years. We use a difference in differences setting and study both the dynamics and the average effect of the pandemic outbreak on the county x week calls per 100,000 inhabitants according to the severity of the abuse (against children, sexual or physical, redirected to the police).

To account for the fact that many cells in our data have zero values (0 calls), we use a two-part model for mixed discrete-continuous outcomes (Liu et al., 2010; Grüning et al., 2019). We first estimate the effect of the outbreak on the probability of observing the abuse using a logit model (step 1); we then estimate the effect of the outbreak on the increase in the number of calls per 100,000 inhabitants conditional on observing the abuse (step 2), using a generalized estimation equation with gamma density function[16].

[16]To account for the left-skewed distribution of abuses per 100,000 inhabitants

We display the marginal effects of the pandemic outbreak on calls per 100,000 inhabitants for the combined logit and GLM steps of the two-part model in Table 5.A6. Equivalent estimates expressed in percentages and absolute number of abuses are displayed Table 5.A7 . On average, we find that the introduction of emergency measures increased the detection of all types of severe abuses by 0.85 cases per 100,000 inhabitants (specification 1). This represents an increase in weekly detection of 47%, or an additional 400 abuses detected, between March and July 2020 compared to pre-pandemic years. This increase is mostly driven by calls related to sexual or physical abuse (Table 5.A6, specification 2).

We then estimate a +53% increase in calls redirected to the police (specification 3), which might stem from: (i) an increase in severe cases of abuse that require police intervention; (ii) counsellors' change in redirection strategies in the context of the pandemic, during which shelters' and health facilities' capacity had been severely diminished. While both channels might drive this result, the fact that reported sexual / physical abuse (specification 2) increased by the same proportion (47%) is suggestive of a increase in extreme violence against women and children. In the next section, we investigate the extent to which those redirected calls might have translated into proper protection by the police. Calls for abuse against children (specification 4) increased by 0.33 cases per 100,000 inhabitants aged less than 25, which is equivalent to an additional 50 cases of abuse detected between March and July, compared to pre-pandemic levels.

We therefore document that the pandemic outbreak caused a large decrease in detection of abuse through the healthcare system, which was partly offset by an increase in help-seeking behavior of victims through the emergency hotline. We note however that violence in the domestic environment was even harder to report for women during quarantine, while being on the watch of the violent partner. Thus, we consider reported cases of extreme abuse as a lower bound of the actual violence suffered by women and children.

We next explore whether alternative institutions for protection against MBV (police, women's centers) might have compensated the loss in detection through the healthcare system, and if the increase in calls to the hotline might have translated into specific actions.

5.4.2 Police and women's centers

We first exploit quarterly data on police reports for domestic abuse against women and young girls . We note a small increase in the number of reports following the introduction of emergency measures during the 2nd quarter of 2020 (Figure 5.A7). For both women and young girls, the number of additional interventions relatively to 2019 does not compare in magnitude with the estimated loss of detected cases through the healthcare system (between 20 and 100 cases)[17].

We next turn to women's centers (SIAD), which are distributed across the region (Map 5.A5) to offer information, counseling services and emergency support to women and young girls in situations of domestic abuse (finding emergency shelter, making a psychological assessment, etc.). During the pandemic, assistance through direct calls was reinforced to attend women in need. Physical visits were also available in some centers to attend emergency situations. We exploit information on each visit/call received in centers (age of victim, type of violence suffered, actions suggested, year of the visit) and compare the number of visits in 2019 and 2020 - unfortunately, information on the exact date of the visit is not available. Likewise, we note a very small increase in women and children attended to by SIAD professionals in 2020, which suggests that this channel unlikely compensated the loss in detection and care that we estimated through the Healthcare system (Table 5.A2).

5.5 Conclusion

We document a very strong and long-lasting decrease in the detection of male-based violence against women and children after the pandemic outbreak in March 2020, particularly during the months of strict lockdown. Loss of detection was even stronger for the most vulnerable categories of victims: individuals from low-income households, and children aged 14 years old or less.

Overall, we estimate that the pandemic caused a drop in detection of at least 32%, which might have prevented needed intervention for more than 4000 victims - among which 800 children - between March and December 2020. This loss in detection was likely driven by : (1) quarantine measures, which limited victims' mobility to attend health services - except for emergencies (2) a congestion effect, whereby health workers were over-burdened by the COVID-19 cases. However, we do report that part of the reduction in detection is offset by an increase in help-seeking behav-

[17] Against 4,000 undetected cases through the healthcare system

ior of victims through the emergency hotline, which experienced strong increases in the number of calls for extreme forms of abuse: physical and sexual abuse (+47%), and abuse against children and young adults (+33%). Although the likelihood of being redirected to the policy through the emergency line also increases, our data shows that those victims remained effectively unprotected by alternative channels such as police and women's centers. Our paper therefore suggests that the shadow pandemic caused by the COVID-19 went beyond an increased exposure to abuse of victims. Given the lingering effects of violence exposure, failure of the existing channels of protection to detect victims during the pandemic is likely to have important and long run consequences. This points to the urgency of public interventions to counterbalance the negative physical and mental health consequences of untreated violence against women and children. In particular, as children's possibility to report violence are much scarcer than women's, significant efforts should be engaged to detect abused children not only through the health system, but also in the school context.

