Title

Essays in Economic Development and Health Economics

Author

Sharma, Manasvi

Publication Date

2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Essays in Economic Development and Health Economics

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Economics

by

Manasvi Sharma

Dissertation Committee:
Professor Priyaranjan Jha, Co-Chair
Assistant Professor Vellore Arthi, Co-Chair
Assistant Professor Meera Mahadevan

2024

Dedication

To my family and friends who inspired me to pursue this journey of learning, unlearning, and relearning.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I am deeply indebted to Professor Priyaranjan Jha, my committee co-chair, whose guidance and support were instrumental in making this dissertation achievable. Thank you for being so enthusiastic and continually motivating me to strive for excellence. I appreciate Professor Vellore Arthi, committee co-chair, for graciously dedicating time and offering valuable comments to improve my work. Our discussions were instrumental in fostering creative thinking and offering a well-rounded critique. I am also thankful to Professor Meera Mahadevan for her constant support, encouragement, and detailed feedback throughout.

Chapter 1 of this dissertation is a reprint of the material as it appears in Sharma, M. (2024). Ethnic fertility and exposure to armed conflict: the case of Sri Lanka. Rev Econ Household https://doi.org/10.1007/s11150-024-09703-y, used with permission from Springer Nature. I want to thank the Center for Global Peace and Conflict Studies, UCI for providing the grant that facilitated the development and publication of Chapter 1. Furthermore, I also appreciate the Department of Economics at UCI for providing various platforms to present my work and receive feedback, as well as for the summer research fellowships that enabled me to continue my dissertation work.

I am grateful to my family whose unwavering support has been my source of strength throughout this journey. A heartfelt thanks to my husband and cohort mate, Karan Talathi whose enduring love and dedication have been my guiding light. Last but not least, I am thankful for my friends who never failed to find humor and kept my spirits high.

Vita

Manasvi Sharma

EDUCATION

University of California, Irvine
Ph.D. in Economics *2024*
M.A. in Economics *2022*

Delhi School of Economics, University of Delhi, India
M.A. in Economics *2019*

Shri Ram College of Commerce, University of Delhi, India
B.A. (Hons.) in Economics *2017*

RESEARCH INTERESTS

Development Economics, Health Economics, Labor Economics, Economic Demography

PUBLICATIONS

- Ethnic Fertility and Exposure to Armed Conflict: The Case of Sri Lanka, Review of Economics of the Household, April 2024.

WORKING PAPERS

- Early Life Health Shocks and Intergenerational Mobility: The Impact of the 1848 Public Health Act in England and Wales

- Social Mobility in the Nineteenth-Century England and Wales

WORKS IN PROGRESS

- Females Over Males in Politics? Evidence from Disparity in Regional Economic Performance in India (with Karan Talathi and Rekha Sharma)

- Dietary Diversification and Food Security in India: An Empirical Analysis of Determinants (with Aishwarya Chakravarty and Aakash Khanna)

TEACHING EXPERIENCE

Teaching Assistant, *UC Irvine* *2020 - Present*

- Probability and Statistics in Economics II *Summer 2020, Winter 2023*
- Basic Economics I *Winter 2022, Spring 2022, Spring 2023*
- Basic Economics II *Winter 2024*
- Intermediate Economics I *Fall 2020*
- Intermediate Economics III *Spring 2021, Summer 2021, Fall 2021, Spring 2024*
- Applied Econometrics I *Summer 2020*
- Tax Policy *Winter 2021*
- Intermediate Quantitative Economics I *Fall 2022*

Head Teaching Assistant, *UC Irvine*

- Basic Economics I *Fall 2023*

RELEVANT EXPERIENCE

- Summer Associate Intern**,** *Analysis Group (Los Angeles)* *Summer 2023*
- Research Assistant, *Prof. Vellore Arthi (UC Irvine)* *Summer 2022*
- Intern, *Inspire Analytics Ltd. (New Delhi, India)* *Summer 2017*
- Intern, *The Wordsworth Project (New Delhi, India)* *Winter 2015*
- Research Assistant, *Prof. Surender Kumar (Delhi School of Economics)* *Summer 2015*

PROFESSIONAL AND PEDAGOGICAL DEVELOPMENT

- Student Success Mentor Training, Mentor Collective *2023*
- DE&I Foundations Training, Mentor Collective *2023*
- Public Speaking Course "Activate to Captivate" *2023*
- Teaching Assistant Professional Development Program *2020*

FELLOWSHIPS AND AWARDS

- Summer Research Fellowship, *UC Irvine* *2021 – 2023*
- Graduate Dean's Recruitment Fellowship, *UC Irvine* *2019*
- Delhi University Post Graduate Merit Scholarship *2018*
- Suniti Goyal Gold Medal, Topper in Girls - B.A.(Hons) Economics *2017*
- First Position in B.A(Hons.) Economics Third Year *2017*
- Third Position in B.A(Hons.) Economics Second Year *2016*

GRANTS

- Center for Global Peace and Conflict Studies Research Grant, *UC Irvine* *2021, 2022*

CONFERENCES, SEMINARS, AND POSTER PRESENTATIONS

- All-UC Demography Conference, Graduate Poster Presentation *2023*
- Center for Global Peace and Conflict Studies, Graduate Student Conference *2022, 2023*
- Graduate Poster Presentation, UC Irvine *2021, 2022*
- Graduate Student Applied Micro Workshop, UC Irvine *2020*

POSITIONS OF RESPONSIBILITY

- Peer Mentor, UCI Next Gen Pathways by Mentor Collective *Summer 2023 - Present*
- Member, National Service Scheme, Shri Ram College of Commerce *2014 - 2016*
- Campus Ambassador, Rio+22 Power India Certification Program, IARC *2015*

SOFTWARE SKILLS

- Stata, MATLAB, MS Office, QGIS

Abstract of the Dissertation

Essays in Economic Development and Health Economics

By

Manasvi Sharma

Doctor of Philosophy in Economics

University of California, Irvine, 2024

Professor Priyaranjan Jha, Co-Chair

Assistant Professor Vellore Arthi, Co-Chair

The three chapters in this dissertation examine aspects of economic development and health outcomes. In the first chapter, I investigate the impact of exposure to armed conflict on fertility in Sri Lanka. Using a difference-in-difference methodology, I find that exposure to civil war led to a reduction in female fertility in Sri Lanka, with evidence of an increased female age at marriage in high-conflict districts as a mechanism. The paper further focuses on ethnic disparities in demographic adjustments triggered by exposure to conflict. It determines if conflict altered the fertility patterns of the Sinhalese majority and the Sri Lankan Tamil minority differently. Estimates suggest that there is a differential in fertility adjustments of the two ethnic groups in response to conflict: the reduction in crude birth rate was significantly smaller for the Sri Lankan Tamils compared to the Sinhalese across various model specifications. The presence of an ethnic group-level replacement effect led to a lesser reduction in fertility for Sri Lankan Tamils. These results contribute to the literature on the impact of armed conflict and underscore the importance of studying demographic adjustments by sub-groups, specifically ethnicity in this context, as the intensity of adjustment often varies with the socio-political vulnerability of the group. Understanding these disparities is crucial as a sustained demographic differential has the potential to impact the ethnic composition of Sri Lanka and may further crystallize the ethnic divide in an already volatile political setting.

The second chapter examines the impact of early childhood health shocks on intergenerational mobility, in the historical context of the 1848 Public Health Act of England and Wales. The Act made the state the guarantor of public health and environmental quality for the first time in England and Wales, introducing sanitation measures for districts with a mortality rate of 23 deaths per 1000

people or more. Preliminary analysis reveals that the Act led to a reduction in mortality, yet solely relying on improvements in a crude measure such as mortality may overlook intrinsic changes in population health and abilities. To address this, the study employs a novel dataset to analyze the impact of shocks to the health environment in early childhood on a comprehensive measure such as intergenerational mobility. Using a difference-in-differences strategy, the analysis reveals that sons exposed to the act during early childhood were 5% more likely to pursue a different occupation than their fathers and 16% more likely to be in a better-ranked occupation than their fathers. This was driven by a transition of sons away from farming and unskilled jobs into skilled/semiskilled jobs. Furthermore, sons with early childhood exposure to the Act were more likely to be in occupations that required literacy, indicative of acquiring higher cognitive abilities. Evidence also suggests that spatial mobility played a role in these outcomes. The results emphasize the enduring effects of early childhood health shocks and thus underscore the role of public health interventions in shaping economic opportunities across generations.

The third chapter explores social mobility in nineteenth-century England and Wales, leveraging a large dataset linking the full count population censuses of 1851 and 1881. By employing a comprehensive dataset, it addresses previous data limitations and contributes revised estimates of historical occupational mobility, providing updated figures for both intergenerational and intragenerational mobility. New sample estimates of absolute intergenerational mobility suggest that 45.60% of the sons worked in a different occupational category than their fathers. The study also computes relative intergenerational mobility estimates using Altham statistics. International comparisons indicate England and Wales had lower intergenerational mobility than several New World economies. Evaluating total intragenerational mobility, the paper finds evidence of career stability among men, with more than 50% of the men engaging in the same occupational category 30 years later. This paper does not find evidence of an open and mobile society in nineteenth-century England and Wales.

Introduction

In the first chapter of this dissertation, I estimate the effect of exposure to armed conflict on female fertility in Sri Lanka. Given the increasing global instability, comprehending how war shapes demographic shifts is essential. While past research offers varied findings regarding the relationship between conflict exposure and fertility, the effects of the Sri Lankan civil war in this regard remain understudied.

Utilizing the Uppsala Conflict Data Program Georeferenced Event Dataset (UCDP GED) on conflict events, I classify districts above the mean of the total number of conflict events during 1992-2009 as treatment districts (high-conflict zone) and the others as control (low-conflict zone). Employing vital statistics data, I calculate crude birth rates as a metric for fertility. Using a difference-in-difference methodology, I estimate that on average, fertility declined by 5.42 births per 1000 people in the high-conflict districts of Sri Lanka. Next, I explore the possibility of increasing female age at marriage as a mechanism. Estimates suggest that the female age at marriage increased by 2 years in the high-conflict areas, implying that a higher age at marriage would delay fertility and thus result in the observed decline in the birth rates.

The paper further examines ethnic disparities in demographic adjustments triggered by exposure to the conflict. The historical backdrop of the Sri Lankan civil war underscores the significant role ethnicity played in shaping its dynamics. The rift between the Sri Lankan Tamil minority and the Sinhalese majority evolved over time, culminating in hostility between the two communities and the eruption of war in 1983. Consequently, examining diversity across ethnic groups is instrumental in comprehending their sensitivities and their capacity to adapt their fertility preferences or reactions to such upheavals, offering valuable insights for policymakers.

The difference-in-difference estimates provide evidence of a differential in fertility adjustments of the two ethnic groups in response to the conflict. Across various specifications, the results suggest that the Sinhalese reduced their fertility by 8-10 live births per 1000 people while Sri Lankan Tamils reduced theirs by 0-4 live births per 1000 people in response to exposure to armed conflict. This implies that the Sinhalese had a more pronounced fertility reduction. Additionally, using data on conflict-specific mortality from the UCDP database, the study finds the presence of a replacement effect among the Sri Lankan Tamils, but an absence of such an effect for the Sinhalese. These results contribute to the literature on the impact of armed conflict and underscore the importance of studying demographic adjustments by sub-groups, specifically ethnicity in this context, as the intensity of adjustment often varies with the socio-political vulnerability of the group.

The second chapter examines the impact of early-life health shocks on intergenerational mobility. I study this in the historical context of the Public Health Act of 1848 in England and Wales which was the first effort of the state to take charge of public health and environmental quality. This legislation marked a pivotal moment in public health history, addressing the dire sanitary conditions prevalent in the early nineteenth century, characterized by high mortality rates and rampant diseases, especially among children. The population of England and Wales had doubled in the five decades preceding the act, but there was inadequate expansion in the infrastructure to support this population. There was a lack of shelter, water supply, and drainage. These unlivable conditions came to attention in the 1830s as waves of the cholera epidemic and disease-led mortality crippled England and Wales. The poor state of public health was a pressing concern that required government intervention. In Aug 1848, the parliament passed the Public Health Act). It established the General Board of Health as the central authority to ensure the execution of the Act, comprising a President and two other board members. It was established for five years after passing the Act and continued annually thereafter until 1858 when the board ceased to exist. The Act legislated that the supply of water, sewerage and drainage, cleansing, and paving be under the supervision of a local body. Thus, a Local Board of Health was to be compulsorily formed if a place had high mortality (equal to or more than 23 deaths per 1000 people).

The paper employs a difference-in-differences methodology to first evaluate gains in mortality due to the implementation of the water, sanitation, and hygiene (WASH) measures in districts that were eligible for treatment. Using newly digitized historical vital statistics data from 1841-1880, I compute the seven-year average mortality rate of all districts in 1848 and identify districts that are eligible for treatment outlined by the Act. While this does not provide the actual location of the Local Boards of Health, it is a proxy for their locations. A potential issue with a binary classification of districts into treatment and control districts is that due to the contagious spread of diseases, a district's environment can be influenced indirectly by the adoption of the Public Health Act in neighboring districts. Additionally, 'yardstick competition' posits that public goods provision can spill over into neighboring regions as governments may mimic positive behaviors. To mitigate these considerations, I account for contiguous non-treatment districts, identified as spillover districts, in the main regression specification. In line with previous literature, I find evidence of a 4.05% mortality reduction in districts that were eligible for treatment under the Act, highlighting the effectiveness of the Act in improving public health. However, given the limitations of mortality as an indicator of health improvements, the paper explores intergenerational mobility as a more comprehensive measure.

The study assesses occupational intergenerational mobility as a long-run labor market outcome by analyzing a novel dataset of father-son pairs from the 1851 and 1861 censuses. I generate a primary measure of intergenerational mobility using self-reported occupations in the census, determining whether the son is in a better-ranked occupation than his father. To employ the difference-in-difference methodology, I utilize two key sources of variation – the variation in district eligibility for the Act and the age at exposure to the Act. The paper emphasizes the effect of early childhood exposure since access to the WASH measures in the first 1000 days of life (i.e., early life) is crucial. This is when 80% of the child's brain develops, forming the foundation for cognitive and psychosocial functions. Therefore, access to WASH in this phase reduces the likelihood of suffering from a vicious cycle of diarrhea and undernutrition. This implies the child is less likely to be stunted and have better brain and physical development—therefore, higher cognitive abilities that translate into higher skilled jobs. The analysis reveals that sons exposed to the Act during early childhood in treatment districts were more likely to experience upward mobility relative to their counterparts in control districts. This upward mobility is predominantly driven by transitions from farming and unskilled occupations to skilled/semiskilled occupations. Further evidence suggests that these sons were also more likely to be engaged in jobs that required literacy, indicating enhanced human capital accumulation. They are also more likely to migrate out of their childhood districts, suggesting that spatial migration was a mechanism for the upward mobility of sons with early childhood exposure to the Act.

The results emphasize the enduring effects of early childhood health shocks and thus underscore the role of public health interventions in shaping economic opportunities across generations.

The third chapter delves into the trends in social mobility in nineteenth-century England and Wales. It examines how individuals are positioned within societal hierarchies, and the transmission of occupation, class, and earnings across generations (intergenerational mobility) as well as throughout their lives (intragenerational mobility or career mobility). While the optimistic Smilesian perspective portrays Victorian Britain as a land of boundless opportunity, others argue the existence of enduring social barriers.

To measure social mobility in this historical context, the paper utilizes a newly linked dataset of fathers and sons from the full count population censuses of 1851 and 1881. Unlike prior studies that relied on partial census data, this dataset offers a significantly larger linked sample, enabling more comprehensive analysis. Intergenerational mobility is estimated using two metrics: absolute and relative. Absolute mobility reflects the proportion of sons in different occupational categories than their fathers. Using an occupational transition matrix between fathers and sons, I find that 45.60% of the sons in this sample worked in a different occupation than their fathers. To systematically compare mobility between transition matrices (to compare estimates for England and Wales with prior literature or to conduct international comparisons), literature has relied upon the Altham Statistic, which is a single metric that summarizes the mobility difference and can be tested for statistical significance. Using the Altham statistic, I estimate the relative intergenerational mobility in England and Wales and conduct international mobility comparisons.

In addition to intergenerational mobility, the paper also examines intragenerational or career mobility, shedding light on the stability of occupational trajectories over time. Contrary to some previous findings, results suggest that men exhibited considerable career stability in nineteenth-century England and Wales, with a majority remaining in the same occupation for three decades.

Overall, this paper enhances the understanding of historical social mobility in England and Wales by addressing measurement errors in previous studies caused by unrepresentative samples due to limited data availability. Utilizing the complete 1851 census dataset provides a more robust analysis of mobility, enriching our comprehension of the historical social mobility dynamics.

Chapter 1: Ethnic Fertility and Exposure to Armed Conflict: The Case of Sri Lanka[1]

1.1 Introduction

Wars have far-reaching effects on various aspects of politics, the economy, and society. On a direct level, wars lead to morbidity, mortality, the destruction of infrastructure and healthcare systems, and the displacement of populations. However, their repercussions often extend beyond the immediate aftermath. According to Catani (2018), war-affected populations may experience long-term psychological stressors. Studies by Devakumar et al. (2014) and Phadera (2021) highlight intergenerational health consequences, indicating that the impact goes deeper and persists across generations.

With rising global instability, understanding the impact of war on demographic changes is crucial. Previous studies have found mixed evidence on the effect of conflict exposure on fertility. For instance, Lindstrom and Berhanu (1999), Agadjanian and Prata (2002), and Blanc (2004) provide evidence of a reduction in fertility due to conflict. On the other hand, Kulczycki and Prem (1999) and Ladier-Fouladi and Hourcade (1997) find no overall effect, while Kraehnert et al. (2018) identify evidence of increased fertility resulting from conflict. Despite a growing body of literature on the impact of war on fertility, the consequences of the Sri Lankan civil war remain relatively unexplored in this context.

[1] I am grateful to Priyaranjan Jha, Vellore Arthi, Damon Clark, Meera Mahadevan, and Tim Bruckner for their valuable feedback. I would also like to thank the participants of the Graduate Student Applied Microeconomics Workshop, UCI, and the Annual Center for Global Peace and Conflict Studies Conference, UCI. This study was funded by the Center for Global Peace and Conflict Studies, UCI. All errors are my own.

This paper aims to explore the effect of exposure to armed conflict on female fertility in Sri Lanka and highlights the role of ethnicity in this context. The 26-year-long civil war in Sri Lanka (1983-2009) is considered one of the world's most violent and prolonged secessionist movements since the end of the Second World War (DeVotta, 2004). Using the Uppsala Conflict Data Program Georeferenced Event Dataset (UCDP GED) on conflict events, I classify districts above the mean of the total number of conflict events during 1992-2009 as treatment districts (high-conflict zone) and the others as control (low-conflict zone). Using vital statistics data, I compute crude birth rates as a measure of fertility. Employing a difference-in-difference approach, estimates suggest that on average fertility declined by 5.42 births per 1000 people in the high-conflict districts of Sri Lanka. Moreover, the study suggests that an increase in female age at marriage in these areas may contribute to this observed fertility decline.

This paper emphasizes that, as one of the understudied aspects of heterogeneity in conflict literature, the effects of exposure to armed conflict differ across ethnic groups. The history of the Sri Lankan conflict suggests that ethnicity had a major role in shaping the dynamics of the conflict. The divide between the Sri Lankan Tamil minority and the Sinhalese majority percolated through the years and ultimately manifested into intolerance between the two communities, resulting in war in 1983. Consequently, exploring heterogeneity across ethnic groups is valuable in understanding their sensitivities and their ability to adjust their fertility preferences or responses to such shocks, which is further useful to guide policymakers.

The difference-in-difference estimates in the study provide evidence of differential fertility adjustments of the two ethnic groups in response to the conflict. Across various specifications, the estimates suggest that the Sinhalese reduced their fertility by 8-10 live births per 1000 people while Sri Lankan Tamils reduced theirs by 0-4 live births per 1000 people in response to exposure to armed conflict. This implies that the Sinhalese had a more pronounced fertility reduction. Additionally, using data on conflict-specific mortality from the UCDP database, this study finds evidence of a replacement effect among the Sri Lankan Tamils, contrasting with the absence of such an effect for the Sinhalese.

This paper contributes to a growing literature by evaluating the dynamics between conflict, ethnicity, and fertility outcomes. In line with recent studies (Schindler & Brück, 2011; Kraehnert et al., 2018; Thiede et al., 2020; Boehnke & Gay, 2020), this paper finds that changes in female age at marriage contribute to an overall reduction in fertility in Sri Lanka and evidence of ethnic group level replacement effects explaining a lower reduction in fertility for Sri Lankan Tamils. To the best of my knowledge, this is the first paper to focus on the effect of exposure to conflict on fertility in Sri Lanka.

The rest of the paper is organized as follows: Section 1.2 discusses related literature; Section 1.3 details the history of the ethnic divide and evolution of the conflict; Section 1.4 describes the data and the methodology; Section 1.5 discusses the overall effect in Sri Lanka and a related mechanism of increasing female age at marriage; Section 1.6 focuses on the ethnic group differential and the presence of replacement effect for Sri Lankan Tamils; Section 1.7 concludes.

1.2 Related Literature

In general, fertility behaviors are governed by a plethora of factors such as income, education, age, and other socio-economic factors. Bongaarts' model of the proximate determinants of fertility categorizes the set of factors affecting fertility levels into two broad categories: behavioral and biological (Bongaarts, 1987). Within this framework, three biological factors influence fertility decisions: permanent sterility, the probability of conception, and intrauterine mortality. Behavioral determinants, on the other hand, include factors like the prevalence of contraceptive use and effectiveness, duration of breastfeeding, practice of induced abortion, and patterns of union formation.

The literature on conflict identifies both biological and behavioral mechanisms through which armed conflict exposure can impact fertility. The four major pathways discussed in previous studies are income effect, marriage market effect, sexual violence against women, and replacement effect.

The income effect posits that factoring in the cost of raising children, coupled with a reduction in wealth during the war (Collier et al., 2003), can often result in reduced fertility. For instance, Kraehnert et al. (2018) found that the death of a woman's sibling during the civil war in Rwanda lowered the probability of giving birth. The study suggests that if the deceased person played a role in monetarily supporting the upbringing of the child, the income effect would explain the reduction in fertility.

Conflicts can also exert an impact on the marriage market and thus impact fertility. This occurs when the conflict results in a demographic deprivation of young men/husbands, either by reduced spousal cohabitation with men recruited and away at war (Blanc, 2004; Urdal & Che, 2013; Vandenbroucke, 2014) or by altering the sex ratio (Boehnke & Gay, 2020;, Kraehnert et al., 2018). Exposure to armed conflict can also result in a shift in marriage patterns due to a change in bride prices (Corno et al., 2017) or a change in the age at marriage. The increased uncertainty and absence of men could encourage postponement in marriage and thus delay marital fertility (Thiede et al., 2020). On the contrary, Hacker et al. (2010) also emphasize that in some situations, war can be a catalyst for marriages. Faust (1996) underscored that war resulted in a dread of “spinsterhood” in southern women during the American Civil War. Barber (2000) reported that the American Civil War had led to a “marriage frenzy” in Richmond.

Additionally, conflict may positively affect fertility rates due to increased sexual and gender-based violence against women by armed combatants or perpetrated by intimate partners (McGinn, 2000; Østby, 2016; Leatherman, 2011; United Nations, 2014; Wood, 2018). Kottegoda et al. (2008) provide accounts of increased vulnerabilities of women and higher intimate partner violence in Sri Lanka during the civil war. The repercussions of conflict extend to the accessibility of reproductive health care for women (Kottegoda et al., 2008; Urdal and Che, 2013; Lindskog 2016). On the contrary, conflict can also result in psychological stressors as well as poor nutritional status of the women and thus negatively impact fertility (Nowrojee, 1996; McKay, 1998; Abu-Musa et al., 2008; Dossa et al., 2015; Catani, 2018).

Finally, the effect of higher mortality on fertility is ambiguous (Schultz 1997, Ben-Porath 1976). The presence of a positive correlation suggests a replacement mechanism. Replacement effects can arise at an individual level if children are considered old-age insurance for the parent and an increase in mortality leads individuals to have more children in the expectation that only some of their kids will survive (Nugent, 1985; Rossi & Godard, 2019) or if people are trying to replace the loss of their loved ones (Schindler & Brück 2011, Kraehnert et al. 2018). At the group level replacement effects can arise if pronatalist sentiments emerge during conflicts among the minority community to maintain the size of their group in the population (Chi et al., 2015; Palmer & Storeng, 2016; Alburez-Gutierrez, 2019). In the Sri Lankan context, there may be a replacement effect in place for the Sri Lankan Tamils based on the qualitative findings of Jordan and Denov (2007) who report evidence of marriage between members of opposite genders being strongly encouraged amongst LTTE militants to inspire a greater sense of commitment and to ensure traditional Tamil customs are practiced in the larger society.

With many counteracting mechanisms, the net effect of conflict on fertility remains to be tested empirically for Sri Lanka.

1.3 The Sri Lankan Civil War

In 1981, Sri Lanka had a total population of approximately 14.8 million. The Sinhalese comprised 74% of the population, followed by Sri Lankan Tamils who were 12.7% of the population, and Indian Tamils who were 5.5% among other groups (Census of Population and Housing Report, 1981). Historically, the Sinhalese have inhabited the western, southern, and central districts, whereas the Sri Lankan Tamils were a majority in the northern and eastern parts of the country (Census of Population and Housing Report, 1971; Keethaponcalan, 2002). In line with that, the Census of Popuation and Housing Report (1981) finds that 65% of the Sri Lankan Tamils resided in the northern and eastern districts in 1981.

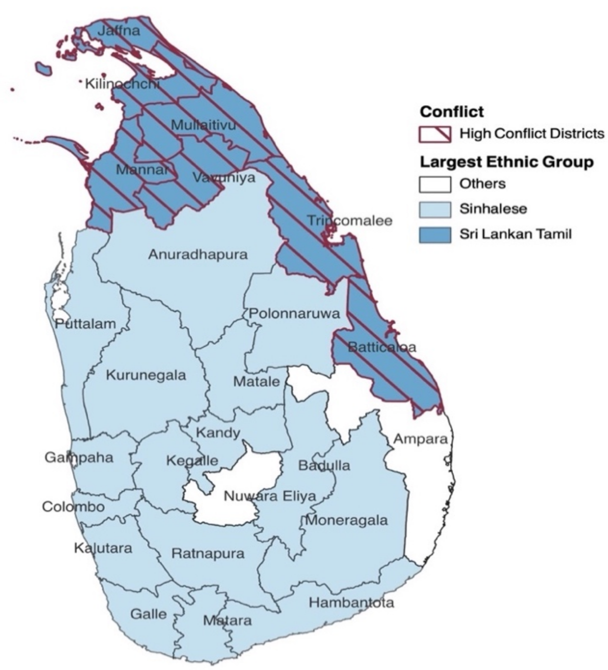

Figure 1.1: Map of Sri Lanka - Areas of High-Conflict and Largest Ethnic Group in Each District

Note: The classification of districts as high and low conflict is based on Appendix Table A.1 using the UCDP conflict dataset. Data for the size of ethnic groups by district is obtained from the Census of Population and Housing Report (1981) and is provided in Appendix Table A.2. This allows a graphical depiction of the largest ethnic group in each district, mapped with the conflict status of the district.

Keethaponcalan (2002) describes that the districts in the Northern and Eastern Provinces (Jaffna, Kilinochchi, Mullaitivu, Mannar, Vavuniya, Trincomalee, Batticaloa, and Ampara) were referred to as the "traditional homeland" by Sri Lankan Tamils. These areas were under the direct control of the Liberation Tigers of Tamil Eelam (LTTE) during the war and had to therefore witness a large amount of conflict between the LTTE and the Government of Sri Lanka (GoSL). [2] Figure 1.1 identifies the geographic variation in conflict intensity in the districts along with the dominant ethnic group in the district. The figure points to an overlap between the conflict zone and the dominant ethnic group in the district being the Sri Lankan Tamils. While there were some instances of conflict reported in almost all districts of Sri Lanka, they were largely concentrated in the Sri Lankan Tamil dominant districts (Appendix Table A.1).

Though the war began in 1983, the roots of the ethnic conflict in Sri Lanka go back to the former as Ceylon under British Rule. With the 'divide and rule' policy underway, there were tensions between the communities even before 1983. It is claimed that there was an over-representation of Tamils in the colonial administrative positions, owing to British favoritism. The British often constructed English medium schools in Tamil areas that allowed greater access for Tamils to it and government jobs at the time. Tamils comprised about 30 percent of upper-level civil service jobs while comprising 22.7 percent of the population as per the census in 1946. The Sinhalese comprised 69.4 percent of the population and 57.6 percent of the upper levels of civil service (Keethaponcalan, 2002). This over-representation was considered unjust by the Sinhalese.

