Long-term health and human capital effects of massive investments in public health and education: Evidence from Cuba*

Giovanni Mellace†
University of Southern Denmark

Rok Spruk‡
University of Ljubljana

October 30, 2025

Abstract

We study the impact of establishing a national health service and investing heavily in education on public health and human capital. Exploiting Cuba's 1960 healthcare reform and using Synthetic Control methods with data from former European colonies in the Western Hemisphere, we find that infant mortality fell by 35%, life expectancy rose by two years, and average schooling increased by 1.4 years. These effects are robust and persistent, highlighting the long-term benefits of universal healthcare and expanded educational access.

Keywords: Public Health, Human Capital, Synthetic Control Method, Cuba.
JEL classification: C23, I15, I18, I25, I30, O43.

*The authors gratefully acknowledge helpful comments from Thomas Barnebeck Andersen, ..., and Anthony Wray.

†Department of Economics, Campusvej 55, 5230 Odense M, Denmark, giome@sam.sdu.dk

‡School of Economics and Business, Kardeljeva ploščad 17, 1000 Ljubljana, Slovenia, rok.spruk@ef.uni-lj.si

1 Introduction

Investments in public health services are fundamental to enhancing population health outcomes and fostering human capital development. The implementation of universal health care systems has been associated with significant improvements in infant mortality rates, life expectancy, and educational attainment (Grossman, 1972). Understanding the long-term effects of such systems, especially in varying socio-economic contexts, is crucial for informing policy decisions.

In 1960, Cuba established a comprehensive National Health Service aimed at providing equitable and accessible health care to all citizens (Keck and Reed, 2012). This significant policy shift transformed Cuba's health care landscape and presents a unique opportunity to examine the causal impact of large-scale public health investments on key health and human capital indicators. While previous research has highlighted Cuba's notable health achievements despite economic challenges (Nikelly, 1988; Feinsilver, 2023), there is a lack of empirical studies quantifying the long-term effects of its National Health Service on public health outcomes and educational attainment.

This study examines the contribution of Cuba's National Health Service to public health and human capital formation from 1870 to 2022. By exploiting the establishment of the National Health Service as a source of variation in long-term public health and human capital formation, we aim to estimate the causal effects of universal health care provision.

Our identification strategy utilizes the synthetic control method developed by Abadie and Gardeazabal (2003) and Abadie et al. (2010) to build a synthetic version of Cuba that closely mirrors its pre-intervention trends as a weighted combination of former European colonies in the Western Hemisphere that did not implement comprehensive universal health care during the same period. Comparing Cuba's actual outcomes with this synthetic control allows us to estimate the causal effects of the National Health Service on infant mortality, life expectancy, and human capital formation.

Our findings indicate that the introduction of equitable and accessible public health care provision is associated with a permanent decrease in infant mortality by around 35 percent, a sustained improvement in life expectancy at birth by two years, and an increase of 1.4 years in the average length of schooling. These effects are robust to an extensive battery of placebo analyses and persist up to the present day.

This study contributes to the literature in several ways. First, it provides empirical evidence on the long-term impact of universal health care on public health outcomes in a developing country context. Second, it explores the spillover effects of health improvements on human capital formation, emphasizing the interconnection between health and education (Bleakley, 2010). Third, it employs a robust identification strategy that accounts for potential confounding factors, ensuring that the estimated effects are attributable to the health care reform. The remainder of the paper is organized as follows. Section 2 introduces the historical and institutional background, section 3 provides the discussion of data and samples. Section 4 presents the identification strategy whereas Section 5 discusses the results and the corresponding robustness checks. Section 6 concludes.

2 Historical and Policy Background

Cuba's health care system is a subject of rigorous scholarly and policy debate that has received substantial attention since the triumph of the revolution in 1959. Its history is both fascinating and perplexing and pinpointed by notable public health achievements that has drew laudable admiration and envy from the rest of the world. In response, Cuban health paradox lately

emerged as a widely circulated term as an exception that proves the rule, colloquially known as the black swan (Spiegel and Yassi 2004). In particular, the kernel of this discussion emphasizes Cuba's envious public health milestones and sustained health performance, comparable to the advanced economies, in spite of its poor economic performance (Evans 2008). Since a detailed discussion of Cuba's health policy and its major breakthrough is a subject beyond the scope and objective of our study and can be found elsewhere (Cooper et. al. 2006, Lamrani 2021), a brief excursus and discussion will be provided here.