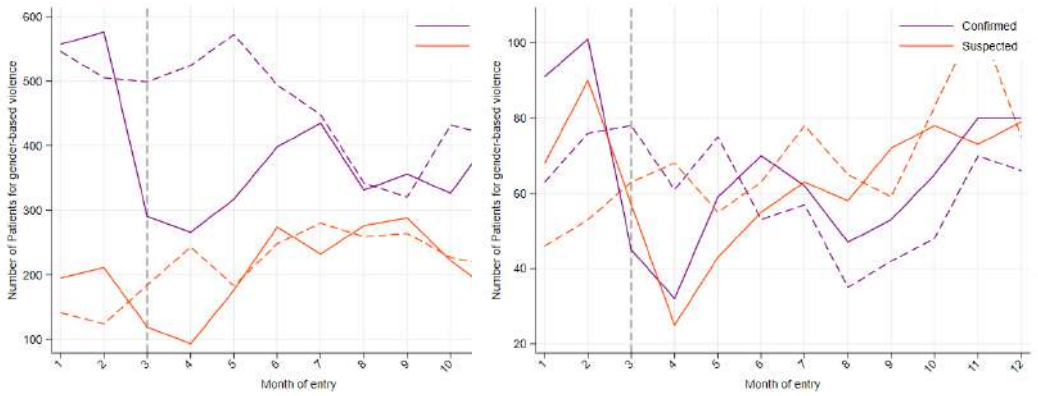
5.A Appendix, Chapter 5

5.A.1 Additional Tables and Figures

Table 5.A1: COVID-19 Restrictions in 2020

Month of 2020	Restrictions
March	11th: School closures implemented for all schools of all levels. 12th: All non-essential shops are closed; airports, railways, and ports are closed. 15th: Lockdown – mobility only allowed for essential workers and to buy essential products.
April	26th: Children can now stroll around with an adult for an hour around a 1km radius from home.
May	2nd: Outside activities (sports and stroll around) are allowed between 6am-10am and 8pm-11pm. 11th: The majority of Catalonian counties enter the first phase of the progressive deconfinement plan ("desescalada"). 25th: The metropolitan area of Barcelona enters the first phase of the deconfinement plan, while the rest of the counties enter the 2nd phase.
June	1st: Schools are opened in counties in the 2nd phase of the deconfinement plan, with obligations to wear masks and reduced class sizes.
July-December	Educational plan for the year 2020/2021 was approved: - From kindergarten to high school, with reduced class sizes (mixed format presencial/online for high school students made possible). - Universities: limited capacity of 50%. Online or presencial exams (university decisions).

Figure 5.A1: Monthly number of victims of abuse detected by the Healthcare system, in 2019 (dotted line) and 2020 (full line), by type of abuse (suspected or confirmed)



(a) Adults (b) Children and Young Adults

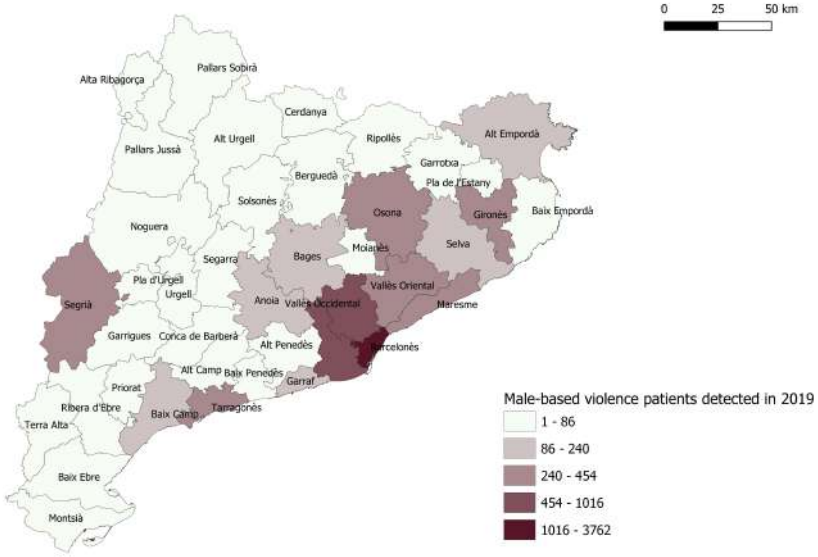
Source: Acquas, authors' own calculations

Table 5.A2: Number of women and children attended to in a SIAD center in Catalunya, in 2019 and 2020

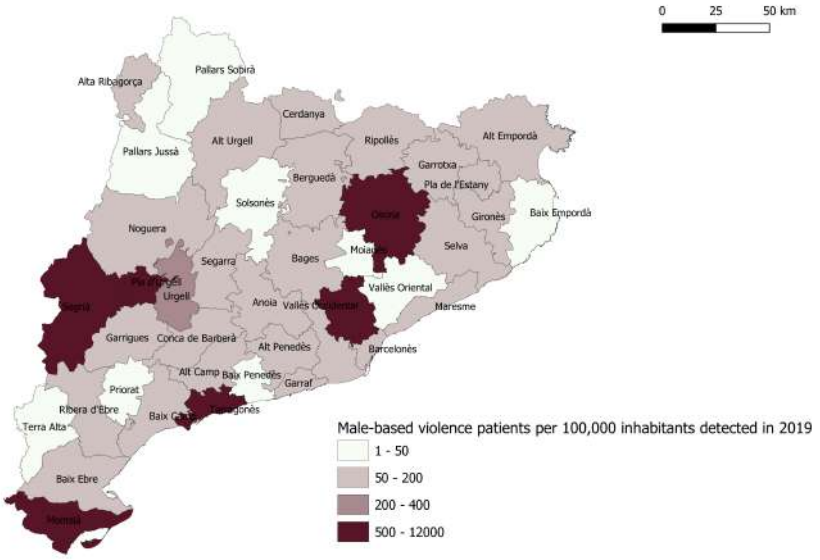
Age Group	**2019**	**2020**
<13 yo	46	64
13 to 18 yo	198	189
Total <18 yo	244	256
Women >18 yo	5,576	5,936

Source: Catalan Institute for Women, SIAD

Figure 5.A2: Victims of abuse detected by the Healthcare system in 2019, by county



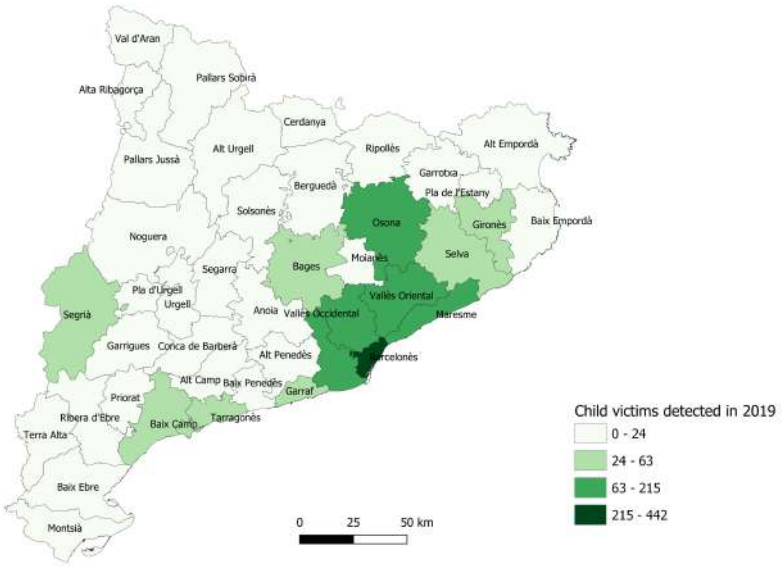
(a) Total



(b) per 100,000 inhabitants

Source: Acquas, authors' own calculations

Figure 5.A3: Number of children and young adults victims of abuse detected through the Healthcare system in 2019