Post-independence in 1948, the bulk of the power in government landed in the hands of the Sinhalese elite who resented Sri Lankan Tamils. They passed several laws that were detrimental to Tamil representation and interests. They disenfranchised Tamil migrant plantation workers from India, also known as the Estate Tamils or Indian Tamils, or Upcountry Tamils. This meant the Tamils lost 40% of their parliamentarians because Estate Tamils could not be represented under this law. The Sinhala Only Act of 1956 declared Sinhala as the only official language of Sri Lanka. This made it mandatory for Tamils in the public sector to qualify in Sinhala within a designated short period, a lack thereof meant losing their jobs. The Sinhalese considered it as a move to reclaim the true Sri Lankan identity that was lost during British Rule. However, to the Tamils, this was an attempt to limit their economic opportunities, as from 1956-1976 of the 189,000 people recruited in the public sector, 95% were Sinhalese. From the 1970s there was virtually no

[2] *Source:* https://samsn.ifj.org/northern-and-eastern-provinces-sri-lanka-legacy-of-conflict/

employment of Tamils in the armed forces and police which are 99.9% and 95% respectively Sinhalese. In 1972, the government changed the country's name from Ceylon to Sri Lanka and made Buddhism the nation's primary religion, emphasizing a sense of culture and national identity. In an attempt to bridge a gap between educational disparities in the country, the 'standardization policy' of 1972 required Tamil students to score higher marks than their Sinhalese counterparts to get admission into Sri Lankan Universities. This act was aimed at providing support to the previously disadvantaged Sinhalese youth. Hoping to undo the injustice to the Sinhalese under the British Era, standardization ended up creating a further rift between Sinhalese and Tamils. Demands for equal status by Tamils radicalized over the years and led to the formation of a militant organization called LTTE in 1976 under the leadership of Velupillai Prabhakaran. Their demands for self-government and an independent state called the Tamil Eelam waged war in 1983.

1.3.1 Eelam War-I (1983-1987)

The first of four Eelam Wars began in 1983 and lasted 4 years. While it was the northern and eastern provinces that were direct conflict zones under direct control of LTTE, there were some instances of violence reported across Sri Lanka. In July 1983- known as "black July"- riots broke out in Colombo and other cities. LTTE had killed 13 army soldiers. In retaliation by the Government of Sri Lanka (GoSL), about 2500-3000 Tamils likely died. Violence escalated over the years as Tamil insurgents used car bombs, human bombs, and landmines against Sinhalese civilians and the military. India tried to intervene by sending peacekeepers in 1987, but its troops had to withdraw by 1990 as Tamil guerillas refused to step back.

1.3.2 Eelam War-II (1990-1995)

After the withdrawal of the Indian Peace Keeping Forces, the war got more intense. LTTE killed about 600-700 Sinhalese police officers in June 1990. A female LTTE suicide bomber assassinated India's Prime Minister Rajiv Gandhi at an election rally in 1991. The government cut off shipments of food and medicines to the Jaffna peninsula and started intensive aerial bombardment. The LTTE responded by massacres of Sinhalese and Muslim villagers which led to Muslim self-defense units doing the same in Tamil villages.

1.3.3 Eelam War-III (1995-2002)

In January 1995, a peace agreement was signed with the new government. This however was violated after 3 months when LTTE destroyed ships and naval gunboats. This was followed by massacres in refugee camps in the Jaffna Peninsula and other civilian sites. By December, Jaffna was under government control for the first time. However, in July 1996, the LTTE carried an 8-day strike in Mullaitivu against 1400 government troops. More than 1200 government soldiers were killed and the LTTE also lost about 332 troops. Suicide bombers attacked Colombo, the Central Bank in Colombo, the Sri Lankan World Trade Center, and other southern cities repeatedly during the late 1990s.1999 also saw an attempt on President Kumaratunga's life by a bomber. Norway tried to negotiate a peace settlement in December 2000. After various rounds of negotiations, LTTE and GoSL signed a peace treaty with LTTE foregoing the demand for a separate Tamil state. However, in October 2003, GoSL declared a state of emergency when Tigers announced their full control over Northern and Eastern Sri Lanka. Norway reports about 300 violations of the ceasefire by the GoSL and 3000 by LTTE during that period.

1.3.4 Eelam War-IV (2006-2009)

Eelam War IV from 2006-2009 was the final war that culminated in peace, and GoSL declared victory. After the October 2006 peace talks in Geneva failed, the government launched a major attack against the LTTE. The next few years were called a bloodbath as thousands died on both sides. Ultimately in May 2009, the GoSL declared victory. The estimated death toll by the UN over the years is 75,000-100,000 people with many other victims of kidnappings, sexual abuse, and disabilities due to war.

1.4 Data

The dataset for this analysis was created by digitizing the data from the Annual Reports of Vital Statistics from 1967-1980 (Vital Statistics, 1967-1980) and 1992-2009 (Vital Statistics, 1992-2009) and Census Reports over the years 1963-2012 (Census of Population and Housing Report, 1963, 1971, 1981, 2001, and 2012). Registration data exists for 1981-1991 as well, but I am unable to obtain the vital statistics for 1981-1991. The data on vital statistics has information on births and deaths at the district level disaggregated by ethnicity. Appendix Table A.3 provides the descriptive statistics of the data for all ethnic groups in Sri Lanka. The registration system for births and deaths in Sri Lanka has been in place since 1867 and is of good quality (Levine, 2007; Jayachandran and Lleras-Muney, 2009) with completeness of at least 90% in most of the years of this analysis (United Nations Statistics Division, 1970-2010; United Nations Statistics Division, 2006). Assuming exponential population growth, I use the census year population by ethnic group for each district to interpolate the population for intercensal years.[3] Using the data on births at the ethnicity-district level, I create measures of crude birth rates. While crude rates have the drawback of not accounting for the age composition of the population, which may be consequential for the size of the next birth cohort, data availability restricts the analysis to using the said measure.

To create a panel dataset of female age at marriage, district-level data on pre-war years female singulate mean age at marriage (SMAM) has been collected for years 1953, 1971, and 1981 from the Census of Population and Housing Report (1981). The vital statistics dataset by GoSL provides information on female age at marriage for the post-war years of this analysis from 1992-2006 (Vital Statistics, 1992–2009). The data classifies females getting married into 3 categories based on their prior marital status: females who were never married, females who were divorcees, and females who were widows. For consistency in age at marriage information before and during the war, since SMAM reports mean age at first marriage, I use the information of female age at marriage for women who were previously never married so that this can mimic SMAM data in war years. This dataset is not available at disaggregated levels by ethnic groups.

I obtain conflict data from the UCDP GED database. UCDP is the world's major data source for organized violence with the oldest ongoing data collection project for civil wars. It defines an armed conflict as a contested incompatibility that concerns government and/or territory over which the use of armed force between the military forces of two parties, of which at least one is the government of a state, has resulted in at least 25 battle-related deaths each year. UCDP has global coverage and provides data on three types of organized violence: state-based conflict, non-state

[3] United Nations Population Fund report 'Population Situation Analysis: Sri Lanka', describes that the population of Sri Lanka has been growing exponentially overtime. *Source:* https://www.unfpa.org/sites/default/files/admin-resource/Sri-Lanka-Population-Situation-Analysis_0.pdf

conflict, and one-sided violence, both spatially and temporally. I aggregate conflict events at the district level. Appendix Table A.1 provides the number of conflict events aggregated at the district level over the years 1992-2009. The district with the greatest number of conflict events is Mullaitivu with 806 total conflict events reported. Jaffna experienced about 559 of all conflict events reported in the dataset. The districts with the lowest conflict events were Ratnapura and Kalutara which reported 0 conflict events as per the dataset. The dataset classifies each conflict event as organized violence with at least one fatality in either the best, low, or high estimate categories and provides its date and location (Sundberg & Melander, 2013). Using the best estimates of conflict-specific fatalities from the dataset, I obtain an estimate of conflict-specific death rates at the district-year level.

Since the boundary of districts changed over time, I adjust the data based on the area of the district in 2012 to allow for comparability over time and thus have 25 districts in this analysis. Wherever districts were split to create new ones, I adjust its population to be based on its share of the parent district's area. For instance, the Gampaha district with a 1387 km sq area was carved from the Colombo district which had an area of 2051 km sq area before the split. So, the population assigned to Gampaha before the split has been adjusted in proportion to the areas of the two districts (i.e., approximately 65% of the Colombo population before the split).

1.5 Overview: Impact in Sri Lanka

Few studies on Sri Lanka have examined trends in fertility (Abeykoon, 1987; Puvanarajan & DeSilva 2001), but none through the lens of the effect of armed conflict. To identify the causal effect of exposure to conflict on fertility, I estimate a difference-in-difference model with crude birth rate as the dependent variable. Using the information on conflict events, I classify districts above the mean of the total number of conflict events during 1992-2009 as treatment districts (high-conflict zone) and the others as control (low-conflict zone) in Appendix Table A.1. Alternatively, given that conflict per capita may be a more appropriate measure of conflict exposure, for each district I calculate the average of the lagged number of conflict events per capita per year. Using this measure of conflict per capita, I assign districts as treatment group if the conflict intensity was greater than the mean across districts and as control group if they are below the mean intensity. Both these classifications result in the same districts being assigned to treatment and control groups. Appendix Figure A.1 demonstrates parallel trends in the average crude birth rate for the treatment and control group in the pre-treatment years. Since the timing and duration of conflict are unpredictable, the treatment can be considered exogenous.

I estimate the following equation to identify the effect of armed conflict exposure on the crude birth rate in Sri Lanka:

$$br_{gdt} = \beta_0 + \beta_1 treated_d \times post_t + \theta_d + \delta_t + \epsilon_{gdt} \text{ .... (1)}$$

where *br* is the crude birth rate for ethnic group g in district d at time t , θ_d *and* δ_t represent district and time fixed effects respectively. *treated* is a dummy variable that takes a value of 1 for high-conflict districts and 0 for others, *post* is a dummy that takes a value of 1 for years after 1983 that mark the beginning of the civil war (i.e., for 1992-2009, due to data constraints I do not have data from 1983-1991).

Thus, *treated x post* is 1 for treatment districts that are the high-conflict districts for years after 1983 (i.e., for 1992-2009). β_1 in this setup is the relevant coefficient, measuring the impact of exposure to conflict on the crude birth rate in Sri Lanka. I estimate the model using robust standard errors (Castelló-Climent et al., 2018).[4]

Estimates from Table 1.1 column 1 and column 2 report the overall average effect of conflict on fertility in Sri Lanka. While Column 1 provides the unweighted estimates, column 2 provides district population weighted estimates.[5] On average, the effect of exposure to high conflict was a reduction of 3.96 to 5.42 births per 1000 people. The pre-conflict average crude birth rate in Sri Lanka was 27.95 and a reduction of approximately 3.96 to 5.42 5.42 in the crude birth rate is a moderate response of 14.16% - 19.39% fertility reduction compared to some other estimates in the literature. A Population Reference Bureau report (Madsen and Finlay, 2019) documents that fertility in conflict usually drops by up to one-third. For instance, it dropped by one-third due to the Khmer Rouge conflict in Cambodia in the 1970s. Curlin et al. (1976) find a similar response

[4] As a robustness check, columns 1 and 2 of Appendix Table A.9 report the estimates corresponding to columns 1 and 2 of Table 1.1 with clustered standard errors at the district level and estimated p-values from the wild cluster bootstrap recommended by Cameron et al. (2008) as a solution to the problem of a few clusters (Angrist & Pischke, 2009).

[5] Estimates with population weights are preferable as these account for heterogeneity in the population size across districts. Using population weights in this case ensures that small population sizes do not contribute disproportionately to the result. The rest of the paper will thus report weighted estimates.

of the age-specific fertility rate of women aged 10-14 in Bangladesh which fell by 33% due to conflict. A much larger reduction is documented by Vandenbroucke (2014) who finds that WWI led to a 50% reduction in France's birthrate.

Table 1.1: Impact of Conflict on Fertility

Dependent Variable: Crude Birth Rate				
	overall effect		heterogeneity by ethnicity	
	(1)	(2)	(3)	(4)
Independent Variable				
treated x post	-5.42***	-3.96***	-8.247***	-10.66**
	(1.131)	(0.963)	(1.159)	(0.864)
treated x post x Sri Lankan Tamil	-	-	4.072**	11.41***
			(1.879)	(1.657)
treated x post x Others	-	-	4.422*	7.15**
			(2.534)	(1.614)
N	2400	2400	2400	2400
R^2	0.326	0.319	0.536	0.694

Note: Treatment is assigned based on classification in Appendix Table A.1. The regression in column 1 reports the coefficient for equation 1, giving the effect of war on fertility in Sri Lanka. The regression includes district and year-fixed effects. Robust standard errors are reported in the parenthesis. Column 2 adds district population weights. Columns 3 and 4 correspond to estimates from equation 3, exploring heterogeneity by ethnic groups. Column 3 does not include district-ethnic group population weights whereas column 4 does. Columns 3 and 4 control for district, year, ethnicity, district x ethnicity, and year x ethnicity fixed effects. Robust standard errors are reported in the parenthesis. ∗p<0.1, ∗∗p<0.05, ∗∗∗p<0.01.

As a robustness check to this overall estimate, I explore alternative specifications in Appendix Table A.4. There, column 1 assigns treatment based on Li et al. (2019), with an alternative measure assigning 8 districts as high-conflict (with more than 95% of all conflict events occurring in these districts and a larger presence of LTTE) and the rest as low-conflict zone. Column 2 assigns treatment and control districts to top and bottom quartile districts (6 in each group) ranked by the mean level of conflict per capita per year. Columns 3 and 4 remove the top 5% and top 10% of conflict districts respectively. The result remains qualitatively similar.

As highlighted by Bongaarts' model (Bongaarts, 1987), one key determinant of fertility is age at marriage (operating via the marriage market effect). Early marriage or cohabitation tends to lead to longer reproductive periods and, therefore, higher fertility rates. In line with this hypothesis, prior studies have found that age at marriage has been an important determinant of fertility in Sri Lanka with a negative relationship for all ethnic groups (Abeykoon, 1987).

To understand if conflict resulted in an increased age at marriage that exerted a negative effect on overall fertility noted in Table 1.1, I estimate the following equation:

$$aam_{dt} = \beta_0 + \beta_1 treated_d \times post_t + \theta_d + \delta_t + \epsilon_{dt} \text{ ....(2)}$$

where *aam* is the female age at marriage in district d and year t. As before, *treated* is a dummy variable that takes a value of 1 for high-conflict districts and 0 for others, *post* is a dummy that takes a value of 1 for years after 1983 that mark the beginning of the civil war (i.e., for 1992-2006) and 0 for pre-war years (1953, 1971, and 1981). The coefficient β_1 measures the effect of conflict

Table 1.2: Impact of Conflict on Age at Marriage

Dependent Variable: Female Age at Marriage	
	(1)
Independent Variable	
treated x post	1.93***
	(0.376)
N	441
R^2	0.853

Note: Treatment is assigned based on classification in Appendix Table A.1. The regression in column 1 reports the coefficient for equation 2, giving the effect of war on female age at marriage in Sri Lanka. The regression includes district and year-fixed effects. Robust standard errors are reported in the parenthesis. Estimates are weighted by district population. The mean, minimum, maximum, and standard deviation for mean age at marriage is (24.062, 17.3, 30, 1.728). ∗p<0.1, ∗∗p<0.05, ∗∗∗p<0.01.

on age at marriage. θ_d and δ_t represent district and time fixed effects respectively. I estimate the model using robust standard errors (Castelló-Climent et al., 2018). [6]

The estimate in Table 1.2 suggests that on average, exposure to high-conflict led to an increase in female age at marriage by approximately 2 years in high-conflict zones in Sri Lanka. A higher age at marriage would delay fertility and thus result in a lower birth rate. This increased age at marriage could be driven by a marriage market squeeze if war impacted the local sex ratio and resulted in demographic deprivation of young men/husbands either due to men being increasingly recruited in combat or due to spousal deaths.

1.6 Digging Deeper: Heterogeneity & Replacement Effect

1.6.1 Heterogeneity by Ethnicity

Next, I evaluate the heterogeneity by ethnic groups. Appendix Figures A.2-A.6 present preliminary evidence of the differential in fertility trends using the size of the birth cohort and the crude birth rate.[7]

To formally estimate the heterogeneity in fertility adjustments to conflict, I estimate the following equation:

---

[6] As a robustness check, Appendix Table A.10 reports the estimates corresponding to Table 1.2 with clustered standard errors at the district level and estimated p-values from the wild cluster bootstrap recommended by Cameron et al. (2008) as a solution to the problem of a few clusters (Angrist & Pischke, 2009)

[7] Appendix Figure A.2 plots the ratio of the size of the birth cohort over time to the size of the birth cohort in 1967 for both ethnic groups. The graph suggests that there was no apparent trend in birth cohort sizes for either ethnic group in the pre-war years. However, during the war years, there is a marked increase in the size of the birth cohort of the Sri Lankan Tamils relative to 1967. Appendix Figures A.4 and A.5 plot the trend in the absolute cohort sizes over time for both ethnic groups as a robustness check. They yield similar trends to those in Appendix Figure A.2. The measure of relative cohort sizes in Appendix Figure A.2 is preferred as it normalizes the scale and allows analyzing trends for both ethnic groups using the same figure. Appendix Figure A.3 graphs the ratio of the birth rate over time to the birth rate in 1967 for both ethnic groups. The differential response to war is evident. The declining trend in the birth rate for Sri Lankan Tamils almost reverses during the war years. On the contrary, Sinhalese birth rates continue on a downward trajectory even during the war years. Appendix Figure A.6 plots the trend in the absolute birth rates over time for both ethnic groups as a robustness check and yields similar trends as in Appendix Figure A.3.

$$br_{gdt} = \beta_0 + \beta_1 treated_d \times post_t + \sum_{g=2}^{n} \beta_g \ ethnicity_g \times treated_d \times post_t + \gamma_g + \theta_d + \delta_t + \alpha_{gt} + \mu_{dg} + \epsilon_{gdt} \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \ (3)$$

where β_g captures the heterogeneity in the effect of conflict by ethnic group. Since my interests lie in looking at the two ethnic groups that were directly involved in the conflict, I simplify my analysis as follows: in the main specification, I classify ethnic groups as Sinhalese, Sri Lankan Tamils, and Others (non-Sri Lankan Tamil and non-Sinhalese category).[8] As a robustness check, in an alternate specification in Appendix Table A.6, I also disaggregate ethnic groups as Sinhalese, Indian Tamils and Sri Lankan Tamils, Moors and Others (aggregating any remaining ethnic groups, < 1% of the total population).[9] Regardless of specification, I included the dummy of n-1 ethnic groups interacted with *treated x post,* omitting it for Sinhalese and making Sinhalese the reference or base category. β_1 in this setup is the average effect of being in a high-conflict zone on Sinhalese fertility. $\beta_1 + \beta_g$ would give the net effect of conflict for community g. I estimate the model using robust standard errors (Castelló-Climent et al., 2018). [10] γ_g, θ_d *and* δ_t represent ethnic group, district, and time fixed effects respectively. I further control for district-ethnic group fixed effects, μ_{dg} and year-ethnic group fixed effects, α_{gt} to account for unobservables and prevent omitted variable bias such as accounting for some geographic differences within ethnic groups (for instance accounting for some behavioral differences between Sri Lankan Tamils in the northern part relative to those Sri Lankans who lived elsewhere) or for accounting for shocks/policy in some years across Sri Lanka that could have affected ethnic groups differently respectively (for instance the introduction of a nationwide family planning program that could differentially affect ethnic groups based on their cultural acceptance of contraceptives).

The regions marked by intense conflict were predominantly populated by Sri Lankan Tamils.[11] The high conflict districts were the districts that that represented the traditional Tamil

---

[8] This gives a total of 2400 observations (25 districts x 32 years x 3 ethnic groups = 2400).

[9] This gives a total of 4000 observations (25 districts x 32 years x 5 ethnic groups = 4000).

[10] As a robustness check, columns 3 and 4 of Appendix Table A.9 report the estimates corresponding to columns 3 and 4 of Table 1.1 with clustered standard errors at the district level and estimated p-values from the wild cluster bootstrap recommended by Cameron et al. (2008) as a solution to the problem of a few clusters (Angrist & Pischke, 2009).

[11] In the high-conflict districts, the Sinhalese ethnic group had an average population size of around 15438.8, while the Sri Lankan Tamils had a considerably larger average size of 151326.9 (before the war). In districts characterized by low conflict, the Sinhalese constituted the majority, with an average population size of 523724.5, in contrast to the Sri Lankan Tamils, whose average size was approximately 25416.78 (before the war).

homeland and were fought over to create the independent state of Tamil Eelam. Given that ethnicity played a crucial role in the evolution of the conflict, I hypothesize that the Sri Lankan Tamils and the Sinhalese would *not* have a similar demographic response since one of them is a majority and the other is a minority. The socio-political vulnerability of the group could shape its demographic adjustments to shocks. As per Table 1.1 column 3, the effect of high-conflict exposure leads to a decline in Sinhalese fertility by 8 live births per 1000 people. Column 4 which accounts for population weights suggests a similar effect.[12] For the effect on Sri Lankan Tamils, the *treated x post x Sri Lankan Tamil* coefficient is positive and significant. While the net effect on Sri Lankan Tamils as per column 3 is of reduction in fertility by 4 live births per 1000 people, column 4 suggests that the net effect on Sri Lankan Tamils is equivalent to a null effect (-10.66+ 11.41=0.75 ~ 0). These estimates in Table 1.1 underscore that considering only the overall effect on the population masks the heterogeneity by the ethnic groups. Sri Lankan Tamils experienced a significantly lesser reduction in their fertility in response to the conflict compared to the Sinhalese. For the Sri Lankan Tamils, their behavioral responses could be driven by not just an effort to maintain the size of their community and thus their political representation (De Silva et. al, 2020), but to also preserve their culture and tradition (Jordan and Denov, 2007). But to the contrary, a majority community would not necessarily be under such pressure and can thus reduce their fertility much more in response to the war.

As a robustness check for these heterogeneous effects by ethnicity, Appendix Table A.5 explores alternative specifications. There, to explore an alternative treatment classification, column 1 assigns treatment based on Li et al. (2019), assigning 8 districts as high-conflict (with more than 95% of all conflict events being in these districts and a larger presence of LTTE) and the rest as the low-conflict zone. Column 2 assigns treatment and control districts to top and bottom quartile districts (6 in each group) ranked by the mean level of conflict per capita per year. Lastly, columns 3 and 4 remove the top 5% and top 10% of conflict districts respectively to ensure that these are not driving the main results. Across these columns, the result remains qualitatively similar to the main estimates in Table 1.1 suggesting that there was a lesser fertility reduction for the Sri Lankan Tamils relative to the Sinhalese.

[12] Estimates with population weights are preferable as there is heterogeneity in the population of ethnic groups across districts and some districts have a lower-than-average minority share of population. Using population weights in this case ensures that small population sizes do not contribute disproportionately to the result.

Appendix Table A.6 describes additional results for a finer disaggregated effect across all 5 ethnic groups as a further robustness check. Columns 1- 4 are similar to those in Appendix Table A.5 described above. The estimates for Sinhalese and Sri Lankan Tamils are qualitatively and quantitatively similar to those in Table 1.1 and Appendix Table A.5, suggesting that aggregating remaining ethnic groups into the 'Other' category does not bias the main estimates.

For completeness, Appendix Table A.7 explores estimates with a continuous treatment intensity which allows variation in treatment intensity and timing across districts in column 1. A placebo check based on this continuous specification in column 2 confirms that changing the timing of treatment intensity to pre-war years produces a null effect of war on fertility. The estimates here are obtained by dropping observations during the war and as a placebo, imposing the continuous conflict intensity from the years 1992 to 2005 on the years pre-war (1967 to 1980). This null estimate using a placebo conflict intensity suggests that the main estimates are not caused by some spurious pre-conflict trends. However, there are several reasons why it is not favorable to consider the continuous treatment specification as the main specification.[13]

As an extension of the difference-in-differences model, I examine the dynamic effects of war on fertility trends using the event-study design. I estimate the following specification to look at ethnicity-specific effects[14]:

$$br_{gdt} = \beta_0 + \sum_{i=1967}^{1979} \beta_i treated_{d,i} + \sum_{i=1992}^{2009} \beta_i treated_{d,i} + \left\{\sum_{j=1967}^{1979} \beta_j treated_{d,j} + \sum_{j=1992}^{2009} \beta_j treated_{d,j}\right\} x\ sri\ lankan\ tamil_g + \left\{\sum_{k=1967}^{1979} \beta_k treated_{d,k} + \sum_{k=1992}^{2009} \beta_k treated_{d,k}\right\} x\ other_g + \theta_d + \delta_t + \gamma_g + \epsilon_{gdt} \quad \ldots\ldots\ldots\ldots\ldots (4)$$

[13] There are two reasons to prefer the DiD specification. The first one is that a continuous treatment effect DiD requires a much stronger assumption than parallel trends has undesirable properties in heterogeneous treatment effect estimation, and has parameters harder to interpret, confounded with selection bias (Callaway et al., 2021). Second, while having more variation in the conflict intensity may be valuable, the UCDP dataset restricts its domain to events that were reported in secondary sources like newspapers, reports, and other media outlets. So arguably, any media bias in reporting these events can lead to biased estimates. Collapsing the information into a binary variable would make estimates less sensitive as I am then identifying a broader intensity of conflict in the district. Therefore, the binary variable of conflict is the preferred specification for this analysis.

[14] Appendix Table A.8 reports the estimates for the overall effect on Sri Lanka in column 1. These estimates are obtained by estimating the following equation (4'): $br_{gdt} = \beta_0 + \sum_{i=1967}^{1979} \beta_i treated_{d,i} + \sum_{i=1992}^{2009} \beta_i treated_{d,i} + \theta_d + \delta_t + \epsilon_{gdt}$ for the whole population. The estimates are hovering around 0 pre-war and become negative for several years in the war period. Appendix Figure A.7 is the corresponding figure for these estimates.

Here, $treated_{d,i}$ is a dummy variable that takes a value of 1 high-conflict districts for year i and 0 otherwise. The years 1967- 1979 form the lag years and the year 1980 is the base category and is therefore dropped from the specification. The years 1992 to 2009 comprise the lead years of this analysis. Sinhalese are the omitted base category, as before. $sri\ lankan\ tamil_g$ is a binary variable that takes the value 1 for Sri Lankan Tamil ethnic group and 0 otherwise. $other_g$ takes a value of 1 for the other ethnic groups. $\gamma_g, \theta_d, and\ \delta_t$ represent the ethnicity, district and time fixed effects respectively.

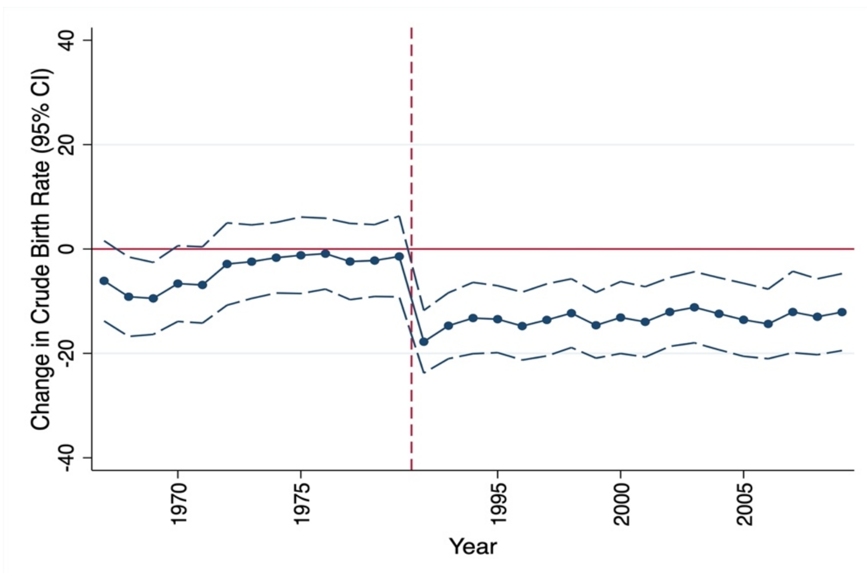

Figure 1.2: Dynamic Effect of Conflict on Crude Birth Rate of Sinhalese

Note: This figure plots the coefficient and 95% CI of the effect of conflict on the crude birth rate for the Sinhalese ethnic group.

The estimates from equation 4 are reported in Appendix Table A.8. Column 2 provides the coefficients for the Sinhalese ethnic group over the years (β_i). Column 3 provides the heterogeneity estimates from the interaction term in equation 4 for the Sri Lankan Tamils (β_j). This represents the differential between the fertility of the two ethnic groups. Adding column 2 and column 3 gives the total effect on Sri Lankan Tamils, which is listed in column 4 ($\beta_i + \beta_j$).

The estimates from columns 2 and 4 have been graphed in Figures 1.2 and 1.3 to show the dynamic effect for the two groups.

For the Sinhalese, Figure 1.2 depicts that the effect of conflict on the Sinhalese birth rate was negative. The coefficient remains negative through all war years, whereas pre-war years saw a fluctuation around 0. A close look at Figure 1.3 suggests that for the Sri Lankan Tamils, fertility did not drop much in the war years. Therefore, comparing Figures 1.2 and 1.3, it is evident that the reduction in fertility was of a larger magnitude for the Sinhalese. The estimates from column 3 of Appendix Table A.8 further indicate that the difference in fertility behavior between the two groups during war years is statistically significant. These results are in line with the heterogeneity evidence from the difference-in-difference estimates of the two groups presented in Table 1.1. Therefore, the dynamic effects are largely consistent with the difference-in-difference estimates.