Before 1959, Cuba's health care system and provision suffered from a number of deficiencies and shortcomings. Despite the independence from Spain in 1902 and revolutionary scientific advances in the field of epidemiology by renowned Cuban scientists[1] , the provision of health care had been characterized by a significant inequalities and stark divide between affluent and relatively prosperous urban areas and poor, undeveloped and deprived rural areas outside Havana. In particular, the rural areas suffered heavily from the absence of basic health-care infrastructure and widespread omnipresence of preventable diseases. Before the revolution, access to health care and strong medical infrastructure were principally limited to the rich and affluent residents of Havana, who enjoyed access to world-class medicine and services provided almost exclusively by private clinics (Danielson 1979; Keck 2012). On the contrary, the rural areas only had one hospital and suffered from the widespread and pervasive presence of infectious diseases such as tuberculosis, malaria and a series of parasitic infections that were paramount among the poor.

A large and pervasive urban-rural gap had been further exacerbated by the shortage of well-trained doctors, widespread illiteracy, and high rates of infant mortality in the rural areas. For instance, a report by Truslow (1951) shows that, inter alia, only eleven percent of farm workers drank milk whilst the rural infant mortality rate stood at 100 per 1,000 live births. In addition, the rural areas experienced deleterious oral health (Watson et. al. 2022), further exacerbated by the excessive consumption of sugar cane. Coupled with extensive child malnutrition and widespread prostitution (Miller and Kennedi 2003, Sippial 2013, Kurlansky 2017), public health disparities between the affluent and poor strata of the society were paramount. On the eve of the revolution, the estimated life expectancy at birth in dense urban areas such as Havana, Santiago de Cuba, Cienfuegos, Camaguey and Santa Clara was between 60 and 62 years whilst the estimated life expectancy in small towns and rural areas was around 45 and 50 years, which corresponding to a urban-rural gap of around 10 to 15 years. Furthermore, around 80 percent of the children in the countryside had intestinal parasites and about 60 percent of the population was chronically malnourished (McGuirre and Frankel 2005, Gorry 2012, Delgado Legon 2018). Astonishing health disparities between the urban and rural areas were further reinforced by the widespread corruption that permeated every aspect of Cuban life such as electoral fraud and authoritarian repression, coaptation and complicity of government officials with the organized crime and mafia syndicates, which reached its peak during the rule of Batista regime (1952-1959).

After the revolution in 1959, equitable provision of health and education Cuban health system underwent a series of tectonic changes, pinpointed by a complete transformation into a

[1]One of the best known advances in Cuba before the revolution is the discovery of link between mosquitos and yellow fever by Carlos J. Finlay in 1881. Finlay presented a theory that Aedes aegypti mosquitoes were responsible for transmitting the disease, and conduct a variety of experiments proving that infected mosquitos could pass yellow fever to healthy individuals. Although initially ignored, Finlay's research and theory was confirmed by the U.S. Yellow Fever Commission in 1900. His research stimulated an important impetus to fields of biomedicine and epidemiology, and promulgated to a series of effective mosquito control programs which, inter alia, led to the eradication of yellow fewer in 1901, facilitated the completion of the Panama Canal between 1904 and 1914 that was halted due to the periodic outbreaks of the fewer, and also influenced modern mosquito-borne disease control such as malaria and dengue (Barnet 1915, Chaves-Carballo 2005).

universal, state-funded and community-based health care.[2] The establishment of the National Health System (i.e. Sistema Nacional de Salud) in 1961 earmarked a formal transition towards a comprehensive and universal health that consisted of several gradual steps. First, the right to universal health care had been enshrined in the Cuban constitution and declared a universal human right and free to all citizens. In this respect, hospitals and clinics were nationalized and brought under the state control. Second, the provision of preventive health care underwent a massive and unprecedented expansion of hospitals and clinics into the rural areas.[3] Third, after 1961, Cuban government massively increased medical school enrollment rates which gradually led to more doctors per capita compared to the advanced, high-income economies. In response, Cuba gradually emerged as one of the global leaders in medical diplomacy by sending doctors to more than sixty countries up to the present day to combat both infections diseases such as the outbreak of Ebola in West Africa in 2013. Against this backdrop, De Vos (2005) identifies four stages of Cuba's health system development: (i) establishment of foundations in the early post-revolutionary period (1959-1970), (ii) consolidation in the succeeding decade (1970-1979), (iii) full expansion and development of family medicine (1980-1990), and (iv) reform and adjustment stage (1990-present).