Source: Acquas, authors' own calculations

Figure 5.A4: Changes in the number of victims of abuse detected through the Healthcare system in 2020, relative to 2019 $((v_{2020} - v_{2019})/v_{2020})$

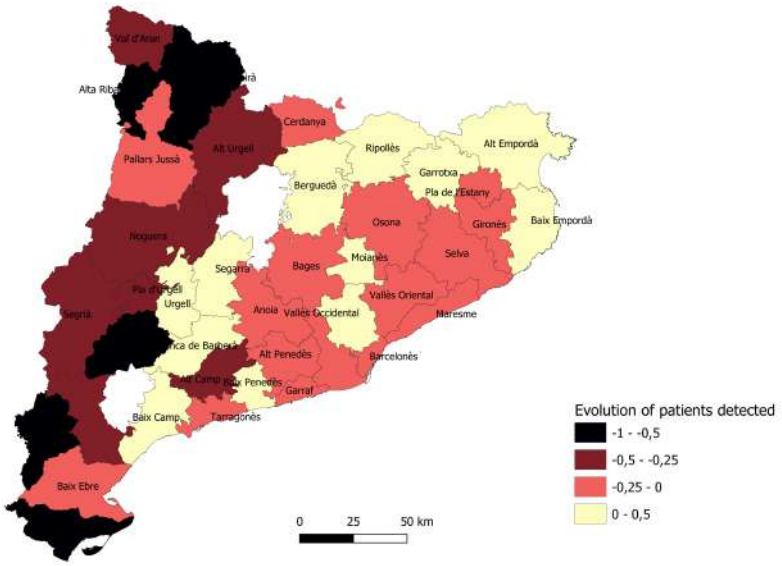


Figure 5.A5: SIAD centers in Catalunya

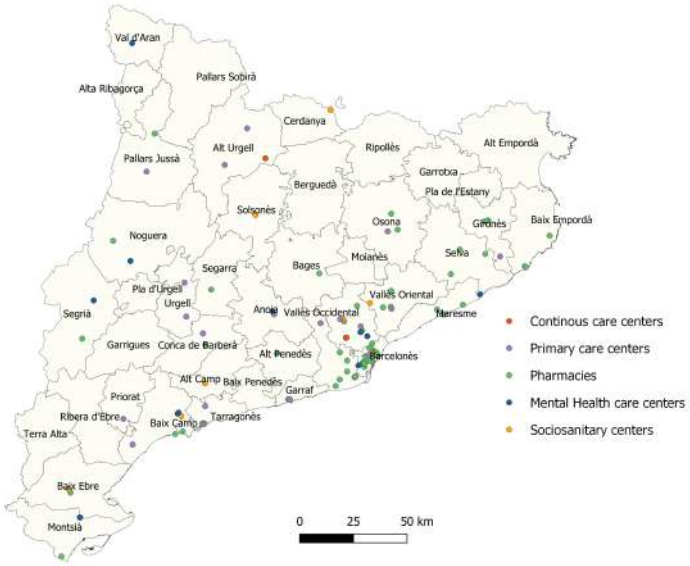
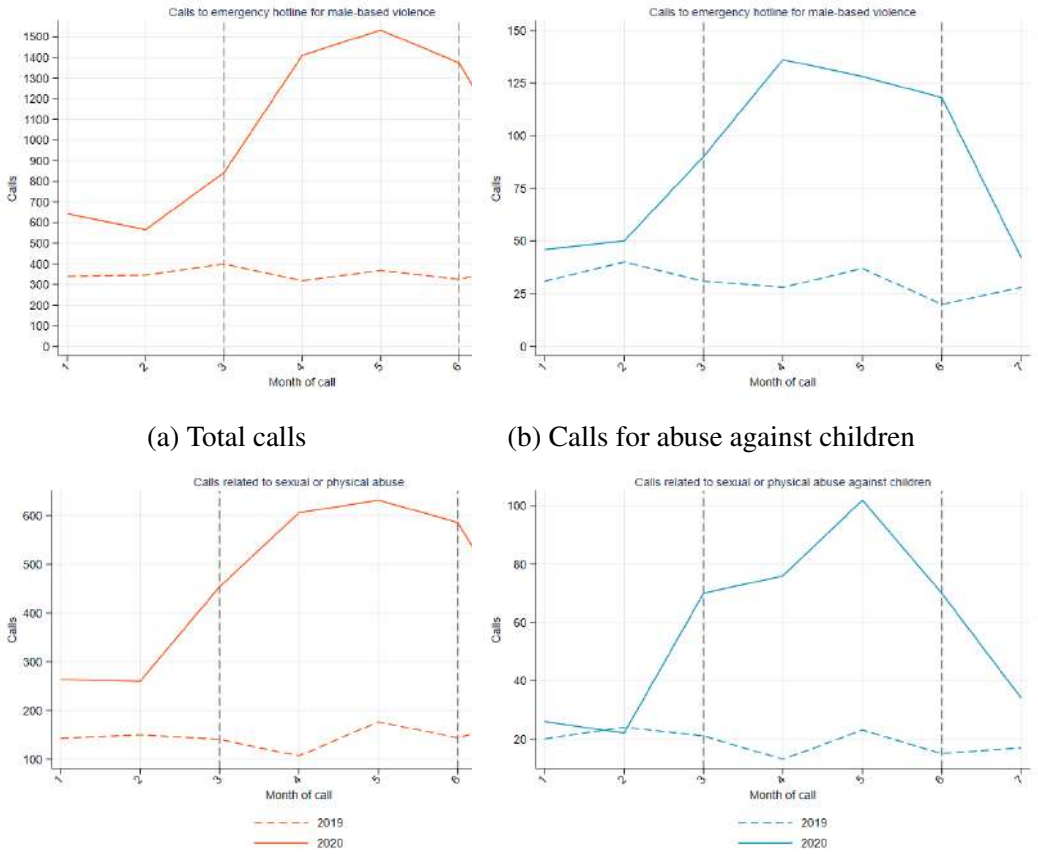


Table 5.A3: Descriptive statistics on abuses reported to the Catalan emergency hotline