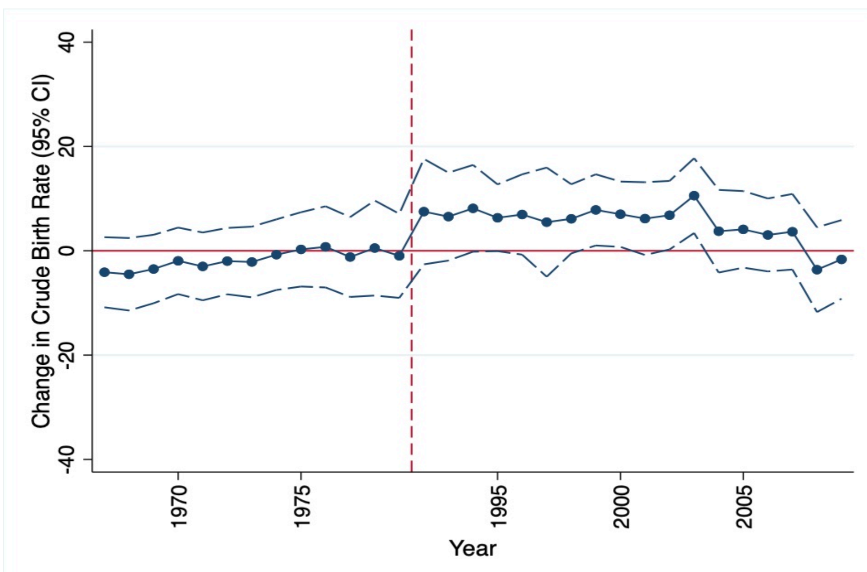

Figure 1.3: Dynamic Effect of Conflict on Crude Birth Rate of Sri Lankan Tamils

Note: This figure plots the coefficient and 95% CI of the effect of conflict on the crude birth rate for the Sri Lankan Tamil ethnic group.

Overall, estimates are qualitatively robust across the definition of treatment assignment. Estimates of the interaction term coefficient for Sri Lankan Tamils across the specifications are significant and positive which implies Sri Lankan Tamils did have a significantly different behavioral response to the war, and most specifications suggest a lesser reduction in fertility compared to Sinhalese. These results are in line with the overall findings of Puvanarajan and De Silva (2001). While their study does not estimate the impact of civil war on fertility, they broadly trace trends in fertility by ethnicity in Sri Lanka and find that the Sinhalese witnessed a stronger reduction in fertility than Sri Lankan Tamils during the 1990s. They also find evidence of an increasing age at marriage in Sri Lanka.

To the extent that I am unable to control for migration, these results should be interpreted with some caution. Several sources highlight that the civil war had resulted in large emigration from Sri Lanka (Fazil, 2019; Ibrahim et al., 2021; Sen, 2020). Sriskandarajah (2002) reports that about 25% of the entire Sri Lankan Tamil population left the country. It is possible that those who left were the ones who had access to wealth and higher education which increased their possibility of finding refuge elsewhere. One might argue if the more educated/ wealthy had higher fertility rates because they could afford to have kids, then their exit could artificially cause the fertility estimates of my analysis. For instance, Zimmerman and Carter (2003), Thiede et al. (2020) and Dehejia and Lleras-Muney (2004) find evidence that with access to greater resources, the educated and wealthy could be more able to buffer the effects of conflict, and thus afford to raise children in the presence of exogenous shocks like wars and famines. However, if 25% of the Sri Lankan Tamils left, because the number leaving is so large, there were likely people from across the entire income distribution who left and not just those on the top (which typically comprises 1% of the population). Further, there is also evidence in the literature that the more educated have a lower fertility rate (Schultz, 1994; Duflo et al., 2015; Jayachandran, 2017). So apriori, it is unclear as to what was the composition of the people who left during war and what it would mean in terms of bias in the estimates of this study.

1.6.2 Replacement Effect

I test for the presence of a replacement effect at the ethnic group level as a mechanism for the heterogeneity findings. Previous studies often utilize survey data to measure this effect at the individual level where they trace the woman's family history to see if she experienced any wartime deaths and altered her fertility in response to it. In this study, since I don't have access to individual-level survey data, I measure the response of fertility at the ethnic group level to lagged conflict-specific mortality rate at the district level.[15, 16]

I estimate the following equation:

$$br_{gdt} = \beta_0 + \beta_1 treated_d \times post_t + \beta_2 treated_d \times post_t \times dr_{d,t-1} + \beta'_g \sum_{g=2}^{n} ethnicity_g \times treated_d \times post_t + \beta''_g \sum_{g=2}^{n} ethnicity_g \times treated_d \times post_t \times dr_{d,t-1} + Pairwise\ Interaction\ terms + \gamma_g + \theta_d + \delta_t + \nu dr_{d,t-1} + \alpha_{gt} + \mu_{dg} + \epsilon_{gdt} \quad \text{……………. (5)}$$

where *br* is the crude birth rate for ethnic group g in district d at time t. Likewise, *dr* is the conflict-specific death rate for ethnic group g in district d at time t-1. *treated* is a dummy variable that takes a value of 1 for high-conflict districts and 0 for others, *post* is a dummy that takes a value of 1 for years after 1983 that mark the beginning of the civil war. As before, I include the dummy variables of n-1 ethnic groups interacted with *treated x post*, omitting it for Sinhalese and making Sinhalese the reference or base category. Thus, as in the main specification, I include the dummy for Sri Lankan Tamils and for others category (n=3). Pairwise interactions include the interaction of *treated* with ethnicity dummies, *post* with ethnicity dummies, *treated* with death rate and its lag, *post* with a lag of death rate, as well as ethnicity dummies with a lag of death rates. $\gamma_g, \theta_d\ and\ \delta_t$ represent ethnic group, district, and time fixed effects respectively. I further control for district-ethnic group fixed effects, μ_{dg} and year-ethnic group fixed effects, α_{gt}. [17]

There would be evidence of a replacement effect if an ethnic group's fertility response was positive in high-conflict zones in response to the lagged conflict-specific death rate. Table 1.3 presents these estimates. The sign of the coefficient *treated x post x death rate lag (i.e.* β_2*)* would reflect replacement effects or lack thereof for Sinhalese, which was the reference category ethnic group. Lagged death rates seemed to negatively impact fertility for Sinhalese in conflict zones

[15] The DHS of Sri Lanka data are not publicly available and could not be obtained for this empirical exercise. Further limitations of the DHS are that most rounds of the survey did not the cover high conflict districts in the North and East, making it difficult to study health in conflict zones.

[16] Since behavioral effects such as fertility require a minimum of 9 months to reflect in data, I consider lagged conflict-specific mortality rate by one year to measure replacement effects.

[17] As a robustness check, Appendix Table A.11 reports the estimates corresponding to Table 1.3 with clustered standard errors at the district level and estimated p-values from the wild cluster bootstrap recommended by Cameron et al. (2008) as a solution to the problem of a few clusters (Angrist & Pischke, 2009).

which implied that with an increase in conflict-specific mortality, Sinhalese in high-conflict districts reduced fertility and thus points to the absence of any replacement mechanism. This could be possible due to various reasons such as the majority status of the ethnic group which allows it to reduce fertility without any immediate threat to its group's political representation in the country. Other reasons for increased mortality resulting in lower fertility could include fear and uncertainty that manifest into psychological and physiological stress and trauma.

Table 1.3: Replacement Effect Mechanism – Impact of Conflict on Fertility

	Dependent Variable: Crude Birth Rate
	(1)
Independent Variable:	
treated x post	-10.336**
	(0.877)
treated x post x deathrate lag	-6.496*
	(3.692)
treated x post x Sri Lankan Tamil	8.001***
	(1.771)
treated x post x death rate lag x Sri Lankan Tamil	38.996***
	(11.279)
N	2250
R^2	0.703

Note: Treatment is assigned based on classification in Appendix Table A.1. The regression includes all pairwise interactions of post, treated, ethnicity, and lag of death rate. The regression also includes district, year, ethnicity, district x ethnic group, and year x ethnic group fixed effects. Robust standard errors are reported in the parenthesis. Estimates are population weighted. The mean, minimum, maximum, and standard deviation for lag of death rate is (0.187, 0, 20.818, 1.135). ∗p<0.1, ∗∗p<0.05, ∗∗∗p<0.01.

On the other hand, *treated x post x Sri Lankan Tamil x death rate lag*, which is the relevant coefficient for Sri Lankan Tamils, is positive *(i.e.* β'_2*)*. It seems that there is a high responsiveness of Sri Lankan Tamils to the conflict-specific mortality around them. The replacement effect coefficients suggest that due to 1 additional conflict-related death per 1000 people in the district, Sri Lankan Tamils responded by increasing their fertility by 32.5 births per 1000 people while the Sinhalese responded by reducing their fertility by 6.496 births per 1000 people. This could reflect some pro-natal sentiments in the minority group to try and maintain their group size during the war, an effort to replace the loss of a loved one or it may even reflect the insurance effect wherein parents increase fertility as a means of securing insurance for their old age if they fear only some of their kids would survive. Thus, there is evidence to suggest that there was an ethnic-group level replacement effect in place for the Sri Lankan Tamils, but not for the Sinhalese.

While this mechanism is not exhaustive, it is still indicative of a replacement effect partially contributing to the differential in fertility trends. Apart from this mechanism, other potential explanations include an increased risk of intimate partner violence and assault in conflict areas coupled with lower usage of contraceptives by Sri Lankan Tamils relative to Sinhalese. GoSL estimates suggest that 2% of all births in 1991-1998 were illegitimate and could potentially account for sexual violence channel during the war, however, that information is not available by ethnic groups at the district level for further analysis.[18] For the most part provision of healthcare was taken up by the LTTE in war zones (Stokke, 2006; Mampilly, 2011). If they were biased in the provision, access to health care in conflict zones might be selectively more available to Sri Lankan Tamils than to Sinhalese since social welfare provision was one aspect of LTTE state-building. This could have further contributed to a lesser negative effect on Sri Lankan Tamil fertility.

[18] The estimates for total births, based on vital statistics, include instances of births outside of marriage if they are officially registered. Consequently, these occurrences aren't distinctly singled out as specific mechanisms. Unfortunately, this statistic on births lacks an ethnic breakdown of illegitimate births, which limits the possibility of conducting an in-depth analysis regarding its implications for the two ethnic groups. Nonetheless, it's conceivable that illegitimate births could be linked to the avenue of sexual violence during the war. However, due to limitations in data availability, I am presently unable to account for this factor.

1.7 Conclusion

This paper provides evidence of a negative relationship between fertility and war in Sri Lanka on average. The increased female age at marriage in high-conflict zones contributes to this lower fertility. An exploration of the ethnic heterogeneity unveils that in response to high-conflict exposure, the Sinhalese had a larger reduction in fertility than their Sri Lankan Tamil counterparts. This study further contributes to the literature by finding evidence of the presence of a replacement effect for the Sri Lankan Tamils. This could reflect pronatalist sentiments in the minority group to maintain their group size during the civil war, an effort to replace the loss of a loved one or it may even reflect the insurance effect wherein parents increase fertility as a means of insurance for their old age if they fear only some of their kids would survive. Future research should explore these in more detail with survey micro-data which often ask questions on desired fertility, and record female fertility history along with information on the loss of loved ones. Future research may also study the marriage market channel disaggregated by the ethnic group if data becomes available.

From a policy perspective, there is enough evidence in the literature (Li et al., 2019) to suggest that children born during such a crisis can often suffer from lower birth weight and lower height-for-age (HAZ) or weight-for-age (WAZ) scores under age 5. If births are not well spaced apart, it can negatively impact the health of the mother and the child. Lower health outcomes could potentially transform into lower human capital in adulthood. So reproductive health policies and accessibility of neonatal care need to be targeted towards high-conflict areas.

It is also key for future research to analyze fertility trends post-war. Jayasundara-Smits (2018) highlights the sense of vulnerability in the minority post-war. She finds that war areas report a dominance of bodily grave crimes relative to the non-war areas where material grave crimes were dominant. With a distinct pattern in the nature of post-war crime in the two regions, the minority may continue to feel threatened. If the replacement effect in Sri Lankan Tamils was in part due to a fear of diminished old-age insurance due to war deaths or due to a pronatalist sentiment to maintain their identity in Sri Lanka, then the differential nature of the post-war crimes in conflict-affected areas could sustain this fertility differential. This has far-reaching consequences in the long term. It can change the demographic composition of the country and thus has the potential to impact the possibility of a future war in an already volatile political environment.

1.8 References

Abeykoon, A. T. (1987). Ethnic models of fertility behavior in Sri Lanka. *Asia-Pacific Population Journal*, 2(4), 29–42. PMID: 12269181.

Abu-Musa, A., Kobeissi, L., Hannoun, A., & Inhorn, M. (2008). Effect of war on fertility: A review of the literature. *Reproductive Biomedicine Online*, 17(Suppl 1), 43–53.

Agadjanian, V., & Prata, N. (2002). War, Peace, and Fertility in Angola. *Demography*, 39, 215–31. https:// doi.org/10.1353/dem.2002.0013.

Alburez-Gutierrez, D. (2019) *What Drives Fertility Recovery after Armed Conflict? A Case Study from the Guatemalan Genocide*, Giornate di Studio sulla Popolazione.

Angrist, J. D. and Pischke, J-S. (2009). *Mostly Harmless Econometrics*. Princeton University Press.

Barber, E. S. (2000). The White Wings of Eros': Courtship and Marriage in Confederate Richmond. In Catherine Clinton ed., Southern Families at War: Loyalty and Conflict in the Civil War South (pp. 120). Cary, United States: Oxford University Press. 2000.

Ben-Porath, Y. (1976). Fertility Response to Child Mortality: Micro Data from Israel. *Journal of Political Economy*, 84(4), S163–S178.

Blanc, A. K. (2004). The role of conflict in the rapid fertility decline in Eritrea and prospects for the future. *Studies in Family Planning*, 35(4), 236–45.

Boehnke, J., & Gay, V. (2020). The Missing Men: World War I and Female Labor Force Participation. *Journal of Human Resources*. https://doi.org/10.3368/jhr.57.4.0419-10151R1.

Bongaarts, J. (1987). The proximate determinants of fertility. *Technology in Society*, 9(3–4), 243–260.

Callaway, B., Goodman-Bacon, A., Sant'Anna, P. H. C. (2021). *Difference-in-Differences with a Continuous Treatment*. Papers 2107.02637, arXiv.org, revised Jul 2021.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-Based Improvements for Inference with Clustered Errors. *The Review of Economics and Statistics*, MIT Press, 90(3), 414–427.

Castelló-Climent, A., Chaudhary, L., Mukhopadhyay, A. (2018). Higher Education and Prosperity: From Catholic Missionaries to Luminosity in India. *The Economic Journal*, 128, 616.

Catani, C. (2018). Mental health of children living in war zones: a risk and protection perspective. *World Psychiatry*, 17, 104–105.

Census of Population and Housing Report. (1963, 1971, 1981, 2001, 2012). Department of Census & Statistics.

Chi, P. C., Bulage, P., Urdal, H., & Sundby, J. (2015). Perceptions of the Effects of Armed Conflict on Maternal and Reproductive Health Services and Outcomes in Burundi and Northern Uganda: A Qualitative Study. *BMC International Health and Human Rights*, 15(1), 7.

Collier, P., Elliot, V. L., Hegre, H., Hoeffler, A., Reynal-Querol, M., & Sambanis, N. (2003). *Breaking the Conflict Trap: Civil War and Development Policy.* World Bank Policy Research Report #26121.

Corno, L., Hildebrandt, N., & Voena, A. (2017). *Age of Marriage, Weather Shocks, and the Direction of Marriage Payments.* National Bureau of Economic Research Working Paper 23604.

Curlin, G. T., Chen, L. C., & Hussain, S. B. (1976). Demographic Crisis: The Impact of The Bangladesh Civil War (1971) on Births and Deaths in a Rural Area of Bangladesh. *Population Studies*, 30, 87–105.

Dehejia, R., & Lleras-Muney, A. (2004). Booms, Busts, and Babies' Health, *The Quarterly Journal of Economics*, 119(3), 1091–1130.

Devakumar, D., Birch, M., & Osrin, D., et al. (2014). The intergenerational effects of war on the health of children. *BMC Medicine*, 12, 57.

DeVotta, N. (2004). *Blowback: Linguistic Nationalism, Institutional Decay, and Ethnic Conflict in Sri Lanka*, Stanford University Press.

De Silva, W. I., De Silva, R., & Goonatilaka, W. S. M. (2020). Paradigm Shifts in Demographic Components in the New Millennium and Possible Implications: A Case Study of Sri Lanka. Sri Lanka, *Journal of Advanced Social Studies*, 10(2), 96–125.

Dossa, N. I., Zunzunegui, M. V., Hatem, M., & Fraser, W. D. (2015). Mental Health Disorders Among Women Victims of Conflict-Related Sexual Violence in the Democratic Republic of Congo. *Journal of Interpersonal Violence*, 30(13), 2199–2220.

Duflo, E., Pascaline, D., & Michael, K. (2015). Education, HIV, and Early Fertility: Experimental Evidence from Kenya. *American Economic Review*, 105(9), 2757–97.

Faust, D. G. (1996). *Mothers of Invention: Women of the slaveholding South in the American Civil War*. Univ of North Carolina Press.

Fazil, M. M. (2019). *Civil War-Induced Forced Migration and Post-Civil War Resettlement: Sri Lanka's Experience*. Kerala, India: Conference: Post-migrants, Refugees, Welfare Policies and the State, at University of Kerala.

Hacker, J. D., Hilde, L., & Jones, J. H. (2010). The Effect of the Civil War on Southern Marriage Patterns. *The Journal of Southern History*, 76(1), 39–70.

Ibrahim, Y., & Rajapillai, V. & Scott, S. (2021). Consuming conflict as Tamil consciousness: the case of second-generation British Sri Lankan Tamils. *Journal of Ethnic and Migration Studies*, 1–19. https:// doi.org/10.1080/1369183X.2021.1940890.

Jayachandran, S. (2017). Fertility Decline and Missing Women. *American Economic Journal: Applied Economics*, 9(1), 118–39.

Jayachandran, S., & Lleras-Muney, A. (2009). Life Expectancy and Human Capital Investments: Evidence from Maternal Mortality Declines. *The Quarterly Journal of Economics*, 124(1), 349–397.

Jayasundara-Smits, S. (2018). Lost in transition: linking war, war economy and post-war crime in Sri Lanka. *Third World Thematics: A TWQ Journal*, 3(1), 63–79.

Jordan, K., & Denov, M. (2007). Birds of Freedom? Perspectives on Female Emancipation and Sri Lanka's Liberation Tigers of Tamil Eelam. *Journal of International Women's Studies*, 9(1), 42–62.

Keethaponcalan, S. I. (2002). Social Cubism: A Comprehensive Look at The Causes Of conflict In Sri Lanka. *ILSA Journal of International & Comparative Law*, 8(3), 9.

Kottegoda, S., Samuel, K., & Emmanuel, S. (2008). Reproductive Health Concerns in Six Conflict- Affected Areas of Sri Lanka. *Reproductive health matters*, 16, 75–82. https://doi.org/10.1016/ S09688080(08)31359-7.

Kraehnert, K., Brück, T., Di Maio, M., & Nisticò, R. (2018). The Effects of Conflict on Fertility: Evidence from the Genocide in Rwanda. *Demography*. https://doi.org/10.1007/s13524-019-00780-8.

Kulczycki, A., & Prem, C. S. (1999). New evidence on fertility transition through wartime in Lebanon. *Genus*, 55(3/4), 131–52. Accessed March 15, 2021.

Ladier-Fouladi, M., & Hourcade, B. (1997). The paradoxes of demographic trends in Iran: fertility in rapid decline. In Jean-Claude Chasteland & Jean-Claude Chesnais (eds*.), La Population du Monde: enjeux et problemes* (pp. 209–224). Paris: Presses Universitaires de France.

Leatherman, J. L. (2011). *Sexual Violence and Armed Conflict* (pp. 176. Cambridge and Oxford: Polity Press. paperback. ISBN-10: 0745641881.

Levine, R. (2007). *Case Studies in Global Health: Millions Saved*. Sudbury, MA: Jones and Bartlett.

Li, J., Yamazaki, K., & Ito, T. (2019). *Probing Indirect Effects of Civil Conflict on Child Health in Non- Conflict Zones: Evidence from Sri Lanka*. GSICS Working Paper Series 35, Graduate School of International Cooperation Studies, Kobe University.

Lindskog, E. (2016). *Effects of Violent Conflict on Women and Children: Sexual Behavior, Fertility, and Infant Mortality in Rwanda and the Democratic Republic of Congo*. PhD Thesis, Stockholm University, Department of Sociology.

Lindstrom, D. P., & Berhanu, B. (1999). The Impact of War, Famine, and Economic Decline on Marital Fertility in Ethiopia. *Demography*, 36, 247–61. https://doi.org/10.2307/2648112.

Madsen, E. L., & Finlay, J. E. (2019). *The Long-Lasting Toll of Conflict on Fertility and Early Childbearing*. Policy Brief, Population Reference Bureau. Retrieved 20 December, 2023, from https://www.prb.org/wp-content/uploads/2019/09/policy-brief-toll-of-conflict-on-childbearing.pdf.

Mampilly, Z. C. (2011). *Rebel rulers: insurgent governance and civilian life during war.* Ithaca, NY: Cornell University Press.

McGinn, T. (2000). Reproductive Health of War-Affected Populations: What Do We Know? *International Family Planning Perspectives*, 26(4), 174–180.

McKay, S. (1998). The Effects of Armed Conflict on Girls and Women. *Peace and Conflict*, 4(4), 381–392.

Nowrojee, B. (1996). Shattered Lives: Sexual Violence during the Rwandan Genocide and its Aftermath. New York: Human Rights Watch.

Nugent, J. (1985). The Old-Age Security Motive for Fertility. *Population and Development Review,* 11(1), 75–97.

Østby, G. (2016). *Violence Begets Violence: Armed Conflict and Domestic Sexual Violence in Sub-Saharan Africa*. HiCN Working Papers 233, Households in Conflict Network.

Palmer, J. J., & Storeng, K. T. (2016). Building the Nation's Body: The Contested Role of Abortion and Family Planning in Post-War South Sudan. *Social Science & Medicine*, 168, 84–92.

Phadera, L. (2021). Unfortunate Moms and Unfortunate Children: Impact of the Nepali Civil War on Women's Stature and Intergenerational Health. *Journal of Health Economics*, 76, 102410.

Puvanarajan, P., & De Silva, I. (2001). Fertility Decline in Sri Lanka: Are All Ethnic Groups Party To The Process? *Asia-Pacific Population Journal*, 16, 23–46. https://doi.org/10.18356/864f2b6c-en.

Rossi, P., & Godard, M. (2019). The Old-Age Security Motive for Fertility: Evidence from the Extension of Social Pensions in Namibia. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3457324.

Schindler, K., & Bruck, T. (2011). *The Effects of Conflict on Fertility in Rwanda*. Policy Research WPS 5715. World Bank.

Schultz, T. P. (1994). Human capital, family planning, and their effects on population growth. *The American Economic Review*, 84(2), 255–260.

Schultz, T. P. (1997). Chapter 8 Demand for children in low-income countries. *Handbook of Population and Family Economics, Elsevier*, 1(Part A), 349–430.

Sen, R. (2020). *Diaspora Engagement Mapping: Sri Lanka.* Retrieved 20 December 2023 from https://diasporafordevelopment.eu/wp-content/uploads/2020/03/CF_Sri-Lanka-v.20.pdf.

Sriskandarajah, D. (2002). The migration-development nexus: Sri Lanka case study. *International Migration*, 40(5), 283–307.

Stokke, K. (2006). Building the Tamil Eelam State: Emerging State Institutions and Forms of Governance in LTTE-Controlled Areas in Sri Lanka. *Third World Quarterly*, 27(6), 1021–1040.

Sundberg, R., & Melander, E. (2013). Introducing the UCDP Georeferenced Event Dataset. *Journal of Peace Research*, 50(4), 523–532. https://doi.org/10.1177/0022343313484347.

Thiede, B. C., Hancock, M., Kodouda, A., & Piazza, J. (2020). Exposure to Armed Conflict and Fertility in Sub-Saharan Africa. *Demography*, 57(6), 2113–2141.

United Nations Statistics Division. (1970–2010). Demographic Yearbook. Retrieved 8 April, 2023, from https://unstats.un.org/unsd/demographic-social/products/dyb/#statistics.

United Nations Statistics Division. (2006). Population and Vital Statics Report. Retrieved 8 April, 2023, from https://unstats.un.org/unsd/demographic-social/products/vitstats/index.cshtml#previous.

United Nations. (2014). *Sexual Violence: a Tool of War*. Retrieved 20 April, 2022, from https://www.un.org/en/preventgenocide/rwanda/assets/pdf/Backgrounder%20Sexual%20Violence%202014.pdf.

Urdal, H., & Che, C. P. (2013). War and Gender Inequalities in Health: The Impact of Armed Conflict on Fertility and Maternal Mortality. *International Interactions*, 39(4), 489–510.

Vandenbroucke, G. (2014). Fertility and Wars: The Case of World War I in France. *American Economic Journal: Macroeconomics*, 6(2), 108–36.

Vital Statistics. (1967–1980). Sri Lanka: Marriages, births, and deaths. Department of Census & Statistics, Sri Lanka.

Vital Statistics. (1992–2009). Department of Census & Statistics, Sri Lanka. Retrieved 13 July, 2021, from http://www.statistics.gov.lk/Population/StaticalInformation/VitalStatistics

Wood, E. J. (2018). Rape as a Practice of War: Toward a Typology of Political Violence. *Politics & Society*, 46(4), 513–537. https://doi.org/10.1177/0032329218773710.

Zimmerman, F. J., & Carter, M. R. (2003). Asset Smoothing, Consumption Smoothing and the Reproduction of Inequality Under Risk and Subsistence Constraints. *Journal of Development Economics*, 71, 233–260.

Chapter 2: Impact of Early Childhood Health Shocks on Intergenerational Mobility: Evidence from the 1848 Public Health Act[1]

2.1 Introduction

Does early childhood health affect intergenerational mobility? While health is desirable in itself to increase quality and length of life, it also enables the development of social and economic opportunities as adults (Currie, 2009; Almond & Currie, 2011a, b; Gutiérrez et al., 2019). Despite its significance, the role of early life health in shaping the generational persistence of economic opportunities (or lack thereof) remains understudied. I address this gap in the literature using historical data and determine whether exposure to improved sanitation and hygiene under the Public Health Act of 1848 in early childhood facilitated intergenerational mobility in England and Wales in the nineteenth century. Previous literature documents the large burden of diseases arising due to a lack of safe drinking water, sanitation, and hygiene (WASH).[2] Public intervention in health in the form of safe drinking water, sanitation, and hygiene can lead to intergenerational mobility

---

[1] I am grateful to Priyaranjan Jha, Vellore Arthi as well as Meera Mahadevan for their invaluable guidance and encouragement. I am also thankful to Vellore Arthi and Myra Mohnen for generously sharing datasets of vital statistics and linked historical census data respectively. All errors are my own.

[2] Lack of WASH can lead to an increased risk of diarrhea, chronic undernutrition, and stunting due to bouts of diarrhea, acute respiratory infections as well as soil-transmitted helminthiasis. It can also cause malaria, trachoma, schistosomiasis, and lymphatic filariasis. See Esrey et al. (1991), Prüss-Ustün et al. (2019), Wolf et al. (2023).

since healthier children become adults with better physical and cognitive abilities and also because healthier children can spend more effective time improving their education and skills.

The Public Health Act of 1848 made the state the guarantor of standards of health and environmental quality for the first time in the history of England and Wales. It is referred to as the 'sanitation revolution' of the Victorian era and is considered one of the greatest milestones in public health history (Fee & Brown, 2005). The act marked the first attempt of England and Wales at a systematic sanitation improvement in the context of a grim situation of public health in the early nineteenth century. There was inadequate infrastructure to meet the drainage and water supply requirements of the growing population. The death rate in the 1840s was thrice of what it is today in England. Diarrhea, cholera, dysentery, and typhoid were the leading causes of death.[3] A majority of these deaths were concentrated among children due to fevers, respiratory illness, or diseases controllable by proper sanitation arrangements. The passing of the Public Health Act in 1848 legislated the establishment of local health boards in districts with an average mortality rate of 23 deaths per 1000 people or more over the previous 7 years. These local boards could act to alter local infrastructure such as sewers and waterworks. Among other duties, they were also responsible for appointing an inspector of nuisances, sweeping streets, closing cesspits, and ensuring the provision of public privies.

As a starting point for this study in line with related literature evaluating mortality decline in the nineteenth century, I assess the impact of exposure to WASH facilities under the Public Health Act on mortality in England and Wales. I classify districts into treatment, and control districts based on the 23 deaths per 1000 people threshold. To account for spatial spillovers, the main specification also identifies spillover districts as districts that are contiguous to treatment districts but themselves not eligible for treatment. With newly collected vital statistics data on mortality from 1841 to 1880, I adopt a difference-in-differences framework and find that the Public Health Act of 1848 led to a mortality reduction of 1 death per 1,000 people. However, since mortality is a crude measure of health improvements it may not necessarily capture meaningful changes in population health. For instance, the provision of better hygiene may strongly affect intrinsic health outcomes such as morbidity and quality of life but a crude measure such as mortality may not fully gauge that. Recent literature

[3] Mortality rates from cholera, diarrhea, diphtheria, dysentery, respiratory diseases, tuberculosis, typhus, and whooping cough in England were 8.5 deaths per 1,000 persons. This were nearly as high as the overall mortality rate today of 9.3 deaths per 1,000 persons. (Kiniria, 2019).

utilizes information on anthropometric outcomes to capture the impact on intrinsic health and well-being (Bhalotra & Rawlings, 2013; Habicht et al., 1974; Gutiérrez et al., 2019). However, a lack of this data in historical settings prevents studies from being able to directly capture changes in such valuable measures. Investigating intergenerational mobility therefore becomes important because it allows this study to gauge the impact using more comprehensive measures - capturing the effects that operate along the intrinsic margin and impact the human capital formation of survivors.