Since the revolution, Cuba implemented a series of significant large-scale investments in preventive care and medicine ranging from 6 percent of the GDP, or roughly one quarter of the national budget, in the early years to 9 percent of the GDP in 2010, 11 percent of GDP in 2015, and 14 percent of the GDP in 2021 (Sixto 2002, Mesa Lago 2002, Anuario Estadístico de Salud 2020). In this respect, Cuba consistently outperformed advanced and industrialized economies as well as the rest of Latin America in terms of the proportion of GDP invested into primary care and preventive medicine. Henceforth, Cuba achieved significant and notable public health improvements that paled favorably towards high-income countries. By way of example, the rate of infant mortality decreased from 60 deaths per 1,000 live births in 1950s to around 4 deaths per 1,000 live births up to the present day, boasting one of the lowest rates of infant mortality in Latin America and the rate that is comparable with Canada. In terms of further example, life expectancy at birth increased from 59 years in 1959 to 80 years in 2023. Up to the present day, together with Chile and Costa Rica, Cuba is the only Latin American country exceeding the eight-year life expectancy threshold so far. The depth and breadth of Cuba's public health improvement is best exemplified in the remarkable eradication of infectious diseases such as malaria (in 1967), outright containment of dengue fever (in 1981), polio (in 1982), measles (in 1993), and mother-to-child transmission of HIV and syphilis (in 2015), receiving commendable noteworthy praise and recognition from the academic and policymaking community.

One of the most pivotal achievements of the large-scale public investment in health care emanates from Cuba's international medical diplomacy. As a part of the large-scale humanitarian missions, Cuba continually dispatched medical teams abroad, especially to developing countries and crisis zones, to combat the outbreak of infectious diseases, natural and nuclear disasters, shortage of skilled medical staff and, more recently, COVID19 pandemic. In particular, Cuban government pioneered and completed its first humanitarian mission in Algeria (in 1963), combating the shortage of skilled doctors and health professionals in response to the mass exodus of

[2]Radical transformation of the health care system based on the principles of equity, solidarity and preventive after the revolution is colloquially known as the revolutionary medicine, and was the subject of the well-known speech by Ernesto Che Guevara in 1960 at the Pan-American Health Organization conference in Havana (Guevara 1969)

[3]One of the most well-known and widely recognized expansion of the preventive health care is the Family Doctor Program, inaugurated by Cuban government in 1984. As a comprehensive plan to increase medical care in the family and community environment, the plan presents the complete application of primary care, focusing on improvement, quality and efficiency of health services in family medicine with emphasis on preventive care.

French doctors after Algeria declared independence from France in the year beforehand. Henceforth, Cuban government launched several comprehensive international humanitarian missions such as Operation Miracle (in 2004), providing eye surgeries to treat blindness to millions in Latin America and Africa, Ebola missions in West Africa (in 2014), and more recently, deploying 38 brigades with a total of 2,544 healthcare professionals dispatched across 39 countries to combat COVID19 pandemic. Cuba also became a global leader in the treatment and recovery of children after the Chernobyl disaster, and offered free-of-charge care and treatment to more than 25,000 children from Ukraine, Belarus and Russia in the years after the nuclear disaster. In addition, Cuban government established Latin American School of Public Medicine (*Escuela Latinoamericana de Medicina*) – the largest medical school worldwide based on enrolment count - offering free-of-charge education to almost 20,000 students from 110 countries, where more than 30,000 doctors have been trained from under-served communities.[4]