	2019	**2020**
Total abuses detected through calls	9,421	14,668
Abuse against children or young adults detected through calls	752	1,154
Whether caller is also the victim (%)	72.3%	74.5%
Unemployed (%)	8.4%	10.5%
Economically dependent (%)	26.4%	26.2%
Psychological abuse (%)	48.7%	48.2%
Sexual abuse (%)	3.5%	3.5%
Physical abuse (%)	37.9%	39.4%
Physical abuse against children (% of children)	50.1%	53.7%
Sexual abuse against children (% of children)	4.9%	5.6%
Psychological abuse against children (% of children)	96.7%	96.5%
Whether victim lives with the aggressor	47.7%	51.8%
Aggressor – partner (%)	60.2%	62.4%
Aggressor – ex partner (%)	33.6%	30.1%
Share of calls for child abuse redirected to police	13.9%	22.5%
Share of calls redirected to police	15.5%	18.3%

Source: Catalan Institute for Women

Figure 5.A6: Calls to the Catalan Emergency Hotline in 2019 (dotted line) and 2020 (full line)

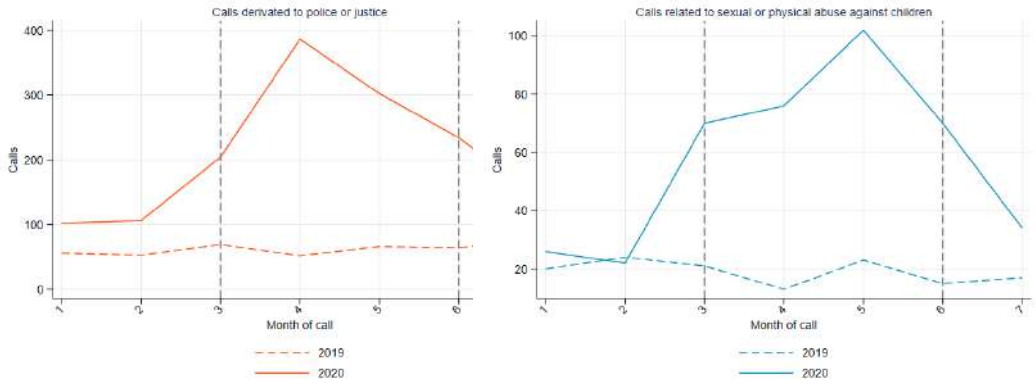


(a) Total calls

(b) Calls for abuse against children

(c) Calls for physical or sexual abuse, all victims

(d) Calls for physical or sexual abuse against children

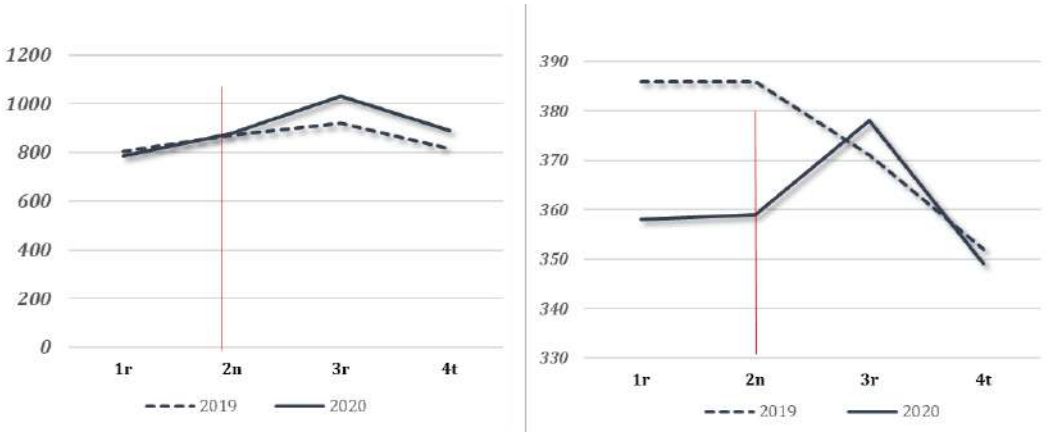


(e) Calls redirected to the Police, all victims

(f) Calls redirected to the Police, abuse against children

Source: Catalan Institute for Women

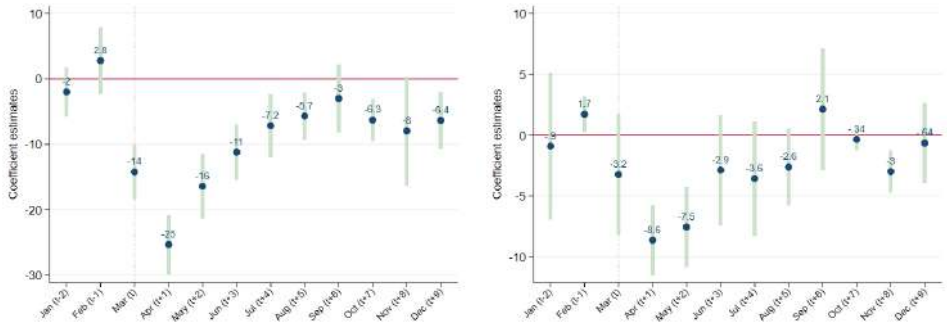
Figure 5.A7: Quarterly number of police reports for domestic violence against women (right) and young girls (left)



Source: Mossos d'Esquadra

5.A.2 Additional Results

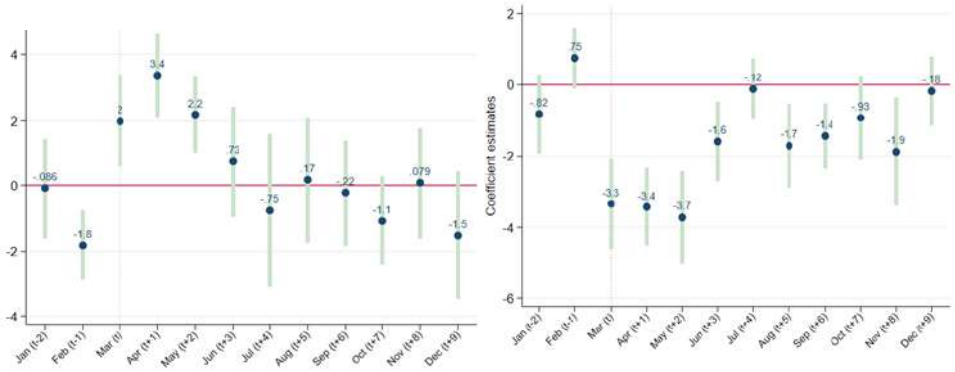
Figure 5.A8: Effect of lockdown on detection of male-based violence against women and children, event study estimates



(a) adult victims by 100,000 (b) <25 yo victims by 100,000

Source: Acquas, authors' own calculations

Figure 5.A9: Effect of lockdown on detection of male-based violence: heterogeneity analyses by subgroups