To understand how this early public health intervention impacted intergenerational mobility, I utilize a novel dataset from Costas-Fernández et al. (2020) which contains a linked sample of more than half a million father-son pairs. The sample was created from the recently digitized 1851 and 1861 full population censuses of England and Wales (Schürer & Higgs, 2020). The census provides information such as the age, district of residence, number of servants in the house, and occupation. However, as is often the case with historical datasets, they do not have measures of income which makes social mobility hard to measure. With a lack of information on other economic outcomes, following Long and Ferrie (2013), I adopt a measure of occupational mobility to study intergenerational mobility in England and Wales during the nineteenth century. I construct a measure of intergenerational mobility by comparing the father's occupation in 1851 to his son's occupation in 1861. Employing a difference-in-differences methodology, the identification strategy exploits spatial variation in treatment at the district level as well as the age of the son in 1848 when the Public Health Act came in place. The age of the son plays a pivotal role in determining his susceptibility to health shocks. An increasing body of evidence suggests that there is a strong link between well-being in the first 1000 days of life and later life outcomes. This early period, spanning from gestation to the second birthday of the child is a crucial period for brain development which is the foundation of the child's eventual cognitive, emotional, and social development. Therefore, I characterize this period as early childhood. I hypothesize that for cohorts receiving the public health intervention in early childhood, access to WASH facilities through the Public Health Act will (net of culling) raise the returns to schooling, thus raising both attendance in childhood and the cognitive skill content of occupations in adulthood. In contrast, for cohorts exposed to the Act later in childhood, the intervention would act as a positive labor productivity shock that, by increasing the child's health capital in the short run, raises the opportunity cost of schooling and so speeds the transition into child labor (Costa, 2015).

In line with this hypothesis, the paper finds evidence that sons who were exposed to the Act in early childhood in treatment districts were 5% more likely to be in a different occupation than their father and

16% more likely to be upward mobile compared to their counterparts in control districts. These findings are robust across a range of subsamples and different specifications. I also conduct a placebo check that randomizes treatment 1000 times to obtain a distribution of estimates. The iterations result in a distribution that is centered at zero, validating that the main estimates are not generated by chance. Further analysis reveals that this upward mobility is largely driven by the transition of sons out of farming and unskilled occupations into skilled/semiskilled occupations. In the absence of data on direct measures of literacy such as years of education, I utilize a historical description of literacy requirements in nineteenth-century jobs based on Mitch (1992) to categorize occupational groups into the ones that are likely to require/use literacy and ones that did not. I find evidence that sons with early childhood exposure to the act were more likely to be in jobs that required literacy, implicitly suggesting that they gained more human capital. A positive shock to health in early childhood due to the WASH measures also increased the likelihood of migration for sons. Analyzing estimates for the upward mobility of movers and stayers suggests that geographic mobility is a potential mechanism for these findings.

This paper contributes to several strands of literature. First, this paper contributes to the literature on intergenerational mobility. With a growing perception that parents' fortune can propagate advantages for children, and thus sustain inequality, intergenerational occupational mobility has been a rising subject of study for economists. The Stand and Rising (2011) OECD report suggests that increasing income inequality "can stifle upward social mobility, making it harder for talented and hard-working people to get the rewards they deserve." With evidence in the literature that intergenerational mobility has varied over time and across countries,[4] studies have analyzed several factors that influence these differences. Andrade et al. (2003) find debt to be an important determinant of intergenerational mobility suggesting that access to capital can be a crucial determinant. Loans allow investment by parents into human capital development such as school enrollment for the child and allow a transfer of higher skills and pushes the child on a higher occupational trajectory. Shocks such as wars (overview of literature provided by Orero et al., 2007) and natural disasters (Caruso, 2017) are also likely to impact social mobility. Costas-Fernández et al. (2020) study the impact of the provision of transportation infrastructure in England

[4] Stand and Rising (2011) report finds that intergenerational earnings mobility is low in countries with high inequality such as Italy, the United Kingdom (Atkinson et al., 1983; Dearden et al., 1997; Clark & Cummins, 2015), and the United States (Solon, 1999; Solon, 2002), and much higher in the Nordic countries (Osterbacka, 2001), where income is distributed more evenly.

and Wales in the nineteenth century.[5] Other public interventions that affect intergenerational mobility are changes in tax schemes (Piketty, 2000; Jones et al., 2020), neighborhood influence (Chetty et al., 2014; Chetty et al., 2016; Guerra & Mohnen, 2022), and caste-based reservation (Lodh et al., 2021).

While some recent studies have investigated the relationship between health and intergenerational mobility (Sacker, 2013; Bhalotra & Venkataramani, 2015; Gutiérrez et al., 2019; Karbownik & Wray, 2021), estimating the effects of public health interventions on intergenerational mobility remains an under-researched area. Specifically in the context of WASH campaigns, prior work has focused on the effect on contemporaneous outcomes such as child mortality, anthropometric outcomes, and nutritional well-being of children (Gamper-Rabindran et al., 2010; Spears & Lamba, 2016; Augsburg & Rodríguez-Lesmes, 2018; Alsan & Goldin, 2018). This paper seeks to fill this gap in the literature by providing evidence on the long-term effects of exposure to WASH measures through its impact on intergenerational mobility, with an emphasis on the historical setting of the Public Health Act of 1848. This act was an important first step toward better living conditions in nineteenth-century England and Wales and findings suggest that it played a significant role in shaping economic opportunities across generations.

Second, this paper also contributes to the literature that underscores the lasting effects of early life shocks on later life outcomes (Crimmins & Finch, 2006; Currie, 2009; Eppig et al., 2010; Almond & Currie, 2011a, b; Conti & Heckman, 2013). Several studies have specifically evaluated the role of policy changes during childhood such as expansions in childhood health insurance coverage (Goodman-Bacon, 2017; Brown et al., 2020) and improvements in living conditions (Gould et al., 2011), access to infant health care centers (Bütikofer et al., 2019) and introduction of sulfa drugs (Jayachandran et al., 2010). However, less is known about the contribution of age at exposure (during childhood) to these shocks on later life outcomes. As emphasized by Karbownik and Wray (2021), this distinction could be important to the extent that shocks experienced early in life could be more consequential than those that occur later on, for example, at age 10 (Heckman, 2006; Barham et al., 2013). This study aims to fill this void in literature by demonstrating that the first 1000 days of life are indeed a more vulnerable period and shocks to the health environment during this development phase have lasting implications.

[5] Their findings suggest that children who grew up closer to the railroad network were 6 percentage points more likely to work in a different occupation than their father and were also 5 percentage points more likely to be upward mobile.

Finally, it contributes to the literature evaluating mortality decline in the nineteenth and twentieth centuries (Cutler et al., 2006). There are broadly two strands of thought. One emphasizes the improvement in nutrition and diet and increased resistance to virulent diseases (McKeown & Brown, 1955; McKeown & Record, 1962; McKeown, 1976; Fogel, 1986) while the other more recent ones emphasize sanitation reforms and improved water supply (Szreter, 1988; Razzell, 1993; Mercer, 2014; Mooney, 2015). Studies across the United States and Europe find evidence of a reduction in mortality due to improvements in water supply and sanitation.[6] In line with the findings of Kiniria (2019), the preliminary results of this study confirm that the Public Health Act of 1848 led to a reduction in mortality in England and Wales.

The paper proceeds as follows: Section 2.2 describes the historical background of the Act. Section 2.3 describes the dataset. Section 2.4 presents the empirical strategy and results. Section 2.5 documents the robustness of the main estimates. Section 2.6 discusses factors that contributed to observed upward intergenerational mobility and Section 2.7 concludes.

2.2 Historical Background of the 1848 Public Health Act

There was rapid growth in the population of England and Wales in the early nineteenth century. In the fifty years leading up to the Act, England's population had increased twofold, rising from 7.75 million to 15.25 million (Kiniria, 2019). The Industrial Revolution resulted in rapid urbanization in towns and cities. The Report from the Select Committee on the Health of Towns (1840) finds that five of the most important provincial towns i.e., Manchester, Glasgow, Birmingham, Leeds, and Liverpool grew in the number of inhabitants twice as fast as the rate of population growth in England and Wales, leading to excess demand for housing and shelter. Simultaneously, there was a disregard for ensuring a reliable supply of clean water, maintaining clean streets, or implementing effective waste disposal systems. The builders were inadequate as they did not supply sewers, water closets, and privies and neglected the supply of fresh water, clean streets, or removal of garbage (Fee & Brown, 2005). The result was overcrowded cities with insufficient sanitation facilities, limited access to clean water, and rampant disease outbreaks.

[6] See Costa (2015), Meeker (1972), Preston & Van de Walle (1978), Meckel (1985), Millward & Bell (1998), Ferrie & Troeskon (2005, 2008), Cutler & Miller (2005), Chapman (2019).

These unlivable conditions came to attention in the 1830s as waves of the cholera epidemic and disease-led mortality crippled England and Wales. The poor state of public health was a pressing concern that required government intervention. In 1839, Edwin Chadwick, a lawyer and social reformer was appointed to investigate the state of sanitation and make recommendations. His report called the "Report on the Sanitary Conditions of the Laboring Population of Great Britain", published in 1842 strongly emphasized the direct relationship between the condition of the physical environment and disease prevalence. He wrote "The great preventives: drainage, street and house cleansing by means of supplies of water and improved sewerage, and especially the introduction of cheaper and more efficient modes of removing all noxious refuse from the towns, are operations for which aid must be sought from the science of the civil engineer, not from the physician, who has done his work when he has pointed out the disease that results from the neglect of proper administrative measures and has alleviated the sufferings of the victims". Chadwick suggested that the money paid out as poor relief to families of men who died due to infectious diseases was more than the cost of improving the sanitation infrastructure for the people. So based on his analysis, in the long run, it would be more economical to develop an improved hygienic environment for the people. Chadwick's report led to the appointment of a Royal Commission for Inquiry into the State of Large Towns and Populous Districts in 1843. He argued strongly in favor of the implementation of necessary sanitary measures (including clean water and ventilation) in his reports.

In August 1848, the parliament for the first time passed a law on public health matters in England and Wales. The Act was considered the beginning of the "commitment to proactive rather than reactive public health" (Hamlin & Sidley, 1998). It established the General Board of Health as the central authority to ensure the execution of the Act, comprising a President and two other board members. It was established for five years after passing the Act and continued annually thereafter until 1858 when the board ceased to exist. The Act legislated that the supply of water, sewerage and drainage, cleansing, and paving be under the supervision of a local body. Thus, a Local Board of Health was to be compulsorily formed if a place exceeded 23 deaths per 1,000 people over seven years. Approximately 200 local boards were established by 1853 with effective access to a population of over 2 million as per the 1851 census. [7]

[7] See Report of General Board of Health on the Administration of the Public Health Act and the Nuisances Removal and Prevention Acts from 1848 to 1854.

Each local board was authorized to designate a qualified medical practitioner or member of the medical profession as the officer of health and appoint a clerk, a surveyor, and a treasurer. They were mandated to hold board meetings at least once per month. The board also appointed an Inspector of Nuisances tasked with investigating complaints and acting against unsanitary conditions such as those in the case of industrial emissions and accumulations of refuse and sewage. Inspectors and Officers of Health conducted house-to-house visitations to inspect sanitation conditions undertook extensive cleanings, oversaw the removal of nuisances, and suggested improved means of ventilation.[8] Under the act, local boards assumed ownership of all public sewers and had the option to purchase private ones. They had the authority to construct, repair, clean, or empty sewers as needed. To enhance sanitation, approximately 8000 miles of drainpipes for sewers were manufactured by 1854, where around 2 miles of sewers served 1000 people in towns.[9] Board inspectors examined plans to combine drainage with existing waterworks and extension of waterworks. Boards also devised schemes to have pure, soft, perennial spring water instead of polluted hard, river water for their supplies by constant instead of intermittent service combined with works for carrying away wastewater. They facilitated the installation of service pipes in houses, installed sinks, filled up cesspools, substituted water closets and self-cleansing house drains, and constructed dust bins.

In addition to these responsibilities, the local boards were in charge of cleaning the streets within their district and removing ashes, dust, dung, and filth. They undertook other tasks such as paving and widening streets, regulating burial grounds, and slaughterhouses, and the provision of public lavatories. Loans could be made to the local boards to allow for expenditure needed to implement facets of the Act. For instance, records suggest that over a million pounds in loans had been sanctioned for sanitary improvements to the boards by the end first decade of the act.[10] The Board also reserved the right to penalize, in an amount not exceeding £5, anyone who obstructed the execution of the act or defaced the board. The effectiveness of the new works was appreciated by the residents and only 1 in 587 cases would necessitate the utilization of compulsory measures.

---

[8] The legal document of the Public Health Act of 1848 (CAP. LXIII. An Act for promoting the Public Health, 31st August 1848) and Lumley (1859) provide detailed descriptions of the roles of Inspectors and Officers of Health.
[9] See Report of General Board of Health on the Administration of the Public Health Act and the Nuisances removal and Prevention Acts from 1848 to 1854.
[10] See Return of Number of Applications for Sanction of General Board of Health to Permanent Works, and Amount of Loans sanctioned, 1854-56.

With the abolition of cesspools in houses and substitution of water closets, together with the introduction of tubular self-cleaning house drains, improved sanitation led to a reduction in mortality in several cases and people reported other intrinsic health improvements such as reduced headaches, diminished symptoms of dyspepsia and a decrease in overall sickness.[11]

2.3 Data

2.3.1 Vital Statistics

I obtain and digitize the annual birth and mortality statistics at the district level for the years 1841-1850 from the reports on Abstract of Birth and Abstract of Death. Information on births and deaths for the years 1851-1880 has been obtained from Arthi et al. (2022). District level populations for 1841, 1851, 1861, and 1881 are obtained from the Decennial Census Reports. I employ the DasGupta interpolation technique to obtain intercensal population estimates.[12] It estimates annual changes in population due to observed births and deaths within each district and predicts the population in the next census year. To account for any residual population change obtained at the decade level, (called the error of closure, a measure of implied net migration) the residual is then distributed smoothly across all intercensal years. Obtaining the intercensal population estimates allows the calculation of the crude death rate annually at the district level to analyze trends in mortality.

2.3.2 Intergenerational Mobility

I obtain the linked father-son database for the years 1851-1861 from Costas-Fernández et al. (2020). They utilize the full-population censuses of England and Wales in 1851 and 1861

[11] Report of General Board of Health on the Administration of the Public Health Act and the Nuisances Removal and Prevention Acts from 1848 to 1854 describes among a case of 500 persons, the mortality reduced to as low as 7 in 1000, when the average rate of mortality for the whole metropolis was close to 23 in 1000. In another favorable instance, the average mortality had reduced to 14 in 1000 based on the experience of 600-700 persons of the working class in the metropolis.

[12] This methodology has been historically used by the United States Census Bureau.

developed by the I-CeM project (Schürer & Higgs, 2020) and link individuals across censuses. [13] The census contains information on geographic variables like house number, street name, parish, registration district, and a range of sociodemographic information: age, gender, marital status, number of children, number of servants, and family structure. The only economic outcome available in this data is self-reported occupation. There are over 400 occupations such as physician, cook, stable keeper, cabinet maker, or farmer.

Within the linked sample of father-son pairs, I analyze sons who were 2 to 17 years old in 1851. This set of children would have been in utero to 14 years old in 1848 when the Public Health Act was first implemented. This gives me 608,624 pairs of father and son linked from the 1851-1861 censuses. Such a link facilitates tracking the son's occupation at a later time in 1861 when he is relatively older (12 to 27 years old) while observing the father's occupation in 1851 when he is younger. Historical settings largely use occupation to compute measures of intergenerational mobility (Boberg-Fazlic & Sharp, 2013; Clark & Cummins, 2015; Long & Ferrie, 2013; Olivetti & Paserman, 2015) due to two main reasons. First, historical data often do not have information on income making occupation the only economic outcome consistently available. Second, occupations withstand transitory shocks that incomes might not. Moreover, as emphasized by Costas-Fernández et al. (2020), occupation and social class encompass several dimensions of an individual's life experience that can be linked to interpretations of social mobility. These facets include aspects like community prestige, workplace autonomy, place of employment, and more.

Based on the Historical Social Class Scheme (HISCLASS) created to make comparisons across different periods, countries, and languages by Van Leeuwen and Maas (2011), individuals' occupations are classified and ranked into 12 groups. This skill-based classification allows us to rank the groups from unskilled farm workers (rank 12) to higher managers (rank 1). It also considers the extent to which some supervise the work of others and the economic sector. Table B.1 in the Appendix describes how these 12 categories have been collapsed into 4 broad categories: white collar, skilled/semiskilled, unskilled, and farmers (Long & Ferrie, 2013) for ease of analysis. Each class can be considered as a group of people with the same opportunities in life and the scheme

[13] Costas-Fernández et al. (2020) rely on birth year, parish of birth, given name, and surname (time-invariant characteristics) to link individuals across censuses.

conforms to the way economic and social historians have looked at social stratification from a historical perspective (van Leeuwen, 2020).

2.3.3 Geolocating Individuals

I geocode each individual in the data at the district level. I account for district boundary changes over time. There were largely two kinds of boundary changes. The first scenario involves splitting up a district over time into several districts. In such a situation, I aggregate observations of the smaller districts post the split to obtain a comparable geographical area. For instance, a district called Kensington and Chelsea in 1841 was dissolved and split into two smaller districts called Kensington and Chelsea respectively in 1847. Thus, for consistency, I agglomerate the smaller districts from 1847 onwards for comparability with the parent district before the split. In the second scenario, several districts merged fully to form a larger district or lost some parts to another district. In such a case, I agglomerate the districts to obtain a larger comparable region in the years before the boundary change. For instance, districts East London and West London in 1861 lost parishes to London City. So, to account for this, I agglomerate the three districts for years before 1861. Repeating this process for all districts over the years gives me 263 districts with consistent boundaries across England and Wales for my analysis. I thus capture the geographic mobility of the individual from 1851 to 1861 at the district level.

2.4 Empirical Analysis

The Act classified areas with a seven-year average mortality of 23 deaths per 1000 people or above to establish a Local Board of Health which would conduct several measures to improve sanitation and hygiene, enforced by the General Board of Health. I estimate the past seven-year average mortality of all districts in 1848 and identify districts that are eligible for treatment outlined by the Act. While this does not provide the actual location of the Local Boards of Health but is a proxy for where they should have been.[14] I find that 61 districts out of 263 should have been treated in

[14] In some cases, local boards could also be created if one-tenth of the taxpayers in the district petitioned to do so. Due to a lack of access to the exact location of the local boards, I am unable to account for such districts in my analysis. However, it can be inferred that a district that would petition for a local health board would be one where the mortality

1848 based on the parliamentary established cutoff of 23 deaths per 1000 people. A direct comparison of treatment districts with all other districts may yield biased estimates due several reasons. First, due to the infectious nature of several diseases prevalent at the time, the disease environment in a district may be indirectly impacted by the adoption of the Public Health Act in any of its neighboring districts. Second, the idea of 'yardstick competition' (first introduced by Ladd, 1992; Case, 1993; Case et al.,1993) suggests that public good provision in an area often generates a spillover in neighboring regions as local governments may try to mimic the good behaviors of their neighbors. Therefore, to account for these spatial spillover effects, districts contiguous to treatment districts (and themselves not eligible for treatment) are classified as spillover districts (91 districts out of 263) for this analysis, and the remaining are classified as control districts (111 districts out of 263)) as displayed in Appendix Figure B.1. The timing and spatial incidence of this health shock is plausibly exogenous for England and Wales. The state of infectious diseases and poor health had been an issue for several decades before 1848 but did not lead to any such provisions despite prior recommendations from Chadwick to improve hygiene and sanitation. Since this was the first-ever parliamentary act for the provision of public health in England and Wales, there was no precedence on how the state would encourage public health infrastructure improvement, making the incidence exogenous.

Table 2.1 provides summary statistics. Panel A reports the summary statistics for the mortality dataset. The average crude death rate from 1841-1880 is reported as 21.25 deaths per 1000 people. Table 2.1, Panel B summarizes the summary statistics for the father-son pairs from the sample from the I-CeM dataset. It suggests that about 24.5% of sons in the sample were in their early childhood in 1848 when the Act was introduced. About 23.8% of the sons in this sample lived in treatment districts, 34.6% in spillover districts, and the rest were in control districts. About 75.3% of the sons stayed in the same district in 1851 as well as in 1861. Based on the broader four-category classification, 37% of the sons have an occupation different from their father. About 22% of sons are in a better-ranked occupation relative to fathers. The absolute difference between the

---

is less than 23 deaths per 1000 people, i.e., districts that are currently classified as spillovers or control districts. In that case, the current estimates I present are lower bounds of the true estimates.

HISCLASS rank of father and son on average is 0.452. The mobility matrices in Tables B.2-B.5 provide additional descriptives by cross-classifying father and son occupations.[15]

2.4.1 Preliminary Analysis - Effect on Mortality

To identify the impact of the Public Health Act of 1848 on mortality in England and Wales, I use data on deaths in England and Wales from the period of 1841-1880.[16] The following difference-in-differences specification is estimated using ordinary least squares:

$$mortality_{dt} = \alpha + \beta_1 treat_d x\, post_t + \beta_2 spillover_d x\, post_t + \delta_d + \lambda_t + \varepsilon_{dt} \quad \ldots\ldots\ldots(1)$$

where $mortality_{dt}$ measures the crude death rate for district d in year t, post is 1 for years 1848-1880 and 0 otherwise, δ_d are district fixed effects and λ_t are year fixed effects. The variables $treat_d$ and $spillover_d$ take a value of 1 for treatment and spillover districts respectively. Control districts are the omitted category for this analysis. The coefficient β_1 is our estimate of interest as that indicates the impact of this act on mortality in treatment districts relative to control districts while accounting for potential spatial spillovers.

Table 2.2 suggests a reduction in the mortality rate in treated districts. Estimates in column 2 show the relevance of identifying the spillover districts and controlling for them in the specification as that increases the magnitude of the main coefficient. On average column 2 suggests that the Act reduced mortality by resulting in 1 lesser death per 1000 people. This is a 4.05% reduction in the crude death rate in treatment districts where the average crude death rate from 1841-1847 (before the Act came in place) was 25.52 deaths per 1000 people. This finding aligns

[15] These matrices report the pattern of intergenerational mobility in the linked father-son pairs across the three types of districts: treatment, spillover, control districts, and all districts in Table B.2, B.3, B.4 and B.5 respectively. There are important differences worth highlighting in these mobility matrices in treatment districts relative to spillover and control districts. First, the treatment districts see a larger movement of sons into the skilled/semi-skilled category conditional on the father's occupation (from 41-79% across columns in row 2). For spillover and control districts this value ranges from 28-75% and 24-70% respectively. Second, sons whose father's occupation was unskilled saw larger upward mobility in treatment districts relative to spillover and control districts. For instance, in treatment districts, 63% of sons whose father was unskilled became skilled/semiskilled and white-collar workers. In contrast, this fraction ranges from 34-38% in the spillover and control districts. This is also true for sons whose father's occupation was farming. They saw larger upward mobility in treatment districts than in spillover and control districts.

[16] Age-specific mortality data are not available at the district level. Therefore, I utilize a measure of aggregate mortality.

Table 2.1: Summary statistics

	Obs	Mean	Std Dev	Min	Max
Panel A: Mortality Statistics (Observations are at district-year level)					
Deaths	10,520	1,707.09	1,735.51	245.00	22,652.00
Population	10,520	77,859.18	71,617.33	14,128.56	998,074.60
Crude Death Rate	10,520	21.25	4.22	7.02	69.65
Treatment	10,520	0.23	0.42	0	1
Spillover	10,520	0.35	0.48	0	1
Note : 263 districts x 40 years = 10,520 observations					
Panel B: Linked I-CeM Dataset					
Son					
Age of son (in 1851)	608,624	8.99	4.19	2	17
Early childhood: In-utero to 2yo in 1848	608,624	0.24	0.43	0	1
Treatment	608,624	0.24	0.43	0	1
Spillover	608,624	0.39	0.49	0	1
Stayed in same district in 1851 & 1861	608,624	0.75	0.43	0	1
Occ. Rank of son in 1861	608,624	2.28	0.71	1	4
Upward Mobility	608,624	0.22	0.41	0	1
Occ. Category son ≠ Occ. Category father	608,624	0.37	0.48	0	1
| Occ. Rank son - Occ. Rank father|	608,624	0.45	0.66	0	3
Father					
Age of Household head in 1851	607,844	41.58	9.49	0.02	99.00
Number of Servants in 1851	608,118	0.20	0.72	0	17
Number of houseful members in 1851	608,624	7.11	9.63	2	2,680
HISCLASS Rank of father in 1851	608,624	2.38	0.79	1	4

Note: The sample in Panel A consists of district-year observations from 1841-1880. The sample in Panel B consists of father-son pairs living in districts in 55 counties. Sons are 2-17 years old when their father's occupation is measured in 1851 and 12-27 years old when their own occupation is measured in 1861. Measures of mobility are based on the four-category HISCLASS.

with Kiniria (2019) who estimates a 3.7% mortality rate reduction attributable to the Public Health Act in England and Wales.

Table 2.2: Effect of the Public Health Act on mortality

	(1)	(2)
treat x post	-0.846**	-1.034**
	(0.378)	(0.418)
Obs.	10520	10520

Note: The dependent variable is the mortality rate in the district. Each cell reports the coefficient of treat x post from equation (1). Both columns control for district and year fixed effects. Column 1 does not account for any spillovers. Column 2 includes a spillover district dummy and its interaction with earlychildhood. Standard errors are clustered at the district level in the parenthesis. ∗p<0.1; ∗∗p<0.05; ∗∗∗p<0.01.

2.4.2 Main Results – Effect on Intergenerational Mobility

Having established the effect of the Public Health Act on mortality, I turn to the labor market outcomes. To estimate the effect of the Public Health Act on intergenerational occupational mobility, I utilize the variation in the son's age in 1848 as well as spatial variation in treatment to determine the son's exposure to the act. There is burgeoning evidence in the literature suggesting a strong link between the first 1000 days of life since conception and later life outcomes.[17] These first days are a critical period for brain development which is the foundation of the child's eventual cognitive, emotional, and social development.[18] One of the key ingredients of healthy early childhood development is access to safe water, sanitation, and good hygiene. It prevents diseases and promotes adequate nutrition absorption in the body, a lack of which can translate into a lasting impact on growth and development (Barker, 1995; Dhamija & Sen, 2023; Victora et al., 2008).

[17] See Guerrant (2008), Victora et al. (2008), Victora et al. (2010), Doyle et al. (2009), Barham et al. (2013), Hoddinott et al. (2013).

[18] This is the period of fastest neural connection formation (up to 1 million per second). The brain is highly vulnerable and needs adequate nutrition during this sensitive period. Negative experiences can slow down and alter the formation of these neural connections in prefrontal region of the brain which can affect the child's ability to learn, solve problems and communicate. *Source*: https://www.unicef.org/sites/default/files/press-releases/glo-media-UNICEF_Early_Moments_Matter_for_Every_Child_report.pdf

Therefore, I hypothesize that for cohorts receiving new infrastructure in-utero and early post-natal (thus defined as exposure to the Act in the first 1000 days of life i.e., in early childhood) access to better sanitation, hygiene, and clean water through the Public Health Act will (net of culling) raise the returns to schooling, thus raising both attendance in childhood and the cognitive skill content of occupations in adulthood. In contrast, for cohorts exposed to the Act later in childhood, the intervention would act as a positive labor productivity shock that, by increasing the child's health capital in the short run, raises the opportunity cost of schooling and so speeds the transition into child labor (Costa, 2015). I first estimate the following equation using ordinary least squares:

$$f\left(Rank_{i,d,c,1861}^{son}, Rank_{i,d,c,1851}^{father}\right) = \alpha + \beta_1 treat_d \; x \; earlychildhood_i + \beta_2 treat_d + \beta_3 earlychildhood_i + \beta_4 X_{i,d,c,1851} + \rho_c + \varepsilon_{i,d,c,1851} \text{..........(2)}$$

The dependent variable takes several forms: (1) an indicator variable equal to one if the son works in a different HISCLASS occupation than his father, (2) an absolute difference between the HISCLASS rank between father and son, (3) an indicator variable equal to one if the son's HICLASS rank is better than his father's (upward mobility). The variable $treat_d$ takes the value of 1 for treatment districts and 0 otherwise. [19] In line with the importance of better hygiene, sanitation, and access to clean water during the first 1000 days of life, $earlychildhood_i$ is a dummy variable that takes a value of 1 for children in-utero and up to 2 years of age in 1848 at the time this act was introduced. The dummy takes a value of 0 for children aged 3 to 14 years in 1848. The coefficient β_1 is our estimate of interest as that indicates the impact on intergenerational mobility in treatment districts for children who had early childhood exposure to the Public Health Act of 1848. I include the variable $X_{i,d,c,1851}$ to control for household characteristics such as the size of the houseful.[20] To account for initial family wealth that may influence the intergenerational mobility outcomes for children, $X_{i,d,c,1851}$ also includes the number of servants, which is a proxy

[19] Due to a lack of data that details the exact location of the child in 1848, I assume that the child was in the same district in 1848 and 1851. This allows me to use the treatment status of his district in 1851 to analyze whether he was potentially exposed to health and sanitation improvements. To the extent that some of these children may not have been in the same district in 1848, this introduces measurement errors that I am unable to account for. However, it can be assumed that the size of this error would be small since migration rates within the gap of 3 years would be smaller than the migration rate I observed over a decade from 1851-1861 (which was approx. 25% in this sample).
[20] Includes household members, relatives and residential inmates like boarders, lodgers etc.

for wealth generally used in historical settings. ρ_c accounts for county fixed effects to net out any regional labor market trends.