One of the principal outcomes of Cuba's health-care transformation is best exemplified by the development of cutting-edge vaccines and biotechnological vaccines. Despite the U.S economic sanctions and embargo in place since 1958 and the adverse effects on economy and health (Garfield and Santana 1997, Barry 2000, Gordon 2016), Cuba developed a series of low-cost and effective vaccines and treatments. In 1989, Cuba developed the Meningitis B vaccine as the first-ever effective vaccine against the deadly diseases and received wide-standing international recognition and praise (Randall 2000, Pérez et. al. 2014, Sierra-González 2019). In 1992, Cuba developed a recombinant Hepatitis B vaccine, which spearheaded an effective and dramatic reduction in the infection rates. Before the development of vaccine, Cuba had an estimated Hepatitis B incidence rate of around 10 to 15 cases per 100,000 inhabitants. By early 2000s, the incidence rate had fallen to less than one case per 100,000, which represents a dramatic reduction that has attributed to a high coverage of the two-stage vaccination program (Ivón et. al. 2017).[5] More recently, Cuba developed ground-breaking therapeutic lung-cancer vaccine (i.e. CIMAvax-EGF) for non-small cell lung cancer in 2011, capable of extending the lives of patients (Crombet Ramos et. al. 2015, Tagliamento et. al. 2018). Despite the limited resources and U.S embargo, Cuba also developed two variants of COVID 19 vaccines in 2021 (i.e. Abdala and Soberana), based on protein subunit and conjugate types, which demonstrated efficacy rates of 92 percent and 91 percent in clinical trials, which are comparable to the leading global vaccines (Más-Bermejo et. al. 2022).

Compared to the initial phase of the post-revolutionary development, Cuba continually allocates more than 10 percent of GDP to health-care and has achieved envious levels of public health exemplified by high life expectancy comparable to the richest countries worldwide, high rates of prenatal care, lowest rate of maternal mortality in the region, and also one of the highest number of centenarians per capita. In 1965, Cuba was the first country in Latin America to legalize abortion[6] and, more recently, introduced a very modern family code approved in 2022 referendum,

[4]It should be noted that students primarily come from Africa, Latin America, Carribean, Asia and also from the United States. The graduates work in low-income regions, providing health care to population with limited health service, and commit to public service rather than private practice in the effort to reduce brain drain of medical professionals to wealthier countries. Countries benefitting the most from trained doctors and graduates include Haiti, Venezuela, Honduras, Bolivia and South Africa among several others.

[5]In the first stage, the high-risk groups (i.e. infants born to infected mothers, healthcare workers) were target. In the second stage, the vaccination was extended universally to infants and adolescents.

[6]Compared to Cuba, other countries in Latin America have not introduced abortion earlier than in 2012 (Uruguay), 2020 (Argentina). In some countries such as Colombia, abortion is decriminalized up to 24th week of pregnancy wheres, in Mexico, divorce laws vary widely across states from the availability upon request to the comprehensive ban. In a notable number of countries (i.e. El Salvador, Nicaragua, Honduras, Chile, Dominican Republic), abortion is banned in all circumstances with no exceptions. See Kulczycki (2011) for a detailed

widely considered one of the most progressive ones in Latin America and even globally (Adler 2024).[7] Before the changes introduced into the family code, Cuba has been long regarded as one of the best developing countries for prospective motherhood based on the detailed report from Save the Children foundation.[8]

The key cornerstone behind Cuba's complete health-care transformation after 1959 is posited by a series of large-scale investments in health care that fundamentally transformed the system into one of the most comprehensive ones worldwide (Conover et. al. 1980). One of the first policy shifts was the declaration that health care is a human right, not a privilege. This principle laid the necessary foundations for the post-revolutionary health policy (Boffey 1980). Unlike other countries in Latin America, Cuba replaced private healthcare with a state-run and government-funded health care, available to all Cuban citizens. In the formative period of health policy reform, Cuban government built hundreds of new hospitals and clinics across the island without neglecting rural and mountainous regions. In this respect, private hospitals that were previously accessible only to the wealthy elites and upper middle-classes in major cities, were transformed into public facilities. In the early 1960s, Cuba decentralized health care provision by introducing a polyclinic model which consisted of the network of localized and neighborhood-based health-care centers that offered all-in-one primary care, prevention specialized services, ensuring medical services were available to all communities (Hirschfeld 2017).