(a) Detection of victims of sexual abuse, relative to other victims (all)

(b) Detection of victims of lower income, relative to higher income (all)

Table 5.A4: Additional effect of lockdown on weekly detection of MBV per 100,000 (full sample), heterogeneity analyses

D_g	Victim is >25yo (1)	Abuse confirmed (2)	Entry by emergency (3)	Abuse from partner (4)	Sexual abuse (5)	Foreign born (6)	Low income family (7)
DiD:	-0.65^{+}	-0.28**	-0.57***	-0.49***	-0.60***	-0.43***	-0.16*
$D_{w\geq 11}D_{y>2019}$	(0.34)	(0.09)	(0.12)	(0.14)	(0.15)	(0.12)	(0.08)
Post:	0.49	0.50***	0.24	0.28^{+}	0.30	0.22	0.34*
$D_{w\geq 11}$	(0.35)	(0.15)	(0.17)	(0.16)	(0.21)	(0.15)	(0.14)
Treated:	-0.06	0.01	0.47**	0.03	-0.07	-0.05	-0.03
$D_{y>2019}$	(0.40)	(0.14)	(0.17)	(0.19)	(0.21)	(0.17)	(0.12)
DDD:	-0.05	-0.10	0.48**	0.33*	0.48**	0.18	-0.36*
$D_g D_{w\geq 11} D_{y>2019}$	(0.37)	(0.17)	(0.16)	(0.16)	(0.18)	(0.15)	(0.16)
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Week and year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	838	827	830	810	771	847	784
R^2	0.22	0.52	0.50	0.70	0.76	0.57	0.76

Notes: This table contains the triple differences estimates of the effect of lockdown introduction on the weekly cases of abuse detected through the healthcare system, by subgroups. the DDD estimate for the 1st specification represents the additional effect of the introduction of lockdown on detection of victims of less than 25 yo, respective to victims of 25 yo or more. The total effect of lockdown for the detection of abuse against victims of less than 25yo is equal to $DD + DDD$. ***p $< .01$; **p $< .05$; *p $< .1$

Table 5.A5: Loss in detection between March 2020 and January 2020, back of the envelope calculations

Patients	Victims detected b/w March and December 2019	Esti. change in detection Table 5.1 (1)	Esti. change in detection Table 5.1 (3) - log model	Loss in number of patients compared to 2019
All victims	9,013	-32%	-36%	-3,245
Victims of <25yo	2,164	-32%	-38%	-822
Victims of <14yo	896	-47%	-	-469

Table 5.A6: Impact of lockdown introduction on number of severe abuses reported to the Catalan emergency hotline

	Severe abuse all ages			Abuse against children and young adults	
Average marginal effects	(1) Total extreme abuse	(2) Sexual or Physical abuse	(3) Abuse derived to the police	(4) Total abuse	(5) Sexual or Physical
$D_{w\geq}$	0.15	0.14	0.00	-0.12	-0.05***
	(0.11)	(0.11)	(0.03)	(0.15)	(0.01)
$D_{y=2020}$	0.29***	0.32***	0.00	0.37***	0.07***
	(0.04)	(0.03)	(0.03)	(0.09)	(0.01)
A: $D_{y=2020}D_{w\geq 11}$	**0.85*****	**0.70*****	**0.33*****	**0.33****	**0.02****
	(0.14)	(0.12)	(0.04)	(0.11)	(0.01)
B: Mean	1.8	1.48	0.62	1.01	0.20
(AxB): Est. effect in %	47%	47%	53%	32%	10%
N	1552	1552	1552	1552	1552
Province FE	Yes	Yes	Yes	Yes	Yes
Week and year FE	Yes	Yes	Yes	Yes	Yes

Notes: Two-part estimation of the effect of lockdown introduction on weekly calls for abuse to the Catalan hotline. $D_{y=2020}$ is a dummy variable which is equal to 1 for the year 2020, and 0 otherwise; $D_{w\geq 11}$ is a dummy variable which is equal to 0 up to the quarantine implementation week (11th), and 1 up to the last week of the year. B corresponds to the average weekly calls per 100,000 between 2013 and 2019, calculated over the months of March to July.+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table 5.A7: Estimated increase in reported cases of extreme abuse between March and July through the hotline

Reported abuse	Number of calls b/w March and July average over 2013-2019	Effect of lockdown on reported abuse (%) (Table 5.A6)	Esti. increase in calls due to lockdown
All calls*	841	+47%	+395
Physical or sexual abuse	696	+47%	+327
Redirected to police	291	+53%	+154
Victime has <25yo	120	+32%	+38

Notes: *Calls included here refer to cases of severe abuse: either physical, sexual, redirected to the police, or against children

6 Concluding remarks

This thesis explore the causes and consequences of violence against minorities, with a particular focus on the expression of violence in the familiar setting. I use both historical and contemporary data and adopt a multifaceted approach: both in terms of the events of violence studied (institutional or interpersonal), the targeted minority (racial or gender), and the context (historical or contemporary).

The first section, composed of Chapters 2 and 3, studies violence during French Caribbean Slavery and its consequences for families after abolition in 1848. The work presented in this section is the fruit of a large scale digitization effort of handwritten administrative archives (nearly 50,000 pages) on enslaved individuals and their descendants. To collect this new data, I built an optical character recognition pipeline tailored for French Handwritten Archives, which I present in Chapter 4. The second section and final Chapter of my thesis turns to a contemporary context. I focus on state's responses to *systemic* violence against women and children during the COVID-19 pandemic (Chapter 5).

Both sections are complementary both in the *origins* of violence that are being investigated (institutions or exogenous stressful shock) but also in their approach to the problem of violence. While the first section focuses on the *mechanisms* through which violence has lingering effects, the second section focuses on the efficiency of institutional responses to intra-familial abuse. The health dimension is highly present in all chapters: either as consequence of violence, using mortality as outcome of interest in Chapter 2 and 3 [1]; or as a cause of violence - the COVID-19 health shock - and institution of interest (the health sector) in Chapter 5.