Table 2.3: The effect of early childhood exposure to the Public Health Act on intergenerational mobility

	(1)	(2)	(3)
Panel A- Dependent Variable: Occ. Category son ≠ Occ. Category father			
Independent Variable			
treat x earlychildhood	0.017***	0.018***	0.018***
	(0.005)	(0.005)	(0.005)
Panel B- Dependent Variable: | Occ. Rank son - Occ. Rank father |			
treat x earlychildhood	0.037***	0.037***	0.036***
	(0.006)	(0.007)	(0.007)
Panel C- Dependent Variable: Occ. Category son > Occ. Category father			
treat x earlychildhood	0.030***	0.029***	0.029***
	(0.004)	(0.004)	(0.004)
Obs.	608,624	608,624	608,118
County FE	NO	YES	YES
HH Controls	NO	NO	YES

Note: Each cell represents the coefficient of treated x earlychildhood from equation (2). Column 1 reports the coefficient of equation (2). Column 2 includes county fixed effects. Column 3 controls for household characteristics. Standard errors are clustered at the district in the parenthesis. ∗p<0.1; ∗∗p<0.05; ∗∗∗p<0.01.

Table 2.3 reports the estimates for all outcome variables. Estimates in column 1 are obtained without including the county fixed effects and the household controls. Column 2 includes county fixed effects and column 3 further adds the household controls. Results across all columns and outcome variables suggest that exposure to the Public Health Act in early childhood led to a positive and significant impact on intergenerational occupational mobility. It led to severing ties between father and son's occupation (Panel A), a movement further away from the father's occupational rank (Panel B), and a movement of the son to a higher-ranked occupation relative to his father (i.e., upward intergenerational mobility) (Panel C).

However, the specification in equation (2) does not account for spillovers. As previously discussed, the infectious nature of diseases means a district's environment can be affected indirectly by neighboring districts' Public Health Act adoption. Additionally, 'yardstick competition' posits that public goods provision can spill over into neighboring regions as governments mimic positive behaviors. To address these concerns, I modify the specification to control for contiguous non-treatment districts classified as spillover districts. Therefore, the full specification is the following:

$$f\left(Rank^{son}_{i,d,c,1861}, Rank^{father}_{i,d,c,1851}\right) = \alpha + \beta_1 treat_d \; x \; earlychildhood_i + \beta_2 spillover_d x \; earlychildhood_i + \beta_3 treat_d + \beta_4 spillover_d + \beta_5 earlychildhood_i + \beta_6 X_{i,d,c,1851} + \rho_c + \varepsilon_{i,d,c,1851} \; \ldots\ldots\ldots.. (3)$$

The variables $treat_d$ and $spillover_d$ take the value of 1 for treatment and spillover districts respectively and 0 otherwise. Control districts are the omitted category for this analysis. All other variables are as described before.

The estimates of β_1 so obtained are reported in Table 2.4. Estimates in column 1 are obtained without including the county fixed effects and the household controls. Columns 2 and 3 include additional controls such as county fixed effects and household characteristics respectively. Column 3 is the baseline specification for the rest of the analysis, as described above in equation (3). Overall, the estimates in Table 2.4 are larger than the estimates in Table 2.3 which underscores the importance of accounting for spillover effects in this setting. Estimates from column 3 suggest that sons with early childhood exposure to the Act in treatment districts are 2.2 percentage points (Panel A) more likely to be in a different occupational group and 3.6 percentage points (Panel C) more likely to be in a better-ranked occupational group than their fathers or upward mobile as compared to sons who were in early childhood in 1848 in control districts. Based on the sample average reported in Table 2.1, this translates into a 5% increase in the probability of the younger sons working in a different occupational group than his father and a 16% increase in upward mobility. These estimates are comparable to some of the other estimates in the literature. For instance, Costas-Fernández et al. (2020) find a 5% increase in the probability of sons being in a different occupational group in England and Wales using data from 1881-1901 in response to proximity to railways. In another study, Karbownik and Wray (2021) find that poor childhood health leads to a

10% reduction in upward mobility in London using inpatient records from three large hospitals between 1870-1902.

Table 2.4: The effect of early childhood exposure to the Public Health Act on intergenerational mobility, complete specification

	(1)	(2)	(3)
Panel A- Dependent Variable: Occ. Category son ≠ Occ. Category father			
Independent Variable	0.022***	0.023***	0.022***
treat x earlychildhood	(0.006)	(0.006)	(0.006)
Panel B- Dependent Variable: | Occ. Rank son - Occ. Rank father |			
treat x earlychildhood	0.045***	0.045***	0.044***
	(0.007)	(0.007)	(0.007)
Panel C- Dependent Variable: Occ. Category son > Occ. Category father			
treat x earlychildhood	0.037***	0.036***	0.036***
	(0.005)	(0.005)	(0.005)
Obs.	608,624	608,624	608,118
Spillover Controls	YES	YES	YES
County FE	NO	YES	YES
HH Controls	NO	NO	YES

Note: Each cell represents the coefficient of treated x earlychildhood from equation (3). Column 1 reports the coefficient of equation (3). Column 2 includes county fixed effects. Column 3 controls for household characteristics. Standard errors are clustered at the district in the parenthesis. ∗p<0.1; ∗∗p<0.05; ∗∗∗p<0.01.

To the extent that I cannot observe the exact location/timing of the local health board establishment due to lack of access to such data, these results should be interpreted with caution. If districts proxied as treated did not get local boards of health until much later in the 1850s, the estimates I present may be biased upwards.

2.5 Robustness

In this section, I provide evidence of the robustness of the main results on intergenerational mobility. First, I show that the baseline estimates of labor market outcomes are robust to performing the analysis based on the finer twelve HISCLASS categories instead of being collapsed into four broader groups. While there are advantages to analyzing the four broader categories, Table 2.5 column 1 reports that the qualitative results remain unchanged.[21]

Second, I replicate the analysis with historical Cambridge Social Interaction and Stratification Scale (HISCAM) occupational rankings, which provides a continuous measure of social status. The scale provides 359 unique scores, with a higher score indicating a better position in society. Critics argue that large classes hide too much variation, and microclasses are needed to uncover that. On the other hand, proponents of large classes (like HISCLASS) have argued that it reflects the true fuzziness of social stratification (van Leeuwen, 2020). Using the continuous HISCAM scale also results in a large reduction in sample size as this scale does not rank all occupations in the census data that can be ranked using HISCLASS. Results in Table 2.5 column 2 report estimates using HISCAM rankings and are qualitatively similar to the main estimates. I also re-estimate the main coefficients using this subsample employing the four-category HISCLASS classification and report it in column 3. These estimates are very similar to the main estimates and suggest that the results obtained in column 2 are not biased due to a smaller sample.

Third, I adopt an alternate method to account for spillover effects. Following Pérez (2018), I re-run my main specification on a sample that excludes neighbors of treated districts (i.e., the spillover districts). The results are reported in Table 2.5 column 4 and are qualitatively and quantitatively comparable to the baseline results.

Fourth, I re-run my main specification on a sample that drops observations from London as it was exempt from this act. The results in column 5 of Table 2.5 remain robust. This suggests inclusion of London in the baseline sample is not driving the results.

Fifth, I confirm that the labor market results are robust to an alternative definition of early childhood. The results are robust to redefining early childhood to include children aged 3 and

[21] Long and Ferrie (2013) discuss that using four categories reduced the sparseness of the mobility tables. Fewer categories also allow for an easier analysis of occupational transitions as done in Section 2.6 below.

under, instead of restricting to children aged 2 and under. Estimates in Table 2.5 column 6 are significant and qualitatively similar. However, note that the coefficients drop in magnitude relative to the baseline specification in Table 2.4 column 3 implying that while the results are robust and remain significant, they were driven by children aged 2 and under. This is in line with the scientific theory that suggests that days up to the 2nd birthday are the most crucial.

Sixth, I run a placebo check by randomizing treatment. I randomly assign treatment to 61 out of 263 districts, identify spillover districts, and re-estimate equation (3). I do this 1000 times and plot the distribution of coefficients along with my baseline results in Appendix Figures B.2-B.4. The estimates are concentrated around zero and column 1 in Appendix Table B.6 suggests that randomization estimates almost never outperform the main estimates of the regressions. So, it is unlikely that these estimates are generated by chance and are not due to any underlying differences in mobility across younger and older age groups.

Finally, I run another placebo check by considering sons aged 3-8 in 1848 as the placebo early childhood group and ages 9-14 as the later childhood group. Estimates in Table 2.5 column 7 find no evidence of increased intergenerational mobility for this placebo age group. This indicates that the results are not due to any underlying differences in mobility in the treatment, spillover, and control districts.

2.6 What led to the upward occupational mobility?

Sons exposed to the Public Health Act in treated districts in early childhood were more likely to be in an occupation different than their father's and were more likely to be upward mobile. I now explore the occupational transitions that led to this upward mobility. I estimate the following ordinary least squares specification:

$$\left(OCC_{i,d,c,1861}^{son} = k \;/\; OCC_{i,d,c,1851}^{father} = m\right) = \alpha + \beta_1 treat_d \; x \; earlychildhood_i + \beta_2 spillover_d x \; earlychildhood_i + \beta_3 treat_d + \beta_4 spillover_d + \beta_5 earlychildhood_i + \beta_6 X_{i,d,c,1851} + \rho_c + \varepsilon_{i,d,c,1851} \;\ldots\ldots\ldots\ldots (4)$$

where $\left(OCC_{i,d,c,1861}^{son} = k \;/\; OCC_{i,d,c,1851}^{father} = m\right)$ is 1 if a son of a father in occupational category

Table 2.5: The effect of early childhood exposure to the Public Health Act on intergenerational mobility, robustness

	12 HISCLASS categories	HISCAM	HISCLASS on HISCAM subsample	Drops spillover districts	Drops London	Redefine earlychildhood	Age placebo check
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A- Dependent Variable: Occ. Category son ≠ Occ. Category father							
Independent Variable							
treat x earlychildhood	0.024***	0.061***	0.034***	0.023***	0.018***	0.017***	-0.004
	(0.007)	(0.012)	(0.007)	(0.006)	(0.006)	(0.005)	(0.005)
Panel B- Dependent Variable: | Occ. Rank son - Occ. Rank father |							
treat x earlychildhood	0.088***	0.028	0.055***	0.045***	0.038***	0.036***	0.007
	(0.024)	(0.073)	(0.009)	(0.007)	(0.007)	(0.007)	(0.006)
Panel C- Dependent Variable: Occ. Category son > Occ. Category father							
treat x earlychildhood	0.014**	0.020*	0.040***	0.036***	0.033***	0.031***	-0.002
	(0.006)	(0.010)	(0.006)	(0.005)	(0.005)	(0.004)	(0.005)
Obs.	608,118	345,684	345,684	372,341	573,903	608,118	459,260
Spillover Controls	YES	YES	YES	NO	YES	YES	YES
County FE	YES	YES	YES	YES	YES	YES	YES
HH Controls	YES	YES	YES	YES	YES	YES	YES

Note: Each cell represents the coefficient of treated x earlychildhood from equation (3). Column 1 constructs dependent variables based on the finer 12 HISCLASS categories. Column 2 constructs dependent variables using HISCAM occupational rankings and Column 3 reruns the four-category HISCLASS outcomes on the HISCAM subsample. Column 4 excludes spillover districts from the sample. Column 5 excludes London observations from the sample. Column 6 recategorizes early childhood exposure to the act as including children in utero to 3 years of age. Column 7 performs a placebo test by excluding in-utero to 2-year-olds from the sample and categorizing early childhood exposure as including 3 to 8-year-olds relative to 9 to 14-year-olds as the older age group. Standard errors are clustered at the district in the parenthesis. ∗p<0.1; ∗∗p<0.05; ∗∗∗p<0.01.

m would work in occupational category k, and 0 if the son works in a category other than k. The remaining variables are as defined previously.

Table 2.6 reports the main coefficient β_1 for occupation transitions that comprised upward mobility. The table shows how exposure to the Public Health Act in early childhood related to the likelihood of different father-son occupational transitions. A positive value of β_1 implies that a given transition was more likely in treated districts in younger sons, while a negative value implies it was less likely. Results suggest that upward mobility is driven by sons of farmers and unskilled workers who moved away from farming and unskilled occupations into the skilled/semiskilled category. This is consistent with the hypothesis of increased skill content for these younger sons relative to older sons upon exposure to improved hygiene and sanitation due to the Public Health Act of 1848. Exposure to the Public Health Act in early childhood likely led to a better trajectory of short-term and long-term health, which can change the cost of attaining skills for these children. These findings are similar to Pérez (2018) who finds that children are more likely to transition out of farming than adults in 1869 in Argentina as moving out of farming implied a transition to occupations that required higher skills which adults lacked.

In the absence of any direct measure of literacy such as years of schooling obtained in the census data, it is difficult to analyze if the observed upward mobility is due to increased investment in literacy of the younger cohort. However, using job advertisements published in nineteenth-century English periodicals as well as other contemporaneous descriptions of occupations, Mitch (1992) classifies jobs into four categories: literacy required, literacy likely to be useful, and possible (or ambiguous) unlikely to use literacy. An overall analysis of this classification suggests that jobs that would require literacy were those of the banking sector, accounts, insurance, and management and some that were unlikely to use literacy would be farming, fishermen, laborer in coal yards, coal haulers, etc. Based on this, it is likely that white collar and skilled/semiskilled job categories in my analysis would map to the jobs that required literacy or where literacy was likely to be useful and unskilled or farming occupations where literacy was unlikely to be useful. For ease of interpretation, I combine the first two categories from Mitch (1992) and the last two to create a binary classification. This is similar to the proxy used by Milner (2021) for evaluating literacy in the 1870s in England and Wales. He finds a high correlation of approximately 0.6 between the binary literacy variable thus created and the one that uses finer details of Mitch's

Table 2.6: Upward mobility occupational transitions

	(1)	(2)	(3)	(4)	(5)	(6)
	Father's Occupation					
	Farmer			Unskilled		Skilled/Semiskilled
Son's Occupation	Unskilled	Skilled/Semiskilled	White Collar	Skilled/Semiskilled	White Collar	White Collar
Treated x Early childhood	-0.060***	0.060***	0.010	0.097***	-0.007	-0.006*
	(0.016)	(0.020)	(0.010)	(0.015)	(0.005)	(0.003)
Obs.	51,347	51,347	51,347	192,145	192,145	298,195

Note: The dependent variable is an indicator equal to one if the son works in a specific occupation conditional on the father's occupation. All cells report coefficients of treated x earlychildhood from equation (3). All regressions account for spillover effects and include county fixed effects as well as household controls. Standard errors are clustered at the district in the parenthesis. ∗p<0.1; ∗∗p<0.05; ∗∗∗p<0.01.

classification using string matching with occupational descriptions (suggesting that the binary classification would also be a good proxy). Therefore, previous literature links the type of occupation itself as an indication of whether or not you are likely to be literate. I run the following regression to test if sons exposed to the Act in early childhood are more likely to be in jobs that require literacy to measure if they were investing more in education:

$$f\ (job\_req\_literacy_{i,d,c}) = \alpha + \beta_1 treat_d\ x\ earlychildhood_i + \beta_2 spillover_d x\ earlychildhood_i + \beta_3 treat_d + \beta_4 spillover_d + \beta_5 earlychildhood_i + \beta_6 X_{i,d,c,1851} + \rho_c + \varepsilon_{i,d,c,1851} \ \ldots\ldots\ (5)$$

The dependent variable takes two forms: (1) $job\_req\_literacy_{i,d,c}$, which is 1 if the sons is in a job that required literacy or was likely to use literacy (i.e., white collar and skilled/semiskilled jobs based on Mitch (1992)) and 0 otherwise. (2) $upward\_literacy\_mobility_{i,d,c}$, which is 1 if the son is in a job that is likely to require literacy, but his father was not and 0 otherwise. The remaining variables are as defined previously and β_1 is the relevant coefficient.

Table 2.7: The effect of early childhood exposure to the Public Health Act on the likelihood of being in a job that requires literacy and on the likelihood of being a mover

	(1)
Panel A- Dependent Variable: Job Requires Literacy	
treat x earlychildhood	0.111***
	(0.010)
Panel B- Dependent Variable: Literacy Mobility	
treat x earlychildhood	0.043***
	(0.005)
Panel C- Dependent Variable: Mover	
treat x earlychildhood	0.047***
	(0.006)
Obs.	608,118

Note: The columns report the coefficient of treated x earlychildhood from equations (5) and (6). All regressions account for spillover effects and include county fixed effects as well as household controls. Standard errors are clustered at the district in the parenthesis. ∗p<0.1; ∗∗p<0.05; ∗∗∗p<0.01 ∗∗∗p<0.01.

Table 2.7 reports that the coefficient in Panel A is positive and significant which provides suggestive evidence that sons with early childhood exposure to better sanitation and hygiene were more likely to be in a job that required them to be literate. Further, the coefficient in Panel B is also positive and significant which suggests that these sons were also more likely to be in literacy-requiring jobs relative to their fathers. It is implicit that they were likely to invest more in literacy which led them to a better-ranked occupation than their father. Therefore, early-life positive health shocks influence later-life human capital formation.

If one way that better sanitation led to upward mobility is by facilitating geographic mobility, migration can also be a potential mechanism. Access to improved WASH facilities in early childhood can influence the probability of migrating to a better job for several reasons. Lack of sanitation and hygiene leading to poor health can directly affect one's physical capacity to move and settle in a new city. It can reduce the accumulation of human capital and productivity of the individual, reducing economic gains from migration. Thiede et al. (2022) emphasize that poor early life conditions can undermine socio-economic attainment and effectively trap individuals in places of birth through negative effects on human capital development. This implies that migration patterns may reflect the enduring developmental consequences of early life exposure to health interventions such as the sanitation facilities that the act provided (Hecht & McArthur, 2023).[1] In a related study, Ramírez-Luzuriaga et al. (2021) find that early life improved nutrition and health led to an increased probability of migration in Guatemala. Their results suggest that an individual's height-for-age during childhood (which captures exposure to early life shocks) was positively associated with the odds of moving in later years. They suggest that a positive health shock in early childhood raises overall human capital formation via increased years of education. This opens better job prospects which increases the probability of migration. Similarly, Logan (2009) suggests that poor health was one of the factors that reduced the geographic mobility of African Americans in the nineteenth century.

Therefore, to investigate if exposure to the Public Health Act led to an increased probability of migration for sons, I run the following regression by ordinary least squares:

[1] While migration may occur as a response to poor conditions of sanitation or high mortality attributes of the districts, such a response is not likely to differ by age at exposure to the Public Health Act and is thus not likely to bias these estimates.

$$mover_{i,d,c} = \alpha + \beta_1 treat_d \; x \; earlychildhood_i + \beta_2 spillover_d x \; earlychildhood_i + \beta_3 treat_d + \beta_4 spillover_d + \beta_5 earlychildhood_i + \beta_6 X_{i,d,c,1851} + \rho_c + \varepsilon_{i,d,c,1851} \; \ldots\ldots\ldots (6)$$

where mover is an indicator variable with a value equal to 1 if the son is not in his childhood district in 1861 and the rest of the variables are as defined previously. Results in Table 2.7 Panel C indicate that sons exposed to Act in early childhood were 4.7 percentage points more likely to move. Next, I run equation (3) separately for movers and stayers. Column 1 of Table 2.8 reports the coefficient for the entire sample, column 2 reports the coefficient for stayers, and column 3 reports it for movers. The estimate in column 3 is larger than the estimate in column 2, suggesting that migration allowed people to be more upward mobile and thus is one potential channel in this case.

Table 2.8: Effect of early childhood exposure to the Public Health Act on mobility: stayers vs. movers

	All (1)	Stayers (2)	Movers (3)
Dependent Variable - Upward Mobility			
Independent Variable			
treat x earlychildhood	0.036***	0.029***	0.038***
	(0.005)	(0.005)	(0.008)
Obs.	608,118	458,033	150,085

Note: Each cell represents the coefficient of treated x earlychildhood from equations (2). Column 1 reports the estimates from the total sample. Column 2 reports estimates for the stayers. Column 3 reports the estimates for movers. All regressions account for spillover effects and include county fixed effects as well as household controls. Standard errors are clustered at the district in the parenthesis. ∗p<0.1; ∗∗p<0.05; ∗∗∗p<0.01.

This result is in line with Buscha et al. (2021) who a positive relationship between spatial migration and upward mobility in 20th century England and Wales. Ward (2022) also finds that sons can improve job outcomes relative to fathers by moving to better opportunities elsewhere. Among other mechanisms, Karbownik and Wray (2021) evaluate the role of the persistence of poor health in adulthood due to poor early life in preventing intergenerational mobility. They analyze hospital visit records from childhood and find that poor health led to an increased likelihood of physical disabilities in adulthood and thus made upward mobility less likely. Due to the lack of such data, this is beyond the scope of the present study. If historical records of hospitalization become available for this period, future research can explore the role of physical and mental disabilities as a channel in this setting.

2.7 Conclusion

The Public Health Act of 1848 marked the first attempt of England and Wales at a systematic sanitation improvement in the context of a grim situation of public health in the early nineteenth century. This study finds that exposure to sanitation improvements implemented by the Public Health Act of 1848 resulted in lower mortality. However, many intrinsic improvements in health such as reduced headaches, diminished symptoms of dyspepsia, and a decrease in overall sickness are not captured by this crude measure. Lack of data in historical contexts often prevents examination of a more comprehensive measure of health and ability. Using a novel data set, this study contributes to the literature by evaluating the impact on a more sensitive measure, i.e., intergenerational mobility.

Using over half a million father-son pairs from historical census data linked from 1851 to 1861, I document the evidence of a relationship between intergenerational occupational mobility and the provision of sanitation infrastructure, better hygiene, and clean water. Utilizing the variation in age at exposure to the act and the spatial variation in treatment, results suggest that early childhood exposure to positive health shocks had a meaningful long-term impact. It led to severing ties between father and son's occupation, a movement further away from father's occupational rank, and increased upward intergenerational mobility. Children from treatment districts who were exposed to the Act in early childhood are 5% more likely to be in a different occupational category than their father and 16% more likely to be upward mobile relative to their counterparts in the control districts. An analysis of occupational transitions that are contributing to this improved upward mobility for sons exposed to the act in early childhood finds evidence of a movement out of farming and unskilled jobs into skilled/semiskilled jobs. In the absence of any direct measure of literacy in the census, I utilize the description of literacy requirement in nineteenth-century jobs by Mitch (1992). I find that sons with early childhood exposure to the act are more likely to be engaged in jobs that required literacy. This implies that their exposure to the act and different returns to gaining further human capital led them to invest more in literacy and thus facilitated a better-ranked occupation than their father. Additionally, exposure to the act in early childhood also increased the geographic mobility of the sons. Returns from moving are larger than from staying for upward mobility, suggesting that migration is a partial mechanism. Additional exploration of mechanisms are avenue for future research.

Similar health environments as those of nineteenth-century England and Wales remain a reality in the developing world in rapidly growing cities (Fogel, 2004; Floud et al., 2011; Currie & Vogl, 2013; Mercer, 2014). Health shocks remain a significant threat to economic opportunities in these countries (Gertler & Gruber, 2002). Thus, while the study is rooted in history, it can guide current discussions. Access to public health interventions in early childhood may contribute to increased productivity and higher human capital for children upon reaching adulthood. The findings of this study thus have broader implications and may serve as a relevant guide to future policymakers.

2.8 References

Almond, D., & Currie, J. (2011a). Human Capital Development Before Age 5. In Ashenfelter, Orley and David Card, eds. *Handbook of Labor Economics*. New York: Elsevier.

Almond, D., & Currie, J. (2011b). Killing Me Softly: The Fetal Origins Hypothesis. *Journal of Economic Perspectives*. 25 (2011b): 153-72.

Alsan, M. & Goldin, C. (2018). Watersheds in Child Mortality: The Role of Effective Water and Sewerage Infrastructure, 1880 to 1920. *Journal of Political Economy*.

Andrade, E.,Veloso, F., Madalozzo, R. & Ferreira, S.G. (2003). *Do Borrowing Constraints Decrease Intergenerational Mobility in Brazil? A Test Using Quantile Regression*. Insper Working Paper WPE: 045/2003.

Arthi, V., Beach, B. & Hanlon, W.W. (2022). Recessions, Mortality, and Migration Bias: Evidence from the Lancashire Cotton Famine. *American Economic Journal: Applied Economics*.

Atkinson, A. B., Maynard, A. K. & Trinder, C. G., (1983). *Parents and Children: Incomes in Two Generations,* London: Heinemann

Augsburg, B. & Rodríguez-Lesmes, P. A., (2018). Sanitation and child health in India. *World Development*, *Elsevier*, vol. 107(C), pages 22-39.

Barham, T., Macours, K. & Maluccio. J.A. (2013). Boys' Cognitive Skill Formation and Physical Growth: Long-Term Experimental Evidence on Critical Ages for Early Childhood Interventions. *American Economic Review*, 103 (3): 467-71

Barker D. J. (1995). Fetal origins of coronary heart disease. *BMJ* (Clinical research ed.), 311(6998), 171–174.

Bhalotra, S.R. & Rawlings, S. (2013). Gradients of the Intergenerational Transmission of Health in Developing Countries. *The Review of Economics and Statistics*, MIT Press, vol. 95(2), pages 660-672.

Bhalotra, S.R. & Venkataramani, A. (2015). *Shadows of the Captain of the Men of Death: Early Life Health Interventions, Human Capital Investments, and Institutions.* SSRN Working Paper.

Boberg-Fazlic, N. & Sharp, P.(2013). *North and South: Social mobility and welfare spending in preindustrial England.* European Historical Economics Society (EHES) Working Paper 37

Brown, D.W., Kowalski, A.E. & Lurie, I.Z. (2020). Long-Term Impacts of Childhood Medicaid Expansions on Outcomes in Adulthood. *Review of Economic Studies*, Vol. 87, No. 2, pp. 792–821.

Buscha, F., Gorman, E., & Sturgis, P. (2021). Spatial and Social Mobility in England and Wales: A Sub-national Analysis of Differences and Trends Over Time. *The British Journal of Sociology*, 72(5), 1378–1393.

Bütikofer, A., Løken, K. V., & Salvanes, K. G. (2019). Infant health care and long-term outcomes. *Review of Economics and Statistics*, *101*(2), 341-354.

CAP. LXIII. An Act for Promoting the Public Health, 31st August 1848 (Legal Document)

Case, A. (1993). Interstate Tax Competition After TRA86. *Journal of Policy Analysis and Management* 12, 136–148.

Case, A. C., H. S. Rosen & J. R. Hines. (1993). Budget Spillovers and Fiscal Policy Interdependence. *Journal of Public Economics* 52, 285–307.

Caruso, G. (2017). The Legacy of Natural Disasters: The Intergenerational Impact of 100 Years of Disasters in Latin America. *Journal of Development Economics.*

Chapman, J. (2019). The Contribution of Infrastructure Investment to Mortality Decline: Evidence from England and Wales, 1861-1900. *Economic History Review* 72, no. 1: 233–59.

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States. *The Quarterly Journal of Economics* 129, 1553–1623.

Chetty, R., Hendren, N., & Katz, L.F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review* 106, 855– 902.

Clark, G. & Cummins, N. (2015). Intergenerational Wealth Mobility in England, 1858–2012: surnames and social mobility. *The Economic Journal* 125, 61–85.

Conti, G. & Heckman, J. (2013). The Developmental Approach to Child and Adult Health. *Pediatrics,* 131 (2): S133-S141.

Costa, D. L. (2015). Health and the Economy in the United States from 1750 to the Present. *Journal of Economic Literature*, 53 (3): 503-70.

Costas-Fernández, J., Guerra, J.A. & Mohnen, M. (2020). *Train to Opportunity: The Effect of Infrastructure on Intergenerational Mobility.* Documento CEDE No. 48.

Currie, J. (2009). Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development. *Journal of Economic Literature*, 47 (1): 87-122.

Currie, J. & Vogl, T. (2013). Early-Life Health and Adult Circumstance in Developing Countries. *Annual Review of Economics*, Vol. 5, No. 1, pp. 1–36.

Cutler, D., & Miller, G. (2005). The role of public health improvements in health advances in the twentieth century United States. *Demography*, 42, 1–22.

Cutler, D., Deaton, A., & Lleras-Muney, A. (2006). The Determinants of Mortality. *Journal of Economic Perspectives,* 20 (3): 97-120.

Crimmins, E.M. & Finch, C. (2006). Infection, Inflammation, Height, and Longevity. *Proceedings of the National Academy of Sciences* 103 (2): 498-503.

Dearden, L., Machin, S., & Reed, H. (1997). Intergenerational Mobility in Britain. *The Economic Journal*, *107*(440), 47–66.

Dhamija, G., & Sen, G. (2023). Lasting impact on health from natural disasters, potential mechanisms, and mitigating effects. *Environment and Development Economics,* 28(1), 1-24.

Doyle, O., Harmon, C.P., Heckman, J.J. & Tremblay, R.E. (2009). Investing in Early Human Development: Timing and Economic Efficiency. *Economics and human biology*, 7(1), 1–6.

Eppig, C., Fincher, C., & Thornhill, R. (2010). Parasite Prevalence and the Worldwide Distribution of Cognitive Ability. *Proceedings of the Royal Society B: Biological Sciences,* 277 (1701): 3801-8.

Esrey, S. A., Potash, J. B., Roberts, L., & Shiff, C. (1991). Effects of Improved Water Supply and Sanitation on Ascariasis, Diarrhoea, Dracunculiasis, Hookworm infection, Schistosomiasis, and Trachoma. *Bulletin of the World Health Organization,* 69(5), 609–621.

Fee, E. & Brown, T.M. (2005). The Public Health Act of 1848. *Bulletin of the World Health Organization, 83* (11), 866 - 867. World Health Organization.

Ferrie, J., & Troesken W. (2005). *Death and the City: Chicago's Mortality Transition, 1850–1925*. Working Paper # 11427. National Bureau of Economic Research. Cambridge, Massachusetts.

Ferrie, J., & Troesken, W. (2008). Water and Chicago's mortality transition, 1850-1925. *Explorations in Economic History*, 45, 1–16.

Floud, R., Fogel, R.W., Harris, B. & Hong, S.C. (2011). *The Changing Body: Health, Nutrition, and Human Development in the Western World since 1700*, Cambridge: Cambridge University Press.

Fogel, R. (1986). Nutrition and the decline in mortality since 1700: Some preliminary findings, In S. Engerman and R. Gallman, eds., *Long-term factors in American economic growth*. Chicago: University of Chicago Press, 439–555.

Fogel, R. W. (2004). *The Escape from Hunger and Premature Death, 1700-2100,* Cambridge: Cambridge University Press.

Gamper-Rabindran, S., Khan, S. &Timmins, C., (2010). The impact of piped water provision on infant mortality in Brazil: A quantile panel data approach. *Journal of Development Economics*, Elsevier, vol. 92(2), pages 188-200, July.

Gertler, P., & Gruber, J. (2002). Insuring Consumption Against Illness. *American Economic Review*, 92 (1): 51-70.

Goodman-Bacon, A. (2017). *The Long-Run Effects of Childhood Insurance Coverage: Medicaid Implementation, Adult Health and Labor Market Outcomes.* Working Paper.

Gould, E.D., Lavy, V. & Paserman, M.D. (2011). Sixty Years After the Magic Carpet Ride: The Long-Run Effect of the Early Childhood Environment on Social and Economic Outcomes. *Review of Economic Studies*, Vol. 78, No. 3, pp. 938–973.

Guerra, J.A. & Mohnen, M. (2022). Multinomial Choice with Social Interactions: Occupations in Victorian London. *The Review of Economics and Statistics*.

Guerrant, R. L., Oriá, R. B., Moore, S. R., Oriá, M. O., & Lima, A. A. (2008). Malnutrition as an Enteric Infectious Disease with Long-Term Effects on Child Development. *Nutrition Reviews*, 66(9), 487–505.

Gutiérrez, J.P., Shamah-Levy, T., Bertozzi, S. & Rivera-Dommarco, J.A. (2019) *Intergenerational Social Mobility Based on the Investments in Human Capital: Evidence of the Long-Term Results of Prospera in Health.* World Bank Policy Research Working Paper No. 9001

Habicht, J. P., Martorell, R., Yarbrough, C., Malina, R. M., & Klein, R. E. (1974). Height and weight standards for preschool children. How relevant are ethnic differences in growth potential?. *Lancet (London, England)*, *1*(7858), 611–614. https://doi.org/10.1016/s0140-6736(74)92663-4

Hamlin, C., & Sidley, P. (1998). Revolutions in public health: 1848, and 1998?. *BMJ*, 317(7158), 587-591.

Hecht, K., & McArthur, D. (2023). Moving on up? How Social Origins Shape Geographic Mobility within Britain's Higher Managerial and Professional Occupations. *Sociology*.

Heckman, J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children. *Science*, Vol. 312, No. 5782, pp. 1900–1902

Hoddinott, J., Alderman, H., Behrman, J.R., Haddad, L., & Horton, S. (2013).The economic rationale for investing in stunting reduction. *Maternal & child nutrition.*

Jayachandran, S., Lleras-Muney, A. & Smith, K.V. (2010). Modern Medicine and the Twentieth Century Decline in Mortality: Evidence on the Impact of Sulfa Drugs. *American Economic Journal: Applied Economics*, Vol. 2, No. 2, pp. 118–146.

Jones, M.R., Simeonova, E. & Akee, R. (2020). *The EITC and Intergenerational Mobility.* Working Papers 20-35, Center for Economic Studies, U.S. Census Bureau.

Kiniria, M. D. (2019). Essays on Social Policy in Victorian England. PhD Dissertation, Cornell University.

Karbownik, K. & Wray, A. (2021). *Educational, Labor-Market and Intergenerational Consequences of Poor Childhood Health*. IZA Discussion Paper No. 14127.

Logan, T. D. (2009). Health, Human Capital, and African-American Migration Before 1910. *Explorations in Economic History,* 46 (2): 169-85

Long, J. & Ferrie, J. (2013). Intergenerational Occupational Mobility in Great Britain and the United States since 1850. *American Economic Review*, 103 (4): 1109-37.

Lumley, W.G. (1859). *The New Sanitary Laws: The Public Health Act, 1848, and the Local Government Act, 1858; An Introduction, Notes, and Index.* Shaw & Sons.

McKeown, T., & Brown, R. G. (1955). Medical Evidence Related to English Population Changes in the Eighteenth Century. *Population Studies*, 9, 119–141.

McKeown, T., & Record, R. G. (1962). Reasons for the Decline of Mortality in England and Wales during the Nineteenth Century. *Population Studies*, *16*(2), 94–122.

McKeown, T. (1976). *The modern rise of population.* London: Edward Arnold

Meckel, R.A. (1985). Immigration, Mortality and Population Growth in Boston, 1840-1880. *Journal of Interdisciplinary History* 15:393-417.

Meeker, E. (1972). The Improving Health of the United States, 1850-1915. *Explorations in Economic History*. 9: 353-374.

Mercer, A. (2014). *Infections, Chronic Disease and the Epidemiological Transition: A New Perspective*. Rochester, NY: Rochester University Press.

Millward R. & Bell, F. (1998). Public Health Expenditures and Mortality in England and Wales, 1870–1914, *Continuity and Change* 13(2): 221–249.

Milner, B. (2021). *The Impact of State-Provided Education: Evidence from the 1870 Education Act,* mimeo.

Mitch, D. (1992). *The Rise of Popular Literacy in Victorian England: The Influence of Private Choice and Public Policy.* University of Pennsylvania Press.

Mooney, G. (2015). *Intrusive interventions: Public Health, Domestic Space and Infectious Disease Surveillance in England, 1840-1914*. Rochester, NY: Rochester University Press.

Olivetti, C. & Paserman, M.D. (2015). In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850-1940. *American Economic Review* 105, 2695–2724.

Orero, M., Heime, C. , Cutler, S. & Mohaupt, S. (2007). *The Impact of Conflict on the Intergenerational Transmission of Chronic Poverty: An Overview and Annotated Bibliography*. Chronic Poverty Research Centre Working Paper No. 71

Osterbacka, E. (2001). Family Background and Economic Status in Finland. *Scandinavian Journal of Economics.* September, 103:3, pp. 467- 84.

Pérez, S. (2018). *Railroads and the Rural to Urban Transition: Evidence from 19th-Century Argentina*, mimeo

Piketty, T.(2000). Theories of Persistent Inequality and Intergenerational Mobility. *Handbook of Income Distribution* 1, 429–476.

Preston, S. H., & van de Walle, E. (1978). Urban French Mortality in the Nineteenth Century. *Population Studies*, 32(2), 275–297.

Prüss-Ustün, A., Wolf, J., Bartram, J., Clasen, T., Cumming, O., Freeman, M. C., Gordon, B., Hunter, P. R., Medlicott, K., & Johnston, R. (2019). Burden of Disease from Inadequate Water, Sanitation and Hygiene for Selected Adverse Health Outcomes: An Updated Analysis with a Focus on Low- and Middle-Income Countries. *International Journal of Hygiene and Environmental Health*, 222(5), 765–777.

Ramírez-Luzuriaga, M. J. & Hoddinott, J. & Martorell, R. & Ramirez-Zea, M. & Stein, A. (2021). Early-Life Nutrition and Subsequent International Migration: A Prospective Study in Rural Guatemala. *The Journal of Nutrition.* 151. 10.1093/jn/nxaa379.

Razzell, P. (1993). The Growth of Population in Eighteenth-Century England: A Critical Reappraisal. *Journal of Economic History*, 53(4), 743–771.

Ladd, H. F. (1992). Mimicking of Local Tax Burdens Among Neighboring Counties. *Public Finance Quarterly* 20, 450–467.

Lodh, T., Roy, P. & Roy, M. (2021). Intergenerational Occupational Mobility in India Across Social Groups. *Indian Economic Review*, Springer, vol. 56(2), pages 405-433, December.

Sacker, A., Kelly, Y., Iacovou, M., Cable, N., & Bartley. M.(2013). Breast Feeding and Intergenerational Social Mobility: What Are the Mechanisms? *Archives of Disease in Childhood*. 98(9):666-71.

Schürer, K. & Higgs, E. (2020). Integrated census microdata (I-CeM); 1851-1911. UK Data Service. [data collection]. Colchester, Essex: UK Data Archive [distributor], April 2014. SN: 7481.

Solon, G. (1999). Intergenerational Mobility in the Labor Market. In Orley C. Ashenfelter and David Card (editors). *Handbook of Labor Economics, Volume 3A*. Amsterdam: North-Holland

Solon, G. (2002). Cross-Country Differences in Intergenerational Earnings Mobility. *Journal of Economic Perspectives*. Vol. 16, pp. 59–66.

Spears, D. & Lamba, S. (2016). Effects of Early-Life Exposure to Sanitation on Childhood Cognitive Skills: Evidence from India's Total Sanitation Campaign. *Journal of Human Resources*, University of Wisconsin Press, vol. 51(2), pages 298-327.

Stand, D. W., & Rising, W. I. K. (2011). An overview of growing income inequalities in OECD countries: main findings. *Divided we stand: Why inequality keeps rising*, 21-45.

Szreter, S. (1988). The Importance of Social Intervention in Britain's Mortality Decline, *c*. 1850-1914: A Reinterpretation of the Role of Public Health. *Social History of Medicine*, 1, 1–37.

Thiede, B. C., Randell, H., & Gray, C. (2022). The Childhood Origins of Climate-Induced Mobility and Immobility. *Population and Development Review*, *48*(3), 767-793.

Van Leeuwen, M.H. & Maas, I. (2011). HISCLASS: A historical international social class scheme. Universitaire Pers Leuven.

van Leeuwen, M. H. D. (2020). Studying Long-Term Changes in the Economy and Society Using the HISCO Family of Occupational Measures. In J. H. Hamilton, A. Dixit, S. Edwards, & K. Judd (Eds.), *Oxford Research Encyclopedia of Economics and Finance.* Oxford University Press.

Victora C.G., Adair L., Fall C., Hallal P.C., Martorell R., Richter L. & Sachdev H.S. (2008). Maternal and Child Undernutrition: Consequences for Adult Health and Human Capital. *Lancet.*

Victora, C. G., de Onis, M., Hallal, P. C., Blössner, M., & Shrimpton, R. (2010). Worldwide timing of growth faltering: revisiting implications for interventions. *Pediatrics*, 125(3), e473–e480.

Ward, Z. (2022). Internal Migration, Education, and Intergenerational Mobility: Evidence from American History. *Journal of Human Resources*, University of Wisconsin Press, vol. 57(6), pages 1981-2011.

Wolf, J., Johnston, R. B., Ambelu, A., Arnold, B. F., Bain, R., Brauer, M., Brown, J., Caruso, B. A., Clasen, T., Colford, J. M., Jr, Mills, J. E., Evans, B., Freeman, M. C., Gordon, B., Kang, G., Lanata, C. F., Medlicott, K. O., Prüss-Ustün, A., Troeger, C., Boisson, S., … Cumming, O. (2023). Burden of Disease Attributable to Unsafe Drinking Water, Sanitation, and Hygiene in Domestic Settings: a Global Analysis for Selected Adverse Health Outcomes. *Lancet*.

Chapter 3: Social Mobility in the Nineteenth-Century England and Wales[1]

3.1 Introduction

The exploration of social mobility has long captivated the attention of scholars in economics and social sciences. Mobility is consequential because it affects how society is shaped or structured over time and how that structure is justified morally, culturally, and politically (Miles, 2022). Central to this inquiry is the examination of how individuals are positioned within a hierarchical social structure and the extent to which occupation, class, and earnings are transmitted across generations (intergenerational mobility) and throughout their lives (intragenerational mobility or career mobility). This study revisits the extent of social mobility in nineteenth-century England and Wales, a period marked by profound changes due to the Industrial Revolution.

Understanding social mobility during this transformative period is crucial. The Industrial Revolution fundamentally altered England's social fabric and economic landscape, affecting individuals' lives and the overall structure of society. A key question arises: *Did nineteenth-century England and Wales foster equal opportunities based on merit, or did it exacerbate inequality?*

On one hand, the Smilesian view presents Victorian Britain as "a land of boundless opportunity," asserting that no social barriers existed for those who adhered to the principles of self-help. On the other hand, Miles (1999) argues that equality of opportunity was absent, depicting a society where the middle and working classes remained distinct.

[1] I am grateful to Priyaranjan Jha, Vellore Arthi, and Meera Mahadevan for their invaluable guidance and encouragement. I am also thankful to Myra Mohnen for generously sharing linked historical census data. All errors are my own.

Existing literature on social mobility in this period has encountered significant challenges due to limited and unrepresentative data. Many studies have relied on data from marriage registries to estimate mobility in the nineteenth century (Miles, 1993, 1999; Mitch, 1993, 2005). Recent research leveraging census data (Long, 2013; Long & Ferrie, 2013; Pérez, 2019) created measures of occupational mobility. However, these studies rely on a 2% sample of the 1851 census dataset to estimate mobility, which limits their findings due to the small sample size. Ward (2023) highlights that unrepresentative samples can bias mobility estimates, showing how estimates can more than double when measurement errors are addressed. This raises an important question on the validity of England and Wales's prior mobility estimates and whether the observed trends were an artifact of the measurement error in the data.

To overcome this measurement error caused by data limitation, this study uses a new linked dataset (Costas-Fernández et al., 2020) created from the full 1851 and 1881 censuses of England and Wales. This dataset includes 196,953 father-son pairs for intragenerational mobility analysis and 489,866 pairs for intergenerational mobility analysis, providing a substantial and representative sample to measure social mobility.

With this comprehensive dataset, I estimate intergenerational mobility using absolute and relative measures. *Absolute mobility* or total mobility refers to the fraction of sons in a different occupational category than their father. I estimate that absolute mobility in the nineteenth century was larger than previously estimated in literature by Pérez (2019) and Long and Ferrie (2013), however, it was much smaller than the optimistic estimate by Long (2013). *Relative mobility,* as described by Berger et al. (2023), captures how the chance of entering a particular occupation differs for sons born to fathers from different occupational groups. Altham statistic (Altham, 1970; Altham & Ferrie, 2007) is used in the literature as a way to summarize relative mobility in a single statistic rather than computing several odds ratios. Using the Altham statistic, I estimate the relative mobility and find that this is slightly larger than estimated by Pérez (2019) but is statistically comparable to the estimate by Long and Ferrie (2013).

Next, I conduct international mobility comparisons with the revised mobility estimates for England and Wales. The findings suggest England and Wales had lower mobility than Argentina, US, Canada, and Sweden and a higher mobility than Norway. The ranking of these countries based on their mobility patterns shows that the New World's settler economies exhibited greater mobility than European countries. Pérez (2019) suggests that one factor that likely contributing to these differences is the higher availability of unexploited land in Argentina and the United States relative to Britain and Norway. Based on this factor, he also speculates that Canada might have a higher rate of social mobility than Britain.

In general, the mobility literature has paid less attention to intragenerational or career mobility (Long, 2013). This study finds that men showed career stability in the nineteenth century, with more than 50% remaining in occupations they were in 30 years ago. Career mobility was also less likely than intergenerational mobility. The new sample-based intragenerational mobility estimates are lower than Long's estimate (Long, 2013), suggesting that his surprisingly high estimates of mobility might have been an artifact of the smaller sample size of his data set.

This paper contributes to the existing literature which computes the extent of social mobility by utilizing the complete 1851 census dataset to estimate the extent of occupational mobility in nineteenth-century England and Wales. By addressing the measurement challenges highlighted by Ward (2023), it revises previous mobility estimates and offers new insights into the dynamics of social mobility during a critical period of economic and social transformation.

The rest of the paper is organized as follows. Section 3.2 reviews the existing literature on nineteenth-century social mobility in England and Wales, section 3.3 summarizes the data, section 3.4 discusses the estimates of intergenerational mobility, section 3.5 discusses intra-generational mobility estimates, and section 3.6 concludes.

3.2 Literature Review

In nineteenth-century England and Wales, life's prospects were considered to be largely determined by social position at birth (Koditschek, 2001). It was deemed as a rigid system where family background played an important role (Long & Ferrie, 2013), with the systems of feudal tradition, guild, and apprenticeship contributing to this rigidity.

The conventional assumption was that industrialization would make society more open, and lead to more mobility (Lipset & Zetterberg, 1959). David Landes in *The Unbound Prometheus* wrote 'A competitive industrial system… will increase social mobility, raising the gifted, ambitious, and lucky, and lowering the inept, lazy, ill-fortuned…Industrialization is, in short, a universal social solvent …' (Landes, 2003). However, Clark and Cummins (2013) find evidence that social mobility rates have historically been low in England and were unaffected by the

Industrial Revolution. Studying surname persistence amongst Oxbridge graduates and holders of top positions since the 13th century, they find that social status in Britain is more strongly inherited than height (Clark & Cummins, 2013; Clark, 2015).

Studies of historical social mobility are frequently limited by the accessibility of individual-level data containing indicators of social standing such as income. In the absence of reliable income data, historical mobility literature has largely relied upon occupation as a proxy for status. Miles (1993, 1999) and Mitch each used samples of marriage registrations from 1839 to 1914 to measure intergenerational occupational mobility. Using the information on occupation reported at the registration of marriage, both Miles and Mitch find that approximately 60-68% of the grooms were in the same occupation as their fathers at the time of marriage. However, Long (2013) criticizes the registration datasets that were used since they only comprised couples married in Anglican churches and thus were not representative. Additionally, the occupations of father and son were not measured at approximately the same ages and thus were taken from different points in their life cycles making the comparisons difficult.

To improve the estimates of mobility, Long (2013) derived new estimates of historical social mobility in England and Wales (inter- and intragenerational) using a large new dataset of fathers and sons linked across censuses from 1851–1881 and 1881–1901. He finds surprisingly high rates of social mobility, in favor of the opinion that British society had substantial opportunity across the socioeconomic spectrum. However, he uses a 2% sample from the census dataset for 1851 since the entire dataset was not available then, which limits his sample of linked fathers and sons to 4071 in 1851-1881.

Long and Ferrie (2013) and Pérez (2019) also provide estimates of intergenerational mobility for nineteenth-century England and Wales. Both papers utilize the 2% sample of the 1851 census, resulting in a small number of observations for the analysis but do not find Britain to be as surprisingly mobile as Long (2013).

3.3 Data

I obtain the father-son linked database from Costas-Fernández et al. (2020). They utilize the full-population censuses of England and Wales in 1851 and 1881 developed by the I-CeM project

(Schürer & Higgs, 2020), and link individuals across censuses.[2] The census contains information on geographic variables like house number, street name, and parish, registration district, and a range of sociodemographic information: age, gender, marital status, number of children, number of servants, and family structure. The only economic outcome available in this data is self-reported occupation. There are over 400 occupations such as physician, cook, stable keeper, cabinet maker, or farmer.

Historical settings largely use occupation to compute measures of intergenerational mobility (Boberg-Fazlic & Sharp, 2013; Long, 2013; Long & Ferrie, 2013; Olivetti & Paserman, 2015) due to two main reasons. First, historical data often does not have information on income making occupation the only economic outcome consistently available. Second, occupations withstand transitory shocks that incomes might not. Moreover, as emphasized by Long (2013), occupation and social class encompass several dimensions of an individual's life experience that can be linked to interpretations of social mobility. These facets include aspects like community prestige, workplace autonomy, place of employment, and more. Finally, measuring mobility using occupational data is also important for this study to compare results with those of the prior literature. Therefore, using the information on self-reported occupation from the census, I map the occupation's HISCO code to a corresponding HISCLASS category. The 4 broad HISCLASS categories are white collar, skilled/semiskilled, unskilled, and farmers. [3] Each class can be considered as a group of people with the same opportunities in life and the scheme conforms to the way economic and social historians have looked at social stratification from a historical perspective (van Leeuwen, 2020).

To measure the mobility between the father's mature occupation to the son's mature occupation (i.e., intergenerational mobility), I construct my sample of father-son pairs following Pérez (2019). The sample includes sons who were 16 years old or younger when they were observed living with their father during the initial census year. This is done to account for the possibility that individuals who continue to live with their father until relatively late might demonstrate distinct patterns of mobility compared to those who do not (Xie & Killewald, 2013).

[2] Costas-Fernández et al. (2020) rely on birth year, parish of birth, given name, and surname (time-invariant characteristics) to link individuals across censuses.

[3] Following Pérez (2019), I collapsed the HISCLASS scheme into four broad categories white-collar (HISCLASS 1-5), farmer (HISCLASS 8), skilled/semi-skilled (HISCLASS 6-7,9) and unskilled (HISCLASS 10-12)

Table 3.1: Summary Statistics

	Obs	Mean	Std Dev	Min	Max
Panel A: Intergenerational Mobility Sample					
Father (in 1851)					
Age	489,866	41.577	7.195	30	60
White Collar	489,866	0.102	0.302	0	1
Skilled/SemiSkilled	489,866	0.478	0.500	0	1
Farmer	489,866	0.096	0.294	0	1
Unskilled	489,866	0.325	0.468	0	1
Son (in 1881)					
Age	489,866	37.710	4.859	30	60
White Collar	489,866	0.168	0.374	0	1
Skilled/SemiSkilled	489,866	0.547	0.498	0	1
Farmer	489,866	0.066	0.248	0	1
Unskilled	489,866	0.219	0.413	0	1
Occ. Category Son ≠ Occ. Category father	489,866	0.456	0.498	0	1
Panel B: Intragenerational Mobility Sample					
Males (in 1851)					
Age	196,953	22.067	4.419	17	35
White Collar	196,953	0.072	0.258	0	1
Skilled/SemiSkilled	196,953	0.514	0.500	0	1
Farmer	196,953	0.092	0.289	0	1
Unskilled	196,953	0.322	0.467	0	1
Males (in 1881)					
Age (in 1881)	196,953	50.712	7.666	1	98
White Collar	196,953	0.153	0.360	0	1
Skilled/SemiSkilled	196,953	0.493	0.500	0	1
Farmer	196,953	0.108	0.311	0	1
Unskilled	196,953	0.246	0.431	0	1
Occ. Category 1851 ≠ Occ. Category 1881	196,953	0.385	0.487	0	1

Note: Panel A summarizes the sample used to obtain intergenerational mobility estimates. Fathers and sons are 30-60 years old when their occupations are measured in 1851 and 1881 respectively. Panel B summarizes the sample used to obtain intragenerational mobility estimates. Males are 17 to 35 years old when their occupations are first measured in 1851 and then again in 1881. Measures of mobility are based on the four-category HISCLASS.

Additionally, I restricted the samples to father-son pairs in which both the father and the son were between the ages of 30 and 60 when their occupations were recorded in the census. This is done to account for the fact that occupations measured either too early or too late in the life cycle might be a noisy measure of long-run economic status (Solon, 1992). The resulting sample consists of 489,866 father-son pairs.

Summary statistics from Table 3.1 Panel A suggest that the average age of fathers (41.577 yo) and the average age of the sons in the sample (37.710 yo) were similar. The fraction of fathers engaged in white-collar jobs was 10.2%, and this fraction increased to 16.8% for the sons. The fraction of fathers engaged in skilled/semiskilled jobs was 47.8% for fathers. This fraction also increased to 54.7% for the sons, whereas the fraction of men engaged in farming and unskilled occupations declined from the father's generation to the son's generation.

Economists are also interested in another aspect of mobility, which involves tracing individuals' experiences from their early careers to their subsequent mature occupations. Therefore, following Long (2013), to measure the career mobility of an individual or intragenerational mobility, older individuals in 1851 (17 to 35 years old) are traced after three decades in 1881. This gives a sample of 196,953 men between the ages of 17 to 35 in 1851. Table 3.1 Panel B summarizes this sample. The average age of males when their occupation is first measured in 1851 is 22.067 years old. Thirty years later, in 1881, the average age of these men is reported to be 50.712 years old. There appears to be a reduction in the fraction of men engaged in unskilled jobs over the lifetime and an increase in the fraction in white-collar jobs. Only 38.5% moved to a different occupational category over three decades.

3.4 Estimating Intergenerational Mobility

Table 3.2 reports the pattern of intergenerational occupational mobility in the linked father-son pairs in the form of a transition matrix. It provides the number and percentage of sons attaining each occupational class in 1881, by father's occupational class in 1851 (origin class).

One way to measure total mobility using this matrix is to identify the fraction of sons who were in a different occupational category than their fathers. This can be done by adding the frequency of the non-diagonal cells and dividing that by the total number of observations in the sample. Results suggest that 45.60% of the sons in this sample worked in a different occupation than their fathers. This estimate is higher than the 44% estimated by Pérez (2019) for Britain in the late nineteenth century and is also higher than 42.6% as computed by Long and Ferrie (2013). However, this is much lower than the rate of 50.1% mobility that was put forth by Long (2013). It seems, therefore, that his estimates of the surprising mobility in Victorian Britain could be an artifact of the small dataset he used (~12,000 observations).

Some well-known facts about the occupational structure of England and Wales also emerge from Table 3.2. By 1851, due to the Industrial Revolution, few fathers were employed in farming (Pérez, 2019). The results suggest that farmers were the smallest occupational group. Most of the fathers were in skilled or semi-skilled jobs in 1851 and 68% of their sons remained in these jobs in 1881, thirty years later. Sons of white-collar and unskilled fathers were most likely to be in skilled/semiskilled jobs, in part because these jobs dominated the labor market at the time (Long & Ferrie, 2018). This result differs from Pérez (2019), who finds that the sons are most likely to

Table 3.2: Intergenerational Mobility – Transition Matrix

	Father's Occupation				
Son's Occupation	White Collar	Skilled/SemiSkilled	Farmer	Unskilled	Row Sum
White Collar	19,168	38,050	7,395	17,898	82,511
	{38.46}	{16.25}	{15.76}	{11.26}	{16.84}
Skilled/SemiSkilled	22,914	162,619	13,856	68,695	268,084
	{45.98}	{69.46}	{29.53}	{43.21}	{54.73}
Farmer	1,818	6,233	18,197	5,902	32,150
	{3.65}	{2.66}	{38.78}	{3.71}	{6.56}
Unskilled	5,937	27,231	7,479	66,474	107,121
	{11.91}	{11.63}	{15.94}	{41.82}	{21.87}
Total	49,837	234,133	46,927	158,969	489,866

Note: The entries (in curly brackets) represent the number (the share) of sons working in a row occupation among sons whose fathers work in column occupation based on the four-category HISCLASS classification. The father's occupation is measured in 1851 and the son's occupation is measured in 1881.

follow their father's occupation, regardless of the category. However, comparing mobility rates between different tables isn't sufficient for determining which society is more mobile because raw rates are influenced by marginal frequencies. Therefore, much of the literature on social mobility from the nineteenth and twentieth centuries uses a different method to compare transition matrices.

To systematically compare mobility between tables, literature has relied upon the Altham Statistic, which is a single metric that summarizes the mobility difference and can be tested for statistical significance (Altham, 1970; Long & Ferrie, 2013; Antonie et al., 2022; Berger et al., 2023). The measure treats movements across categories equally, without considering the starting and ending categories. This means there's no assumed hierarchy or distance between occupational groups. This is particularly valuable since in the absence of individual earnings data, there is no commonly agreed-upon ranking of these occupational groups in literature.

To compare two tables P and Q with r rows and s columns, the general form of this statistic d(P, Q) is as follows:

$$d(P,Q) = \left[\sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{m=1}^{r}\sum_{l=1}^{s}\left[\log\left(\frac{p_{ij}p_{lm}}{p_{im}p_{lj}}\right) - \log\left(\frac{q_{ij}q_{lm}}{q_{im}q_{lj}}\right)\right]^2\right]^{1/2} \quad \ldots\ldots\ldots. (1)$$

where *i* and *l* index fathers' occupations and *j* and *m* index sons' occupations. The statistic measures how far the association between rows and columns in table P departs from the association between rows and columns in table Q. It aggregates the log differences between the cross-product ratios in matrix P and matrix Q. As discussed by Long and Ferrie (2013), a likelihood-ratio χ^2 statistic G^2 with (r-1)(s-1) degrees of freedom can be used to test whether the matrix Θ with elements $\theta_{ij} = \log\left(\frac{p_{ij}}{q_{ij}}\right)$ is independent. Rejecting the hypothesis that Θ is independent, is equivalent to accepting that $d(P,Q) \neq 0$. This implies that the degree of association between rows and columns differs across P and Q. Note that the statistic does not tell us which table of the two tables exhibits larger mobility. To determine that, we calculate additional statistics, $d(P,I)$ and $d(Q,I)$. This statistic measures the difference in row-column associations relative to a matrix I, which has independent rows and columns. Hence, if we observe that $d(P,Q) \neq 0$ and that $d(P,I)$ < $d(Q,I)$, we can conclude that *P* shows greater mobility than *Q*.