The first general contribution of my thesis lies in showing that collective, institutionalized violence against minorities can translate into large negative effects on survivors' descendants through trauma and child neglect. This mechanism, which

[1] I use mortality as a direct measure of violence during Slavery in Chapter 2, and child mortality as a measure of poor quality of childhood environment in Chapter 3

has long been discussed by psychologists in the case of slavery (Clark and Clark, 1950; Halloran, 2018), had yet to be explored empirically. The findings in this chapter highlight a significant effect of intergenerational transmission for inequality, underscoring the importance of further investment in this field. The second general contribution of this thesis lies in documenting important heterogeneity in violence exposure within targeted minorities. In Chapter 5, I find a much stronger effect of pandemic on the loss in detection of abuse - and therefore care provided to victims - among the most vulnerable groups: particularly children from low-income households. In Chapter 2, I find sizable differences in coercion intensity depending on planters' economic incentives, which I find reflected in substantial inequality among the next generation - through higher mortality rate of children of extremely coerced fathers. In both chapters, I document that children are highly vulnerable to violence against their parents: during slavery (Chapter 3), or against their mother during the contemporary context of the pandemic (Chapter 5).

The data collection efforts put into this thesis constitute valuable public goods. First, the novel data on enslaved population and their descendants contains precious information to approach the determinants of the experience of slavery, and its legacy. I will continue to enrich this dataset, and make it available to researchers and individuals. The OCR pipeline built to collect these valuable archives will also be released in open source. It contributes to scholars' effort in providing open source solutions for un-digitized text processing (Shen et al., 2021) in several ways. First, by creating an open-source Handwritten Data processing model, tailored from French Handwritten text. Second, by manually constructing a labeled dataset of handwritten French data, which can serve as crucial resource for other researchers looking to fine-tune their models. Third, by writing reproducible research notebook allowing for other researchers to reproduce our steps and train a TrOCR model for another language.

Bibliography

Abramsky, T., C. H. Watts, C. Garcia-Moreno, K. Devries, L. Kiss, M. Ellsberg, H. A. Jansen, and L. Heise (2011, February). What factors are associated with recent intimate partner violence? findings from the who multi-country study on women's health and domestic violence. *BMC Public Health 11*(1), 109.

Acemoglu, D., C. García-Jimeno, and J. A. Robinson (2012). Finding Eldorado: Slavery and long-run development in Colombia. *Journal of Comparative Economics 40*(4), 534–564.

Acemoglu, D., S. Johnson, and J. A. Robinson (2001, December). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review 91*(5), 1369–1401.

Acharya, A., M. Blackwell, and M. Sen (2018). *Deep Roots: How Slavery Still Shapes Southern Politics*. Princeton University Press.

Aizer, A. (2011). Poverty, Violence, and Health: The Impact of Domestic Violence During Pregnancy on Newborn Health. *Journal of Human Resources 46*(3), 518–538.

Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *The Quarterly Journal of Economics 115*(3), 715–753. Publisher: Oxford University Press.

Alacevich, C. and D. Zejcirovic (2020). Does violence against civilians depress voter turnout? evidence from bosnia and herzegovina. *Journal of Comparative Economics 48*(4), 841–865.

Alfred, M. (1935). Trois siècles d'histoire antillaise : Martinique et Guadeloupe de 1635 à nos jours. *Société de l'Histoire des Colonies Françaises*.

Althoff, L. and H. Reichardt (2023). Jim crow and black economic progress after slavery.

Anderberg, D., H. Rainer, J. Wadsworth, and T. Wilson (2016, November). Unemployment and Domestic Violence: Theory and Evidence. *The Economic Journal 126*(597), 1947–1979.

Arenas-Arroyo, E., D. Fernandez-Kranz, and N. Nollenberger (2021, February). Intimate partner violence under forced cohabitation and economic stress: Evidence from the COVID-19 pandemic. *Journal of Public Economics 194*, 104350.

Asik, G. A. and E. Nas Ozen (2021). It takes a curfew: The effect of covid-19 on female homicides. *200*, 109761.

Baron, E. J., E. G. Goldstein, and C. T. Wallace (2020). Suffering in silence: How COVID-19 school closures inhibit the reporting of child maltreatment. *190*, 104258.

Beigelman, M. and J. Vall Castelló (2020). COVID-19 and Help-Seeking Behavior for Intimate Partner Violence Victims. *SSRN Electronic Journal*.

Bensassi, S., P. Mohan, and E. Strobl (2017, October). A storm in a teacup hurricanes and sugar prices in the first half of the nineteenth century. *Weather, Climate, and Society 9*(4), 753–768.

Berniell, I. and G. Facchini (2021). Covid-19 lockdown and domestic violence: Evidence from internet-search behavior in 11 countries. *136*, 103775.

Bertocchi, G. and A. Dimico (2014). Slavery, education, and inequality. *European Economic Review 70*, 197–209.

Besley, T. and M. Reynal-Querol (2014). The legacy of historical conflict: Evidence from africa. *American Political Science Review 108*(2), 319–336.

Bhalotra, S., E. Brito, D. Clarke, P. Larroulet, and F. Pino (2024, February). Dynamic impacts of lockdown on domestic violence: Evidence from multiple policy shifts in chile. *The Review of Economics and Statistics*, 1–29. _eprint: https://direct.mit.edu/rest/article-pdf/doi/10.1162/rest_a_01412/2348510/rest_a_01412.pdf.

Bhalotra, S. and S. B. Rawlings (2011, April). Intergenerational persistence in health in developing countries: The penalty of gender inequality? *Journal of Public Economics 95*(3), 286–299.

Bindler, A. and N. Ketel (2022, October). Scaring or Scarring? Labor Market Effects of Criminal Victimization. *Journal of Labor Economics 40*(4), 939–970. Publisher: The University of Chicago Press.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2005). Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital. *American Economic Review 95*(1), 437–449.

Blattman, C. (2009). From violence to voting: War and political participation in uganda. *American Political Science Review 103*(2), 231–247.

Blattman, C., J. Hwang, and J. G. Williamson (2007, January). Winners and losers in the commodity lottery: The impact of terms of trade growth and volatility in the periphery 1870–1939. *Journal of Development Economics 82*(1), 156–179.

Bullinger, L. R., A. Boy, S. Messner, and S. Self-Brown (2021). Pediatric emergency department visits due to child abuse and neglect following COVID-19 public health emergency declaration in the southeastern united states. *21*(1), 401.

Cabrera-Hernández, F. and M.-a. Padilla-Romo (2020, February). Hidden Violence: How COVID-19 School Closures Reduced the Reporting of Child Maltreatment. *Working Papers*.

Campbell, J. C. (2002, April). Health consequences of intimate partner violence. *The Lancet 359*(9314), 1331–1336. Publisher: Elsevier.

Card, D. and G. B. Dahl (2011, February). Family Violence and Football: The Effect of Unexpected Emotional Cues on Violent Behavior*. *The Quarterly Journal of Economics 126*(1), 103–143.