Table 3.3: Intergenerational Mobility Comparison

	M	$d(P,I)$	$d(Q,I)$	$d(P,Q)$
Panel A: Comparing to Previous England & Wales Estimates				
Full Census, 1851-1881 ***(P)***	45.6	20.11***		4.31***
vs. Pérez (2019), 1851-1881 ***(Q)***	44		20.80***	
Full Census, 1851-1881 ***(P)***	45.6	20.11***		3.50
vs. Long & Ferrie (2013), 1851-1881 ***(Q)***	42.6		22.70***	
Panel B: Comparing to Estimates of Other Countries				
England and Wales, 1851-1881 ***(P)***	45.6	20.11***		11.21***
vs Argentina, 1869-1895 ***(Q)***	54.7		13.45***	
England and Wales, 1851-1881 ***(P)***	45.6	20.11***		13.52***
vs Norway, 1865-1900 ***(Q)***	44.5		25.94***	
England and Wales, 1851-1881 ***(P)***	45.6	20.11***		9.16***
vs United States, 1850-1880 ***(Q)***	45.5		14.67***	
England and Wales, 1851-1881 ***(P)***	45.6	20.11***		9.12***
vs Canada, 1871-1901 ***(Q)***	50.7		16.00***	
England and Wales, 1851-1881 ***(P)***	45.6	20.11***		9.04***
vs Sweden, 1880-1910 ***(Q)***	53.5		17.93***	

Notes: Panel A- Full Census estimates for England and Wales are calculated in this paper. Other previous estimates are calculated from the transition matrices reported in Pérez (2019) and Long and Ferrie (2013). Panel B – England and Wales estimates are calculated in this paper. Estimates for Argentina, Norway, and the United States are calculated from the transition matrices reported in Pérez (2019). Estimates for Canada are calculated from the transition matrix in Antonie et al. (2022). Estimates for Sweden are calculated from the transition matrix in Berger et al. (2023). ∗p<0.1; ∗∗p<0.05; ∗∗∗p<0.01.

Table 3.3 presents mobility measures based on Altham Statistics. In Panel A, I report estimates from prior studies on England and Wales (which used a 2% sample of the 1851 census data to measure mobility).[4] The first measure M is the percentage of sons in a different occupational

[4] Long (2013) cannot be included in Altham Statistics analysis because it applies the W.A. Armstrong's classification system, which is based on the Registrar General's 1921 and 1951 classification schemes (Armstrong & Wrigley, 1972). This classifies each individual into one of five categories, rather than four based on the HISCLASS classification used

category than their fathers (total mobility). In column 2, $d(P, I)$ reports the Altham statistics comparing the transition matrix of England and Wales (Table 3.2) to a transition matrix representing full independence. Likewise, $d(Q, I)$ gives the corresponding statistic for comparing matrix Q with a matrix of full independence. The last column reports $d(P, Q)$ which is the relevant Altham statistic to understand if matrices P and Q differ statistically from each other. The full census Altham statistics suggest that there was a bit more relative mobility than Pérez (2019) finds in the case of England and Wales. Compared to Long and Ferrie (2013), however, the statistic $d(P, Q)$ is not statistically significant.

Next, I turn to international comparisons in intergenerational mobility. Table 3.3 Panel B reports the measures for several countries. The measure M suggests that the level of mobility was the highest in Argentina, with approximately 54.7% of sons in a different occupational category than their fathers, compared to 53.5% in Sweden, 50.7% in Canada, 45.6% in England and Wales, 45.5% in the United States, and 44.5% in Norway.

However, as discussed before, a limitation of this measure is that it can't tell apart the differences in how generations are linked and the differences in occupational structures across countries. Therefore, I turn to the Altham Statistics for comparing the intergenerational mobility across these countries. All estimates of $d(P, I)$ and $d(Q, I)$ are significant suggesting that in none of these countries do I see independence between fathers' and sons' occupations. The statistic $d(P, Q)$ suggests that the distances from England and Wales's mobility patterns are similar for Sweden, Canada, and the United States. However, England and Wales is further away from Argentina's mobility pattern and the furthest from Norway's.

In line with previous literature, results indicate that mobility is higher in Argentina, Canada, United States, and Sweden compared to the new sample estimates of England and Wales. On the other hand, mobility is lower in Norway compared to England and Wales. The ranking of countries in their mobility pattern confirms that the settler economies of the New World were more mobile than Europe (Pérez, 2019; Antonie et al., 2022; Berger et al., 2023).

---

in the paper. This difference in the number of occupational categories implies that the Altham statistic is not applicable.

3.5 Estimating Intragenerational Mobility

Table 3.4 reports the pattern of absolute intragenerational occupational mobility or career mobility in the form of a transition matrix. It provides the number and percentage of males attaining each occupational class in 1881, by class of their occupational class in 1851.

The elements off the main diagonal represent total mobility. Results suggest that the rate of career mobility or intragenerational mobility was 38.46%, lower than the estimated intergenerational mobility for England and Wales during the same period (Section 3.5). This new sample estimate is also lower than the prior intragenerational mobility estimates of 43.6 % by Long (2013) (as well as 42% mobility from the first job measured by Goldethorpe, based on calculations in Long, 2013). This evidence suggests that the total career mobility perhaps was not "reasonably high". Table 3.4 shows that more than 50% of men remain in occupations they were in 30 years

Table 3.4: Intragenerational Mobility - Transition Matrix

	Own Occupation, 1851				
Own Occupation, 1881	White Collar	Skilled/SemiSkilled	Farmer	Unskilled	Row Sum
White Collar	7,650	14,954	1,718	5,838	30,160
	{54.03}	{14.76}	{9.48}	{9.22}	{15.31}
Skilled/SemiSkilled	4,243	70,319	3,171	19,305	97,038
	{29.97}	{69.40}	{17.50}	{30.47}	{49.27}
Farmer	728	5,068	10,280	5,260	21,336
	{5.14}	{5.00}	{56.74}	{8.30}	{10.83}
Unskilled	1,538	10,984	2,948	32,949	48,419
	{10.86}	{10.84}	{16.27}	{52.01}	{24.58}
Total	14,159	101,325	18,117	63,352	196,953

Note: The entries (in curly brackets) represent the number (the share) of males working in a row occupation in 1881 among males who worked in column occupation in 1851, based on the four-category HISCLASS classification.

ago. For instance, 52% of men in unskilled occupations in 1851 remained in this category, and only a meager 9% were able to become white-collar workers in 1881. A similar situation exists for men who started as farmers in 1851. 56.74% of men who were farmers in 1851 persist in the same category, while only 9.22% become white-collared. This stands in contrast to the Smilesian view of boundless opportunities in England and Wales at the time and suggests that it was difficult to move into white-collar jobs from manual backgrounds in Victorian Britain.[5]

3.6 Conclusion

Utilizing a newly constructed linked dataset from the full-count censuses of England and Wales from 1851 and 1881, this paper contributes to the literature measuring historical social mobility by accounting for the measurement issue created by the unrepresentative samples in prior studies, often limited by data availability, as emphasized by Ward (2023).

The paper revises inter- and intragenerational occupational mobility estimates for England and Wales. New sample estimates of absolute intergenerational mobility suggest that 45.60% of the sons worked in a different occupational category than their fathers. This estimate of absolute mobility was marginally larger than previous estimates by Long and Ferrie (2013) and Pérez (2019) but is much smaller than Long (2013) who estimated total mobility to be optimistically at 50.1%. The intergenerational occupation transition matrix suggests that sons of white-collar and unskilled fathers were most likely to be in skilled/semiskilled jobs. This differs from Pérez (2019), who finds that the sons are most likely to follow their father's occupational footsteps, regardless of the origin category (i.e. there was a strong generational persistence regardless of the father's occupational category).

Using the measure of Altham statistics, I find that the relative intergenerational mobility estimates using the full census datasets are slightly larger mobility than what Pérez (2019) estimated but are statistically similar to the mobility rate estimated by Long and Ferrie (2013). Next, I compare the intergenerational mobility estimates of England and Wales with those of

[5] As mentioned in Section 3.5, a further comparison in intragenerational mobility using Altham statistics comparing Long (2013) with the estimates of this study is not possible due to the different nature of occupational classification schemes adopted in Long (2013).

Argentina, US, Norway, Sweden, and Canada using the Altham Statistics. In line with previous literature, these comparisons suggest that mobility in England was low, and ranked towards the bottom among these countries, confirming the presence of a distinct mobility pattern between the New and the Old-World economies.

Few studies have estimated career mobility for men in England and Wales. Using the linked full-count census dataset, I find that absolute career mobility is lower than Long (2013) estimated. About 50-70% of men in each occupational class remained in their prior occupational classes, even after three decades. Overall, only 38.46% of men displayed career mobility, which is less than the estimated total intergenerational mobility.

These results are at odds with the extent of 'surprising mobility' estimated by Long (2013) and overall, do not support the Smilesian view of an open or mobile English society.

3.7 References

Altham, P. M. E. (1970). The Measurement of Association of Rows and Columns for an r× s Contingency Table. *Journal of the Royal Statistical Society: Series B (Methodological)*, *32*(1), 63-73.

Altham, P. M. E., & Ferrie, J. P. (2007). Comparing Contingency Tables Tools for Analyzing Data from Two Groups Cross-Classified by Two Characteristics. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, *40*(1), 3–16. https://doi.org/10.3200/HMTS.40.1.3-16

Antonie, L., Inwood, K., Minns, C., & Summerfield, F. (2022). Intergenerational Mobility in a Mid-Atlantic Economy: Canada, 1871–1901. *The Journal of Economic History*, *82*(4), 1003-1029.

Armstrong, W. A., & Wrigley, E. A. (1972). The use of information about occupation. In *Nineteenth-Century Society* (pp. 191–310). Cambridge: Cambridge University Press.

Berger, T., Engzell, P., Eriksson, B., & Molinder, J. (2023). Social mobility in Sweden before the welfare state. *The Journal of Economic History*, *83*(2), 431-463.

Boberg-Fazlic, N., & Sharp, P. (2013). *North and south: social mobility and welfare spending in preindustrial England* (No. 37). EHES Working Papers in Economic History.

Clark, G. (2015). *The Son Also Rises: Surnames and the History of Social Mobility: Surnames and the History of Social Mobility*. Princeton University Press.

Clark, G., & Cummins, N. (2013). Surnames and social mobility: England 1230-2012 (Working Paper No. 181). *Department of Economic History: The London School of Economics*.

Costas-Fernández, J., Guerra, J. A., & Mohnen, M. (2020). *Train to opportunity: The Effect of infrastructure on intergenerational mobility.* Documento CEDE, (48).

Koditschek, T. (2001). [Review of *Social Mobility in Nineteenth- and Early Twentieth-Century England*, by A. Miles]. *The Journal of Interdisciplinary History*, *31*(4), 630–632. http://www.jstor.org/stable/206873

Landes, D. S. (2003). *The unbound Prometheus: technological change and industrial development in Western Europe from 1750 to the present*. Cambridge University Press.

Lipset, S. M., & Zetterberg, H. L. (1959). *Social mobility in industrial societies* (pp. 11-75). Berkeley: University of California Press.

Long, J. (2013). The surprising social mobility of Victorian Britain. *European Review of Economic History*, *17*(1), 1-23.

Long, J., & Ferrie, J. (2013). Intergenerational occupational mobility in Great Britain and the United States since 1850. *American Economic Review*, *103*(4), 1109-1137.

Long, J., & Ferrie, J. (2018). Grandfathers matter (ed): occupational mobility across three generations in the US and Britain, 1850–1911. *The Economic Journal*, *128*(612), F422-F445.

Miles, A.(1993). "How Open was Nineteenth-Century British Society? Social Mobility and Equality of Opportunity, 1839-1914". In A. Miles and D. Vincent (eds), *Building European Society: Occupational and Social Mobility in Europe, 1840-1940*. Manchester University Press.

Miles, A. (1999). *Social Mobility in Nineteenth and Early Twentieth Century England*. Macmillan Publishers Ltd.

Miles, A. (2022). Social Mobility. In N. Robertson, J. Singleton, & A. Taylor (Eds.), *20th Century Britain: Economic, Social and Cultural Change* (3rd ed., pp. 216-229). Routledge. https://doi.org/10.4324/9781003037118-16

Mitch, D. (1993). "'Inequalities which every one may remove: occupational recruitment, endogamy, and the homogeneity of social origins in Victorian England.'" In A. Miles and D. Vincent (eds), *Building European Society: Occupational Change and Social Mobility in Europe, 1840 – 1940*. Manchester: Manchester University Press

Mitch, D. (2005). Literacy and occupational mobility in rural versus urban Victorian England: Evidence from the linked marriage register and census records for Birmingham and Norfolk, 1851 and 1881. *Historical Methods: A Journal of Quantitative and Interdisciplinary History,* 38(1), 26-38.

Olivetti, C., & Paserman, M. D. (2015). In the name of the son (and the daughter): Intergenerational mobility in the United States, 1850–1940. *American Economic Review*, *105*(8), 2695-2724.

Pérez, S. (2019). Intergenerational occupational mobility across three continents. *The Journal of Economic History*, *79*(2), 383-416.

Schürer, K., & Higgs, E. (2020). Integrated Census Microdata (I-CeM); 1851-1911. UK Data Service. [data collection]. DOI: 10.5255/UKDA-SN-7481-2 . Colchester, Essex: UK Data Archive [distributor], April 2014. SN: 7481.

Solon, G. (1992). Intergenerational income mobility in the United States. *The American Economic Review*, 393-408.

van Leeuwen, M. H. (2020). Studying long-term changes in the economy and society using the HISCO family of occupational measures. In *Oxford Research Encyclopedia of Economics and Finance*.

Ward, Z. (2023). Intergenerational mobility in American history: Accounting for race and measurement error. *American Economic Review*, 113(12), 3213-3248.

Xie, Y., & Killewald, A. (2013). Intergenerational occupational mobility in Great Britain and the United States since 1850: Comment. *American Economic Review*, 103(5), 2003-2020.

Appendix A: Appendix to Chapter 1

Table A.1: Classification of Districts into High and Low-conflict Zones

District	Number of Conflict Events	Zone
Kalutara	0	Low-conflict (Control Group)
Ratnapura	0	
Nuwara Eliya	1	
Hambantota	1	
Matara	1	
Gampaha	2	
Badulla	2	
Galle	2	
Matale	3	
Kurunegala	3	
Kegalle	4	
Kandy	5	
Puttalam	17	
Monaragala	17	
Anuradhapura	38	
Polonnaruwa	65	
Colombo	71	
Ampara	122	
Trincomalee	259	High-conflict (Treatment Group)
Kilinochchi	311	
Batticaloa	363	
Vavuniya	442	
Mannar	455	
Jaffna	559	
Mullaitivu	806	
Total	3549	

Note: The number of conflict events are calculated by district for years 1992–2009 using the UCDP dataset. This table classifies districts as high (treatment) and low-conflict zones (control) based on being above or below the average number of conflict events (141.96).

Table A.2: Distribution of Ethnic Groups by District in 1981

District	All	Sinhalese	Sri Lankan Tamils	Others
Colombo	1,699,241	1,318,835	170,590	209,816
Gampaha	1,390,862	1,279,512	48,182	63,168
Kalutara	829,704	723,483	9,744	96,477
Kandy	1,048,317	778,801	52,791	216,725
Matale	357,354	285,354	20,579	51,421
Nuwara Eliya	603,577	254,375	76,449	272,753
Galle	814,531	769,343	7,271	37,917
Matara	643,786	608,516	4,683	30,587
Hambantota	424,344	412,055	2,500	9,789
Jaffna	394,512	3,163	375,433	15,916
Kilinochchi	436,040	3,496	414,952	17,592
Mannar	106,235	8,683	54,474	43,078
Vavuniya	95,428	15,794	54,179	25,455
Mullaltivu	77,189	3,992	58,209	14,988
Batticaloa	330,333	11,255	233,713	85,365
Trincomalee	255,948	85,503	87,760	82,685
Kurunegala	1,211,801	1,125,912	14,920	70,969
Puttalam	492,533	407,067	32,282	53,184
Anuradhapura	587,929	535,834	8,026	44,069
Polonnaruwa	261,563	238,965	5,267	17,331
Badulla	640,952	443,024	37,520	160,408
Monaragala	273,570	253,572	5,346	14,652
Ratnapura	797,087	677,510	19,094	100,483
Kegalle	684,944	588,581	15,074	81,289
Total	14,846,750	10,979,568	1,886,864	1,980,318

Note: This table provides a snapshot of the ethnic group population across the districts of Sri Lanka. It has been created using the data from the Census of Population and Housing Report (1981).

Table A.3: Summary Statistics of Vital Statistics Data

Variable	no. of obs	mean	std	min	max
A. Sinhalese					
Live Births	800	10339.82	10010.64	0	50265
Population	800	477723.2	457954.9	1461.38	2027287
Crude Birth Rate	800	20.58	10.73	0	66.77
B. Sri Lankan Tamil					
Live Births	800	1664.46	2271.3	0	14102
Population	800	72974.71	103768.7	1227.13	553787.6
Crude Birth Rate	800	24	12.1	0	102.69
C. Indian Tamil					
Live Births	800	950.58	1896.42	0	9390
Population	800	37561.45	71098.86	152.94	371560.4
Crude Birth Rate	800	17.92	15.93	0	159.77
D. Moors					
Live Births	800	1456.48	1672.21	0	7574
Population	800	50939.47	55437.06	791.99	271064.3
Crude Birth Rate	800	27.27	28.92	0	783.23
E. Others					
Live Births	800	87.78	209.73	0	1389
Population	800	4577.92	10020.36	32.93	51524
Crude Birth Rate	800	25.54	57.39	0	702.61

Note: The summary statistics is obtained from district-ethnicity data from prewar (1967-1980) as well as postwar years (1992-2009).

Table A.4: Impact of Conflict on Fertility: Average Effect on Population, Robustness

	Dependent Variable: Crude Birth Rate			
	alternative treatment	compares top & bottom conflict quartile districts	drops top 5% conflict districts	drops top 10% conflict districts
	(1)	(2)	(3)	(4)
Independent Variable				
treated x post	-4.33***	-2.24*	-6.494***	-6.120***
	(0.848)	(1.326)	(0.96)	(1.016)
N	2,400	1,152	2,208	2,112
R^2	0.32	0.318	0.305	0.303

Note: Robust standard errors are reported in the parenthesis. Estimates are weighted by district population. Column 1 assigns alternative treatment based on Li et al. (2019), assigning 8 districts with a large LTTE influence in the north and east as the high-conflict zone and the rest as the low-conflict zone. Column 2 assigns treatment and control districts to top and bottom quartile districts (6 in each group) ranked by the mean level of conflict per capita per year. Column 3 excludes the top 5% of the conflict districts – Mullaitivu and Jaffna. Column 4 excludes the top 10% of the conflict districts – Mullaitivu, Jaffna and Mannar. ∗p<0.1, ∗∗p<0.05, ∗∗∗p<0.01.

Table A.5: Impact of Conflict on Fertility- Heterogeneity by Ethnic Group, Robustness

	Dependent Variable: Crude Birth Rate			
	alternative treatment	compares top & bottom conflict quartile districts	drops top 5% conflict districts	drops top 10% conflict districts
	(1)	(2)	(3)	(4)
Independent Variable				
treated x post	-5.09***	-10.45***	-10.160***	-10.136***
	(1.038)	(1.44)	(0.879)	(0.904)
treated x post x Sri Lankan Tamil	3.23*	8.86***	7.033**	7.074***
	(1.767)	(3.205)	(1.648)	(1.709)
treated x post x Others	2.14	8.55***	6.558***	8.171***
	(1.527)	(2.285)	(1.504)	(1.561)
N	2400	1152	2208	2112
R^2	0.693	0.741	0.721	0.721

Note: Robust standard errors are reported in the parenthesis. The regression includes district, year, ethnicity, district x ethnic group, and year x ethnic group fixed effects. Estimates are population weighted. Column 1 assigns alternative treatment based on Li et al. (2019), assigning 8 districts with the large LTTE influence in the north and east as the high-conflict zone and the rest as the low-conflict zone. Column 2 assigns treatment and control districts to top and bottom quartile districts (6 in each group) ranked by the mean level of conflict per capita per year. Column 3 excludes the top 5% of the conflict districts – Mullaitivu and Jaffna. Column 4 excludes the top 10% of the conflict districts – Mullaitivu, Jaffna and Mannar. ∗p<0.1, ∗∗p<0.05, ∗∗∗p<0.01.

Table A.6: Impact of Conflict on Fertility by Ethnicity, Robustness

	Dependent Variable: Crude Birth Rate					
	main treatment	main treatment (w/ pop. wts)	alternative treatment	compares top & bottom conflict quartile districts	drops top 5% conflict districts	drops top 10% conflict districts
	(1)	(2)	(3)	(4)	(5)	(6)
Independent Variable						
treated x post	-8.25***	-10.66***	-5.09***	-10.45***	-10.160***	-10.136***
	(1.159)	(0.864)	(1.038)	(1.440)	(0.879)	(0.904)
treated x post x Sri Lankan Tamil	4.07**	11.41***	3.23*	8.86***	7.033***	7.074***
	(1.879)	(1.657)	(1.770)	(3.205)	(1.648)	(1.709)
treated x post x Indian Tamil	1.03	8.23***	3.01*	9.33***	7.657***	8.671***
	(1.886)	(1.641)	(1.723)	(1.779)	(1.871)	(2.055)
treated x post x Moors	1.14	3.36*	-2.04	5.09*	2.563	4.212**
	(6.269)	(1.787)	(1.556)	(2.887)	(1.585)	(1.640)
treated x post x Others	-45.29***	-5.04	-10.18***	-1.29	-5.416*	0.258
	(12.714)	(3.071)	(2.747)	(4.531)	(3.148)	(2.547)
N	4000	4000	4000	1920	3680	3520
R^2	0.335	0.669	0.668	0.694	0.713	0.717

Note: Robust standard errors are reported in the parenthesis. The regression includes district, year, ethnicity, district x ethnic group and year x ethnic group fixed effects. Column 1 and 2 assigns treatment based on the classification in Appendix Table A.1. Column 1 does not include population weights whereas column 2 does. Estimates are population weighted in columns 3-6. Column 3 assigns alternative treatment based on Li et al. (2019), assigning 8 districts with the large LTTE influence in the north and east as the high-conflict zone and the rest as the low-conflict zone. Column 4 assigns treatment and control districts to top and bottom quartile districts (6 in each group) ranked by the mean level of conflict per capita per year. Column 5 excludes the top 5% of the conflict districts – Mullaitivu and Jaffna. Column 6 excludes the top 10% of the conflict districts – Mullaitivu, Jaffna and Mannar. ∗p<0.1, ∗∗p<0.05, ∗∗∗p<0.01.

Table A.7: Impact of Conflict on Fertility by Ethnicity- Conflict events per 1000 inhabitants

Dependent Variable: Crude Birth Rate		
	continuous treatment	placebo
	(1)	(2)
Independent Variable		
conflict intensity	-7.831***	3.368
	(2.257)	(10.6)
conflict intensity x Sri Lankan Tamil	2.149	3.718
	(4.199)	(15.530)
conflict intensity x Others	2.4	14.29
	(3.330)	(12.951)
N	2,250	975
R^2	0.6883	0.75

Note: Robust standard errors are reported in the parenthesis. Estimates are population weighted. The regression includes district, year, ethnicity, district x ethnic group, and year x ethnic group fixed effects. Estimates are population weighted. Conflict intensity is the continuous variable: conflict events per 1000 inhabitants in the district in year t-1. The mean, minimum, maximum, and standard deviation of the conflict intensity variable are the following (0.025, 0, 2.746, 0.159). The placebo test in column 2 shifts the conflict intensity from the actual years of occurrence in 1992-2005 to the pre-war years of 1967-1980 and finds no effect on fertility. ∗p<0.1, ∗∗p<0.05, ∗∗∗p<0.01.

Table A.8: Impact of Conflict on Fertility- Event Study

	Dependent Variable: Crude Birth Rate			
	Overall	Sinhalese	Heterogeneity for Sri Lankan Tamils	Total Effect on Sri Lankan Tamils
Independent Variable:	$treated_{d,\,i}$	$treated_{d,\,i}$	$treated_{d,\,j}$ x sri lankan tamil	
	(1)	(2)	(3)	(4) = (2)+(3)
Year:				
1967	-8.18***	-6.1	1.99	-4.11
	(3.058)	(3.920)	(3.389)	(3.428)
1968	-7.39**	-9.14**	4.64	-4.5
	(3.039)	(3.879)	(3.431)	(3.546)
1969	-6.18**	-9.46***	5.97**	-3.49
	(2.860)	(3.518)	(2.839)	(3.348)
1970	-3.83	-6.63*	4.7	-1.92
	(2.910)	(3.696)	(2.987)	(3.253)
1971	-4.03	-6.89*	3.89	-3
	(2.882)	(3.726)	(3.079)	(3.317)
1972	-3.02	-2.88	0.9	-1.98
	(3.132)	(4.021)	(3.348)	(3.244)
1973	-1.13	-2.42	0.28	-2.15
	(3.103)	(3.592)	(2.986)	(3.457)
1974	-1.24	-1.66	0.92	-0.74
	(3.126)	(3.452)	(2.746)	(3.454)
1975	-1.08	-1.2	1.47	0.27
	(3.462)	(3.736)	(3.192)	(3.634)
1976	-1.13	-0.89	1.65	0.76
	(3.515)	(3.464)	(3.195)	(3.970)
1977	-2.93	-2.4	1.21	-1.19
	(3.554)	(3.723)	(3.280)	(3.906)
1978	-1.96	-2.22	2.74	0.53
	(3.753)	(3.511)	(3.812)	(4.647)
1979	-0.2	-1.43	0.46	-0.97
	(3.267)	(3.949)	(3.229)	(4.112)
1992	-7.27*	-17.76***	25.28***	7.51
	(4.092)	(3.061)	(4.587)	(5.155)
1993	-7.06*	-14.70***	21.25***	6.55
	(3.690)	(3.224)	(3.694)	(4.299)
1994	-5.89	-13.23***	21.37***	8.14*

	(3.831)	(3.483)	(3.895)	(4.240)
1995	-5.47	-13.45***	19.77***	6.33*
	(3.411)	(3.269)	(2.533)	(3.263)
1996	-7.60**	-14.76***	21.72***	6.96*
	(3.599)	(3.316)	(3.326)	(3.925)
1997	-4.97	-13.59***	19.08***	5.49
	(5.778)	(3.521)	(5.078)	(5.337)
1998	-6.48*	-12.29***	18.43***	6.13*
	(3.406)	(3.356)	(2.768)	(3.381)
1999	-5.81	-14.61***	22.45***	7.84**
	(3.656)	(3.210)	(2.648)	(3.480)
2000	-6.62*	-13.13***	20.14***	7.01**
	(3.810)	(3.515)	(2.689)	(3.194)
2001	-6.60*	-13.96***	20.13***	6.17*
	(3.533)	(3.436)	(2.998)	(3.559)
2002	-5.49*	-12.06***	18.89***	6.83**
	(3.331)	(3.362)	(2.596)	(3.353)
2003	-4.1	-11.17***	21.75***	10.57***
	(3.588)	(3.471)	(3.147)	(3.659)
2004	-7.22**	-12.40***	16.16***	3.75
	(3.369)	(3.521)	(3.577)	(4.043)
2005	-7.36**	-13.57***	17.67***	4.11
	(3.578)	(3.567)	(3.294)	(3.745)
2006	-6.80*	-14.35***	17.39***	3.04
	(3.502)	(3.401)	(2.972)	(3.567)
2007	-7.79**	-12.07***	15.73***	3.65
	(3.553)	(3.978)	(3.748)	(3.695)
2008	-9.48***	-12.98***	9.36**	-3.62
	(3.300)	(3.710)	(3.940)	(4.145)
2009	-9.46***	-12.10***	10.45***	-1.65
	(3.273)	(3.757)	(3.733)	(3.847)
N	2400		2400	
R^2	0.322		0.64	

Note: Event-study estimates in column 1 are based on overall equation for Sri Lanka from equation 4’ in footnote 14. Robust standard errors are reported in the parenthesis. Estimates in columns 2-4 are based on equation 4. Column 2 reports event-study estimates for Sinhalese. Column 3 reports the heterogeneity coefficient for Sri Lankan Tamils. Column 4 reports the overall effect for Sri Lankan Tamils (adding estimates from columns 2 and 3, with robust standard errors calculated using the Delta method). Treatment is assigned based on classification in Appendix Table A.1. ∗p<0.1, ∗∗p<0.05, ∗∗∗p<0.01.

Table A.9: Impact of Conflict on Fertility, Robustness

Dependent Variable: Crude Birth Rate				
	overall effect		heterogeneity by ethnicity	
	(1)	(2)	(3)	(4)
Independent Variable				
treated x post	-5.42	-3.96	-8.247	-10.66
clustered SE	(2.862)*	(4.015)	(4.959)	(2.653)***
wild bootstrap p-value	0.091*	0.396	0.14	0.048**
treated x post x Sri Lankan Tamil	-	-	4.072	11.41
clustered SE			(4.975)	(4.460)**
wild bootstrap p-value			0.459	0.049**
treated x post x Others	-	-	4.422	7.15
clustered SE			(4.087)	(2.220)***
wild bootstrap p-value			0.311	0.123
N	2400	2400	2400	2400
R^2	0.326	0.319	0.536	0.694

Note: Treatment is assigned based on classification in Appendix Table A.1. The regression in column 1 reports the coefficient for equation 1, giving the effect of war on fertility in Sri Lanka. The regression includes district and year-fixed effects. Column 2 adds district population weights. Columns 3 and 4 correspond to estimates from equation 3, exploring heterogeneity by ethnic groups. Column 3 does not include population weights whereas column 4 does. Columns 3 and 4 control for district, year, ethnicity, district x ethnicity, and year x ethnicity fixed effects. Standard errors clustered at the district level are reported in the parenthesis. The wild cluster bootstrapped p-values are also reported for each coefficient. ∗p<0.1, ∗∗p<0.05, ∗∗∗p<0.01.