Carrell, S. E. and M. L. Hoekstra (2010, January). Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids. *American Economic Journal: Applied Economics 2*(1), 211–228.

Clark, K. B. and M. P. Clark (1950). Emotional factors in racial identification and preference in negro children. *The Journal of Negro Education 19*(3), 341–350.

Constum, T., N. Kempf, T. Paquet, P. Tranouez, C. Chatelain, S. Brée, and F. Merveille (2022). Recognition and Information Extraction in Historical Handwritten Tables: Toward Understanding Early }}20^{th}}}Century Paris Census.

In S. Uchida, E. Barney, and V. Eglin (Eds.), *Document Analysis Systems*, Cham, pp. 143–157. Springer International Publishing.

Cottias, M. (1989). Mortalite et creolisation sur les habitations martiniquaises du XVIIIe au XIXe siecle. *Population (Paris) 44*(1), 55–84.

Cox, D. R. (1972, January). Regression Models and Life-Tables. *Journal of the Royal Statistical Society: Series B (Methodological) 34*(2), 187–202.

Currie, J., M. Mueller-Smith, and M. Rossin-Slater (2022, May). Violence While in Utero: The Impact of Assaults during Pregnancy on Birth Outcomes. *The Review of Economics and Statistics 104*(3), 525–540.

Darity, W. and D. Frank (2003). The economics of reparations. *The American Economic Review 93*(2), 326–329.

Darity, W., A. K. Mullen, and M. Slaughter (2022, May). The cumulative costs of racism and the bill for black reparations. *Journal of Economic Perspectives 36*(2), 99–122.

Davis, A. Y. (1983). *Women, Race & Class* (First Vintage books edition ed.). New York: Vintage Books.

Debien, G. (1974). *Les esclaves aux Antilles françaises (XVIIe et XVIIIe siècles). Basse-Terre et Fort-de-France*. Publication Title: Société d'histoire de la Guadeloupe et Société d'histoire de la Martinique.

Dell, M. (2010). The Persistent Effects of Peru's Mining Mita. *Econometrica 78*(6), 1863–1903.

Dell, M. and B. Olken (2019). The development effects of the extractive colonial economy: The dutch cultivation system in java. *Review of Economic Studies 87*(1), 164–203.

Dippel, C., A. Greif, and D. Trefler (2020, August). Outside options, coercion, and wages: Removing the sugar coating. *The Economic Journal 130*(630), 1678–1714.

Du Bois, W. (1908). The Negro American family. *Atlanta University Press*.

Dupraz, Y. and A. Ferrara (2023, February). Fatherless: The long-term effects of losing a father in the u.s. civil war. *Journal of Human Resources*, 0122–12118R2.

Durand, G. (2011). *Les Noms de Famille de la Population Martiniquaise d'Ascendance Servile – Origine et Signification des Patronymes portés par les Affranchis avant 1848 et par les Nouveaux Libres après 1848 en Martinique*.

Engerman, S. L. and B. W. Higman (2003). The demographic structure of the Caribbean slave Societies in the eighteenth and nineteenth Centuries. *General History of the Caribbean*, 45–104.

Fallope, J. (1983). Les esclaves africains à la Guadeloupe en 1848 d ' après les Archives de la Guadeloupe Les esclaves africains à la Guadeloupe en 1848 d'après les registres d' état civil des nouveaux citoyens conservés aux Archives de la Guadeloupe. *Bulletin de la Société d'Histoire de la Guadeloupe* ((57-58)), 3–25.

Fallope, J. (1987). Les occupations d'esclaves à la Guadeloupe dans la première moitié du XIXe siècle. *Revue française d'histoire d'outre-mer 74*(275), 189–205.

Fogel, R. W. and S. L. Engerman (1974). *Time on the Cross: The Economics of American Negro Slavery*.

Follett, R. J. (2005). *The Sugar Masters: Planters and Slaves in Louisiana's Cane World, 1820-1860*. LSU Press. Google-Books-ID: ul9xsFMcsmoC.

Gay Painter and David I. Levine (2000). Family Structure and Youths' Outcomes: Which Correlations are Causal? *The Journal of Human Resources 35*(3), 524–549.

Grosjean, P. (2014, October). A History of Violence: The Culture of Honor and Homicide in the US South. *Journal of the European Economic Association 12*(5), 1285–1316.

Grüning, T., G. Leifert, T. Strauß, J. Michael, and R. Labahn (2019, September). A Two-Stage Method for Text Line Detection in Historical Documents. *International Journal on Document Analysis and Recognition (IJDAR) 22*(3), 285–302. arXiv:1802.03345 [cs].

Halloran, M. J. (2018). African american health and posttraumatic slave syndrome: A terror management theory account. *Journal of Black Studies 50*(1), 45–65.

Harkness, L. L. (1993). Transgenerational transmission of war-related trauma. *International Handbook of Traumatic Stress Syndromes*, 635–643.

Heckman, J., R. Pinto, and P. Savelyev (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review 103*(6), 2052–2086.

Higman, B. W. (1975). The Slave Family and Household in the British West Indies, 1800-1834. *The Journal of Interdisciplinary History 6*(2), 261–287.

Higman, B. W. (1977, September). Slave population and economy in jamaica, 1807–1834. *The Journal of Economic History 37*(3), 813–815.

Hoeffler, A. (2017, November). What are the costs of violence? *Politics, Philosophy & Economics 16*(4), 422–445. Publisher: SAGE Publications.

Kalil, A., M. Mogstad, M. Rege, and M. E. Votruba (2016, October). Father presence and the intergenerational transmission of educational attainment. *Journal of Human Resources 51*(4), 869–899.

Kestenberg, J. S. (1980). Psychoanalyses of children of survivors from the holocaust: case presentations and assessment. *Journal of the American Psychoanalytic Association 28*(4), 775–804. Place: United States.

Koenig, M. A., R. Stephenson, R. Acharya, L. Barrick, S. Ahmed, and M. Hindin (2010, June). Domestic violence and early childhood mortality in rural India: evidence from prospective data. *International Journal of Epidemiology 39*(3), 825–833.

Koppensteiner, M. F. and L. Menezes (2021, July). Violence and Human Capital Investments. *Journal of Labor Economics 39*(3), 787–823. Publisher: The University of Chicago Press.

Laloux, L. (2019). La bataille du sucre ou la défaite méconnue de Napoléon Ier. *Artefact* (9), 35–56. ISBN: 9782810706235.

Laudares, H. and F. V. Caicedo (2023). Tordesillas, Slavery and the Origins of Brazilian Inequality. *CEPR Press Discussion Paper No. 18245*.