Table A.10: Impact of Conflict on Age at Marriage, Robustness

Dependent Variable: Female Age at Marriage	
	(1)
Independent Variable	
treated x post	1.93
SE	(0.485)***
p-val	0.009***
N	441
R^2	0.853

Note: Treatment is assigned based on classification in Appendix Table A.1. The regression in column 1 reports the coefficient for equation 2, giving the effect of war on female age at marriage in Sri Lanka. The regression includes district and year-fixed effects. Estimates are weighted by district population. Standard errors clustered at the district level are reported in the parenthesis. The wild cluster bootstrapped p-values are also reported for each coefficient. The mean, minimum, maximum, and standard deviation for mean age at marriage is (24.062, 17.3, 30, 1.728). $*p<0.1$, $**p<0.05$, $***p<0.01$.

Table A.11: Replacement Effect Mechanism- Impact of Conflict on Fertility, Robustness

	Dependent Variable: Crude Birth Rate
	(1)
Independent Variable:	
treated x post	-10.336
SE	(2.647)***
p-val	0.051*
treated x post x deathrate lag	-6.496
SE	(6.461)
p-val	0.2773
treated x post x Sri Lankan Tamil	8.001
SE	(4.492)*
p-val	0.104
treated x post x Sri Lankan Tamil x death rate lag	38.996
SE	(15.328)**
p-val	0.275
N	2250
R^2	0.703

Note: Treatment is assigned based on classification in Appendix Table A.1. The regression includes all pairwise interactions of post, treated, ethnicity, and lag of death rate. The regression also includes district, year, ethnicity, district x ethnic group, and year x ethnic group fixed effects. Standard errors clustered at the district level are reported in the parenthesis. The wild cluster bootstrapped p-values are also reported for each coefficient. The mean, minimum, maximum, and standard deviation for lag of death rate is (0.187, 0, 20.818, 1.135). ∗p<0.1, ∗∗p<0.05, ∗∗∗p<0.01.

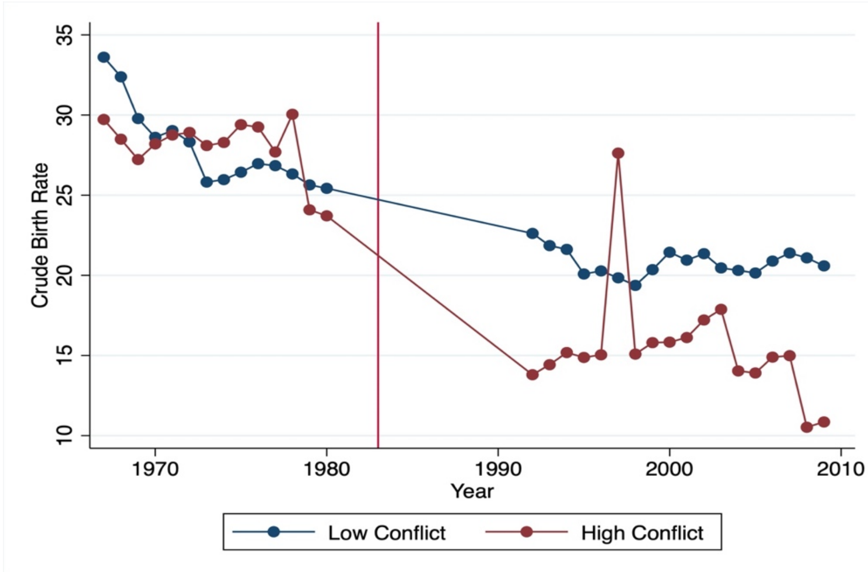

Figure A.1: Parallel Trends in Crude Birth Rate Pre-War

Note: This figure graphs the average crude birth rate (calculated over all ethnic groups) for the low and high-conflict zones over the years. It depicts parallel trends in the pre-war years as the two curves trend closely together. During the war, there was a divergence, and the high-conflict areas witnessed a larger reduction in fertility.

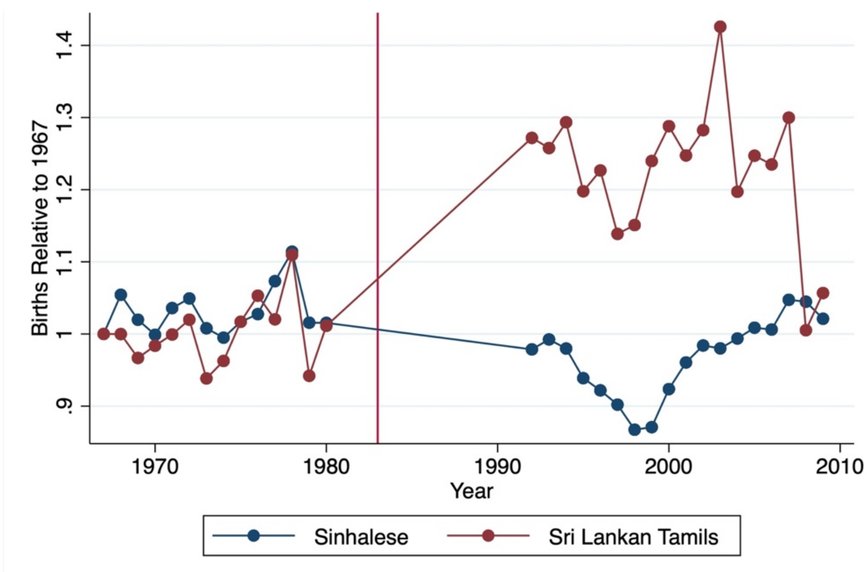

Figure A.2: Trend in Size of Relative Birth Cohort for Sri Lankan Tamils and Sinhalese

Note: This figure depicts the size of the birth cohort over time relative to 1967 levels for the two ethnic groups across Sri Lanka.

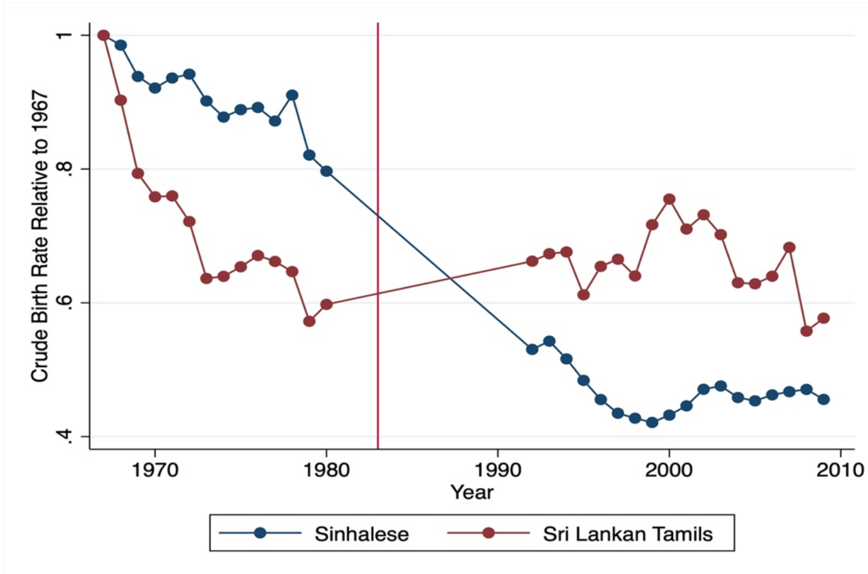

Figure A.3: Trend in Relative Crude Birth Rate for Sri Lankan Tamils and Sinhalese

Note: This figure depicts a trend in crude birth rate over time relative to 1967 levels for the two ethnic groups across Sri Lanka.

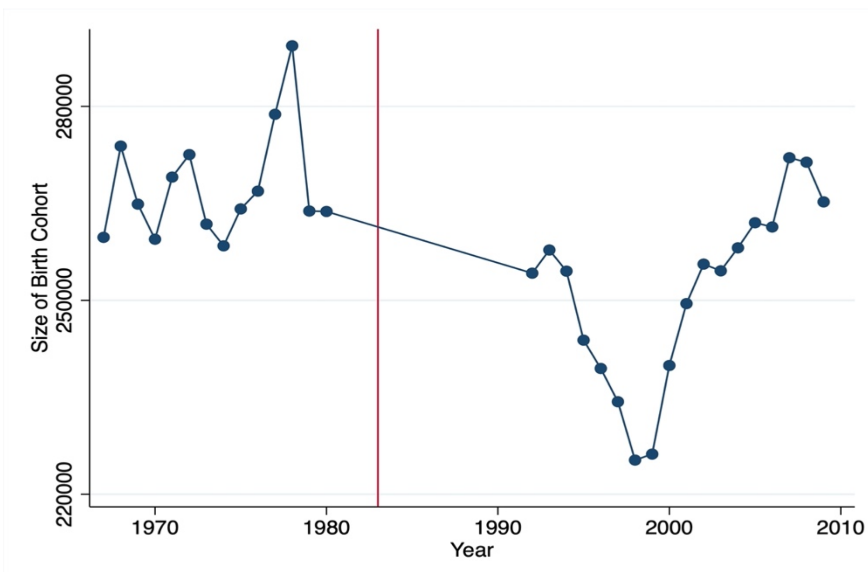

Figure A.4: Trend in Size of Absolute Birth Cohort for Sinhalese

Note: This figure depicts the absolute size of the birth cohort over time for Sinhalese across Sri Lanka.

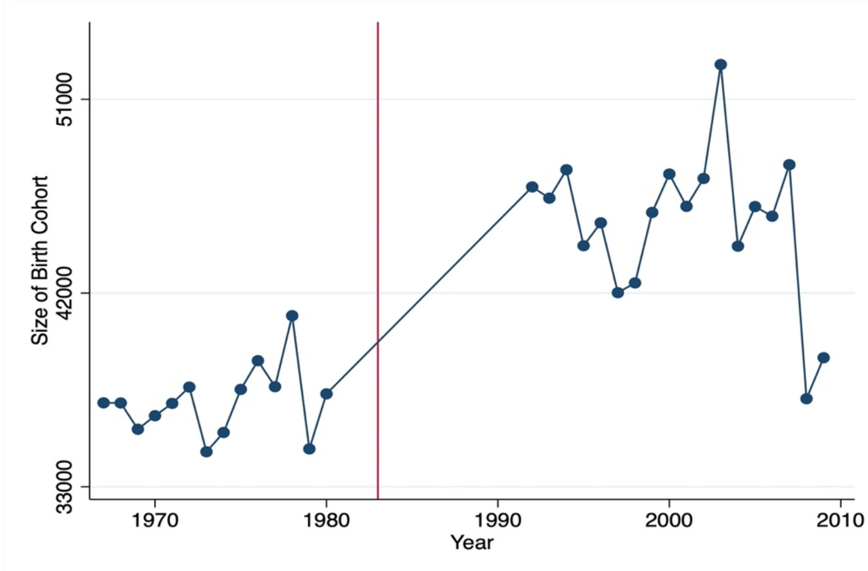

Figure A.5: Trend in Size of Absolute Birth Cohort for Sri Lankan Tamils

Note: This figure depicts the absolute size of the birth cohort over time for Sri Lankan Tamils across Sri Lanka.

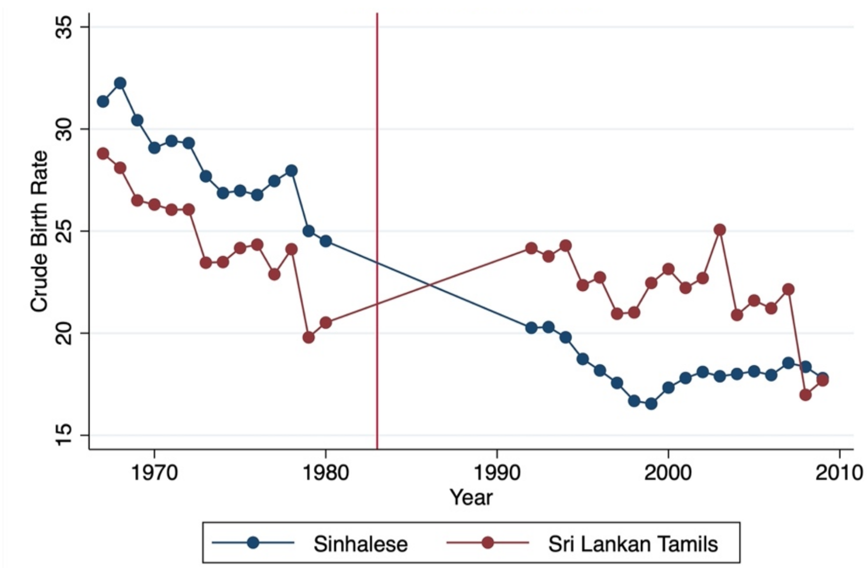

Figure A.6: Trend in Absolute Crude Birth Rate for Sri Lankan Tamils and Sinhalese

Note: This figure depicts the absolute crude birth rate over time for Sri Lankan Tamils and Sinhalese across Sri Lanka.

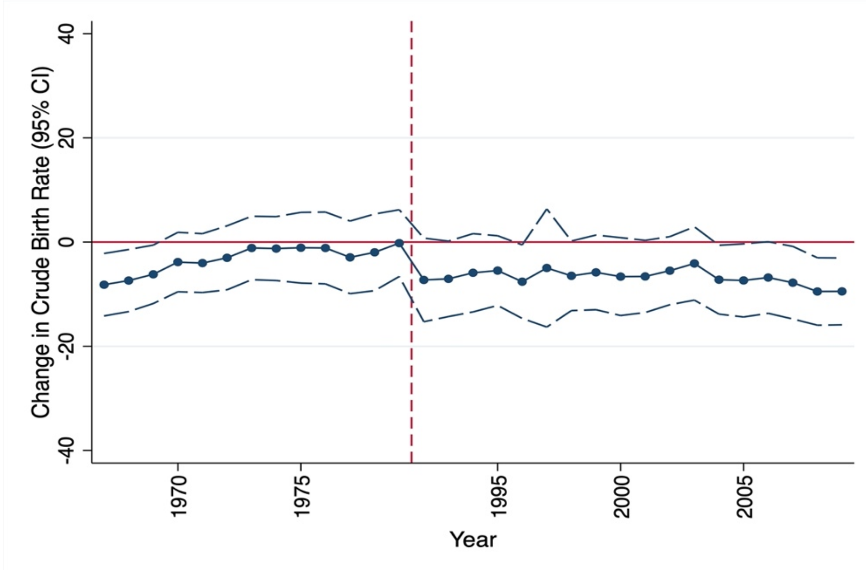

Figure A.7: Dynamic Effect of Conflict on Crude Birth Rate of Sri Lanka

Note: This figure plots the coefficient and 95% CI of the effect of conflict on the crude birth rate of Sri Lanka.

Appendix B: Appendix to Chapter 2

Table B.1: Mapping of HISCLASS Categories

12 HISCLASS Categories	4 HISCLASS Categories
Higher managers (1)	white collar (1)
Higher professionals (2)	white collar
Lower managers (3)	white collar
Lower prof and clerical, sales (4)	white collar
Lower clerical and sales (5)	white collar
Foremen (6)	white collar
Skilled workers (7)	Skilled/semiskilled (2)
Farmers (8)	Farmer (4)
Lower skilled workers (9)	Skilled/semiskilled
Lower skilled farm workers (10)	Skilled/semiskilled
Unskilled workers (11)	Unskilled (3)
Unskilled farm workers (12)	Unskilled

Note: This table maps the 12 HISCLASS categories to the 4 broad categories.

Table B.2: Mobility matrix - Treatment districts

	Father's Occupation				
Son's Occupation	White Collar	Skilled/SemiSkilled	Unskilled	Farmer	Row Sum
White Collar	7,399	9,776	1,996	518	19,689
	{34.45}	{10.16}	{9.00}	{9.97}	
Skilled/SemiSkilled	11,494	76,584	12,081	2,148	102,307
	{53.51}	{79.61}	{54.48}	{41.34}	
Unskilled	2,391	9,347	7,874	983	20,595
	{11.13}	{9.72}	{35.51}	{18.92}	
Farmer	196	489	226	1,547	2,458
	{0.91}	{0.51}	{1.02}	{29.77}	
Total	21,480	96,196	22,177	5,196	145,049

Note: The entries (in curly brackets) represent the number (the share) of sons working in a row occupation among sons whose fathers work in column occupation based on the four-category HISCLASS classification. Observations include sons who are 2-17 years old when their father's occupation is measured in 1851 and 12-27 years old when their own occupation is measured in 1861.

Table B.3: Mobility matrix - Spillover districts

	Father's Occupation				
Son's Occupation	White Collar	Skilled/SemiSkilled	Unskilled	Farmer	Row Sum
White Collar	8,161	9,469	3,414	1,952	22,996
	{33.19}	{8.38}	{4.43}	{9.15}	
Skilled/SemiSkilled	11,867	85,434	25,966	6,102	129,369
	{48.26}	{75.65}	{33.66}	{28.60}	
Unskilled	3,963	17,038	46,931	4,770	72,702
	{16.12}	{15.09}	{60.83}	{22.36}	
Farmer	600	988	839	8,509	10,936
	{2.44}	{0.87}	{1.09}	{39.89}	
Total	24,591	112,929	77,150	21,333	236,003

Note: The entries (in curly brackets) represent the number (the share) of sons working in a row occupation among sons whose fathers work in column occupation based on the four-category HISCLASS classification. Observations include sons who are 2-17 years old when their father's occupation is measured in 1851 and 12-27 years old when their own occupation is measured in 1861.

Table B.4: Mobility matrix - Control districts

Son's Occupation	Father's Occupation				
	White Collar	Skilled/SemiSkilled	Unskilled	Farmer	Row Sum
White Collar	6,911	8,164	3,923	2,359	21,357
	{33.50}	{9.16}	{4.22}	{9.50}	
Skilled/SemiSkilled	9,148	63,000	28,412	6,166	106,726
	{44.35}	{70.65}	{30.58}	{24.82}	
Unskilled	3,917	16,927	59,430	5,472	85,746
	{18.99}	{18.98}	{63.95}	{22.03}	
Farmer	653	1,084	1,160	10,846	13,743
	{3.17}	{1.22}	{1.25}	{43.66}	
Total	20,629	89,175	92,925	24,843	227,572

Note: The entries (in curly brackets) represent the number (the share) of sons working in a row occupation among sons whose fathers work in column occupation based on the four-category HISCLASS classification. Observations include sons who are 2-17 years old when their father's occupation is measured in 1851 and 12-27 years old when their own occupation is measured in 1861.

Table B.5: Mobility matrix - All districts

Son's Occupation	Father's Occupation				
	White Collar	Skilled/SemiSkilled	Unskilled	Farmer	Row Sum
White Collar	22,471	27,409	9,333	4,829	64,042
	{33.69}	{9.19}	{4.85}	{9.40}	{10.52}
Skilled/SemiSkilled	32,509	225,018	66,459	14,416	338,402
	{48.74}	{75.43}	{34.57}	{28.06}	{55.60}
Unskilled	10,271	43,312	114,235	11,225	179,043
	{15.40}	{14.52}	{59.42}	{21.85}	{29.42}
Farmer	1,449	2,561	2,225	20,902	27,137
	{2.17}	{0.86}	{1.16}	{40.69}	{4.46}
Total	66,700	298,300	192,252	51,372	608,624

Note: The entries (in curly brackets) represent the number (the share) of sons working in a row occupation among sons whose fathers work in column occupation based on the four-category HISCLASS classification. Observations include sons who are 2-17 years old when their father's occupation is measured in 1851 and 12-27 years old when their own occupation is measured in 1861.

Table B.6: Placebo check p-values (randomizing treatment)

	(1)	(2)	(3)
Panel A- Occ. Category son ≠ Occ. Category fataher			
	p < 0.00016	*p < 0.01*	*p < 0.05*
Proportion of placebo regressions	0.008	0.053	0.158
Panel B- | Occ. Rank son - Occ. Rank father|			
	p < 0.00000	*p < 0.01*	*p < 0.05*
Proportion of placebo regressions	0	0.056	0.172
Panel C- Occ. Rank son > Occ. Rank father			
	p < 0.00000	*p < 0.01*	*p < 0.05*
Proportion of placebo regressions	0	0.098	0.199

Note: The randomization was iterated 1000 times. Column 1 looks at the proportion of placebo regressions that outperform the original with a p-value less than the p-value of the main estimate in Table 2.4, column 3. Columns 2 and 3 look at the proportion of placebo estimates with $p<0.01$ and $p<0.05$.

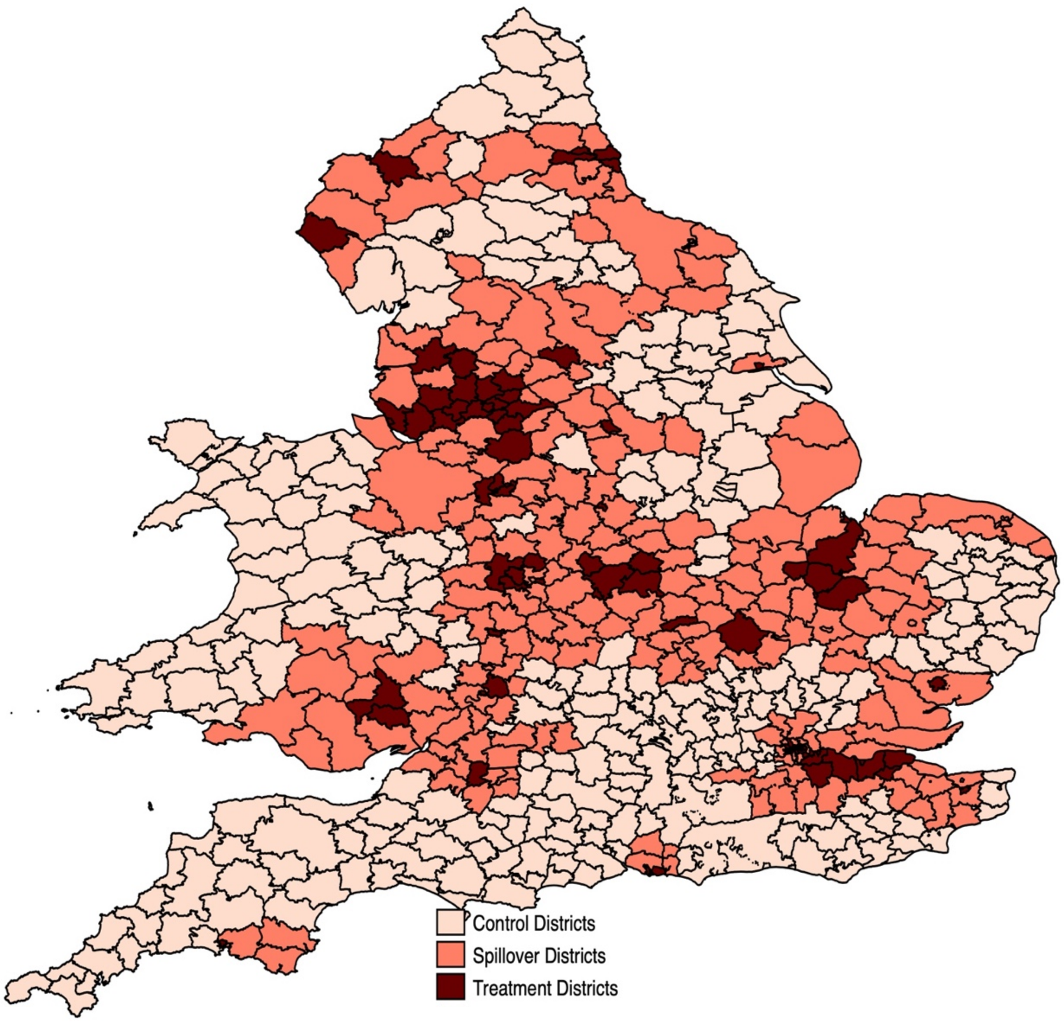

Figure B.1: Treatment, Spillover, and Control Districts in England and Wales

Note: This figure denotes the geographical location of treatment, spillover, and control districts in England and Wales under the 1848 Public Health Act. Treatment districts are classified based on mortality of 23 deaths per 1000 people or more. Neighbors of the treatment districts are classified as spillover districts. The remaining districts are control districts.

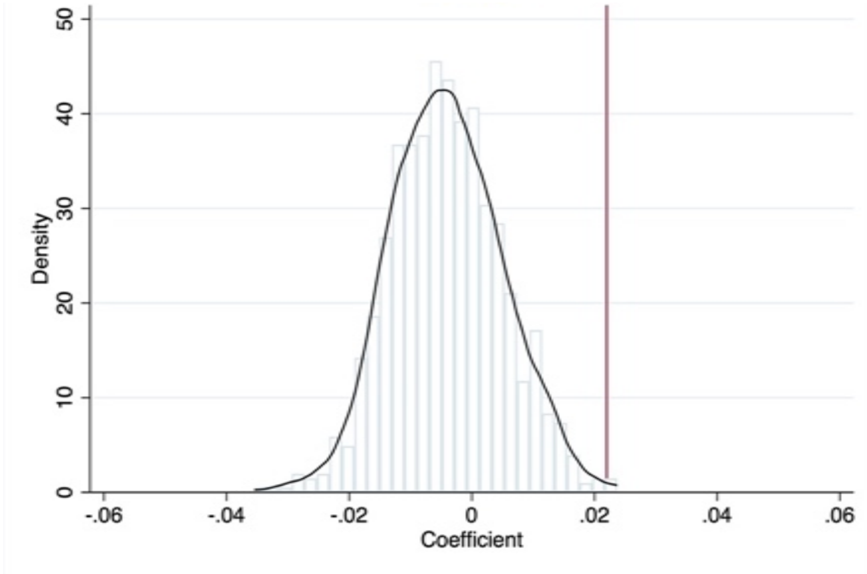

Figure B.2: Placebo Check- Occ. Category son ≠ Occ. Category father

Note: The distribution for the coefficients has been generated by randomizing treatment 1000 times. The red marker indicates the main estimates I obtain in Table 2.4, column 3.

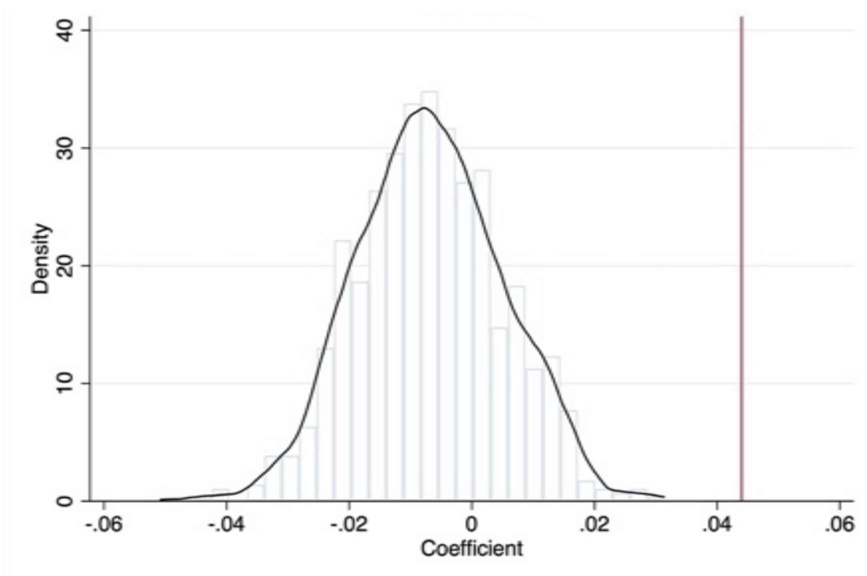

Figure B.3: Placebo Check- | Occ. Rank son - Occ. Rank father |

Note: The distribution for the coefficients has been generated by randomizing treatment 1000 times. The red marker indicates the main estimates I obtain in Table 2.4, column 3.

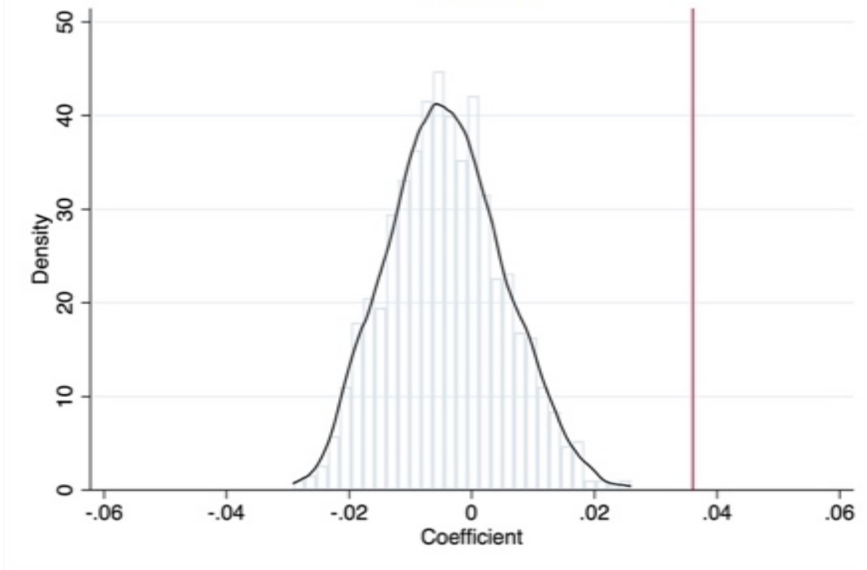

Figure B.4: Placebo Check- Occ. Rank son > Occ. Rank father

Note: The distribution for the coefficients has been generated by randomizing treatment 1000 times. The red marker indicates the main estimates I obtain in Table 2.4, column 3.