Lee, S.-H. (2005, June). Demand for Immunization, Parental Selection, and Child Survival: Evidence from Rural India. *Review of Economics of the Household 3*(2), 171–196.

Leslie, E. and R. Wilson (2020, September). Sheltering in place and domestic violence: Evidence from calls for service during COVID-19. *Journal of Public Economics 189*, 104241.

Levav, I., R. Kohn, and S. Schwartz (1998). The psychiatric after-effects of the holocaust on the second generation. *Psychological Medicine 28*(4), 755–760.

Li, M., T. Lv, J. Chen, L. Cui, Y. Lu, D. Florencio, C. Zhang, Z. Li, and F. Wei (2022, September). TrOCR: Transformer-based Optical Character Recognition with Pre-trained Models. arXiv:2109.10282 [cs].

Liu, L., R. L. Strawderman, M. E. Cowen, and Y.-C. T. Shih (2010, January). A flexible two-part random effects model for correlated medical costs. *Journal of Health Economics 29*(1), 110–123.

McLanahan, S., L. Tach, and D. Schneider (2013, July). The Causal Effects of Father Absence. *Annual Review of Sociology 39*, 399–427.

Miller, A. R., C. Segal, and M. K. Spencer (2020). Effects of the COVID-19 Pandemic on Domestic Violence in Los Angeles. *NBER Working Paper 28068*.

Miller, M. C. (2018). Destroyed by slavery? slavery and african american family formation following emancipation. *Demography 55*(5), 1587–1609.

Mulot, S. (2013, November). La matrifocalité caribéenne n'est pas un mirage créole. *L'Homme. Revue française d'anthropologie* (207-208), 159–191.

Mulot, S., M. Messu, M. Modak, G. Neyrand, and G. Séraphin (2023, March). L'évolution des familles guadeloupéennes de la matrifocalité vers la coparentalité : résistances, adaptations, créolisations.:. *Recherches familiales n° 20*(1), 135–150.

Naidu, S. and N. Yuchtman (2013, February). Coercive Contract Enforcement: Law and the Labor Market in Nineteenth Century Industrial Britain. *American Economic Review 103*(1), 107–144.

Norris, S., M. Pecenco, and J. Weaver (2021, September). The Effects of Parental and Sibling Incarceration: Evidence from Ohio. *American Economic Review 111*(9), 2926–2963.

Nunn, N. (2008). Slavery, Inequality, and Economic Development in the Americas: An Examination of the Engerman-Sokoloff Hypothesis. *Cambrige: Havard University Press*, 148–180.

Nunn, N. and L. Wantchekon (2011, December). The Slave Trade and the Origins of Mistrust in Africa. *American Economic Review 101*(7), 3221–3252.

Pamuk, S. (2007, December). The Black Death and the origins of the 'Great Divergence' across Europe, 1300–1600. *European Review of Economic History 11*(3), 289–317.

Patterson, O. (1982). *Slavery and Social Death: A Comparative Study*. Cambridge, MA: Harvard University Press.

Peterson, C., M. V. Aslam, K. L. Rice, N. Gupta, and M. C. Kearns (2024, February). Systematic Review of Per Person Violence Costs. *American journal of preventive medicine 66*(2), 342–350.

Piketty, T. (2020, October). *Capital and Ideology*. Harvard University Press.

Régent, F. (2010, December). Approche démographique de la population servile en Guadeloupe à la fin de l'ancien régime:. *Annales de démographie historique n° 119*(1), 5–25.

Sabia, J. J., A. K. Dills, and J. DeSimone (2013, May). Sexual Violence against Women and Labor Market Outcomes. *American Economic Review 103*(3), 274–78.

Schnakenbourg, C. (1968). Note sur les origines de l'industrie sucrière en Guadeloupe au XVIIe siècle (1640-1670). *Revue française d'histoire d'outremer 55*(200), 267–315.

Schnakenbourg, C. (1977). Statistiques pour l'histoire de l'économie de plantation en Guadeloupe et Martinique (1635-1835). *Bulletin de la Société d'Histoire de la Guadeloupe* (31), 3.

Schnakenbourg, C. (1987). La disparition des habitation-sucreries en Guadeloupe (1848-1906). Recherche sur la désagrégation des structures préindustrielles de la production sucrière antillaise après l'abolition de l'esclavage. *Revue française d'histoire d'outre-mer 74*(276), 257–309.

Schneider, E. B. (2017, February). Children's growth in an adaptive framework: explaining the growth patterns of American slaves and other historical populations: GROWTH PATTERNS OF AMERICAN SLAVES. *The Economic History Review 70*(1), 3–29.

Shen, Z., R. Zhang, M. Dell, B. C. G. Lee, J. Carlson, and W. Li (2021, June). LayoutParser: A Unified Toolkit for Deep Learning Based Document Image Analysis.

Smith, R. T. (1982). Family, social change and social policy in the west indies. *Nieuwe West-Indische Gids / New West Indian Guide 56*(3/4), 111–142.

Takaku, R. and I. Yokoyama (2021). What the COVID-19 school closure left in its wake: Evidence from a regression discontinuity analysis in japan. *195*, 104364.

Trevon D. Logan, J. B. P. (2018). On the marital status of u. s. slaves: Evidence from touro infirmary, new orleans, louisiana. *Explorations in Economic History* (69), 50–63.

Tur-Prat, A. and F. V. Caicedo (2020). The long shadow of the spanish civil war.

Vanony-Frisch, N. (1987). Les esclaves de la Guadeloupe à la fin de l'Ancien Régime d'après les sources notariales (1770-1789). *Bulletin de la Société d'Histoire de la Guadeloupe* (63-64), 3.

Victor Schoelcher (1842). *Des Colonies Françaises : Abolition Immédiate de l'Esclavage*. Paris.

Villeneuve, R. (1960). Le financement de l'industrie sucrière en France, entre 1815 et 1850. *Revue d'histoire économique et sociale (R.H.E.S.) XXXVIII*(3), 285–319.

Ward, A. J. R. (1978). The profitability of sugar planting in the british west indies, 1650-1834. *The Economic History Review 31*(2), 197–213.

World Health Organization (2018). Violence against women prevalence estimates.

World Health Organization (2020). Covid-19 and violence against women.

Zhao, D., L. Zou, X. Lei, and Y. Zhang (2017, August). Gender Differences in Infant Mortality and Neonatal Morbidity in Mixed-Gender Twins. *Scientific Reports 7*(1), 1–6.