In 1961, Cuban government made medical education free, encouraging thousands of young Cubans to become doctors and nurses without financial constraints. Medical schools emphasized and reinforced community-based training by preparing doctors, nurses and specialists to serve in rural and urban environment. In this respect, Cuba focused heavily on preventive medicine to reduce the burden on hospitals and preserve relative good health of the civilian population. These efforts were principally spearheaded by mass immunization programs against infectious diseases and mandatory vaccinations. Since 1960, Cuban government implemented a series of programs to improve child nutrition, ensure access to clean water and sanitation, leading to dramatic reduction in infant mortality and infectious diseases. In terms of further example, the Cuban government introduced a nation-wide maternal and child health program, providing pre-natal care, free childbirth services and infant vaccinations (Bodenstein 2010). In the 1970, Cuba began sending doctors abroad to assist disaster-stricken regions in Africa, Latin America and elsewhere as a part of its medical diplomacy (Feinsilver 1989, Huish 2014). In 1980, Cuba further developed its biotechnology sector to produce vaccines, medicine and provide cutting-edge research (Villanueva and Espiñeira 2023).[9]

overview of abortion practices and legislation in Latin America

[7]The new family code in 2022 promulgated the legalization of same-sex marriage and child adoption for same-sex couples, permitted altruistic child surrogacy, emphasized gender equality in all aspects of family life, further protected the individuals from domestic violence and discrimination within the family. The new code also recognized a wider range of family structures including multi-generational households, single-parent families and non-traditional partnerships. The code also prioritized children's rights to be heard in family matters and protection from harm and discrimination. The code also explicitly prohibited discrimination on the basis of sexual orientation and gender identity, and massively enhanced the rights of LGBT+ individuals in all aspects of family life. Against this backdrop, the family code has been praised as the most progressive one in Latin America and, in many ways, more progressive compared to the United States (Adler 2023).

[8]The composite index of the quality of motherhood encompassed several criteria such as the quality of health and education system, assistance by skilled health personnel during childbirth, distribution of contraceptive methods, level of women's education, political and economic gender equality, and the rate of participation in political life and equal pay. See: *Informe Sobre El Estado Mundial de las Madres 2011.* Madrid: Save the Children.

[9]One of the best known examples of massive biotechnology investments is the creation of Finlay Institute (in 1991) and the Center for Genetic Engineering and Biotechnology (CIGB) (in 1985) which led many vaccine

In addition to the establishment of the National Health Service in 1961, the Cuban government launched the Literacy Campaign (i.e. *Campaña Nacional de Alfabetización en Cuba*, a nationwide effort aimed at eradicating widespread illiteracy inherited from pre-revolutionary period, in the same year. Existing estimates (Kellner 1989) indicate that the literacy rate before the revolution was between 60% and 76% with notable disparities between urban and rural areas.[10] This campaign, which took place over the course of just eight months, was one of the most significant educational initiatives in Cuban history. During the course of eight months, more than a quarter million volunteers (that is, *brigadistas*) were dispatched and mobilized to teach reading and writing skills in economically and socially disadvantaged rural and urban communities without electricity, running water, and paved roads. The campaign had been centered around the construction of new schools, training new educators and teaching predominantly illiterate peasants (i.e. *guajiros)* to read and write.[11] The campaign resulted in a rapid and dramatic increase in literacy rates. By the end of the campaign, around 707,000 Cubans or approximately 10 percent of the population became literate, effectively reducing the national literacy rate to 3.9%. By 1962, Cuba's literacy rate stood at 96%, and was comparable with the most advanced democracies in Western Europe and North America. By 1981, the literacy rate reached 98% and outperformed many countries at higher level of economic development such as Spain, Portugal and Greece. The timing of the literacy campaign coincides with the foundation of the National Health Service, creating a challenge in disentangling the independent effects of each policy on human capital outcomes such as average years of schooling. However, while the Literacy Campaign had a profound impact on education, its relatively short duration means that it is more likely to have had a complementary lasting effects on trajectory of educational attainment in coalesence with a longer-term impact of the health system reforms. In this paper, we focus on the broader, more sustained effects of health care investments but acknowledge that the simultaneous educational reforms may influence the interpretation of changes in schooling outcomes during this period.

Compared to the other countries in Latin America and United States, Cuban government consistently prioritized healthcare and education even during the economic crisis in the 1990s, colloquially known as the Special Period, after the collapse of Soviet Union and the loss of trade. Despite the limited resources, Cuba maintained its free health-care system and the government covered costs through state revenues and international medical services. By maintaining a steady investment in medicine, research and public health infrastructure, Cuba enshrined long-term commitment to health as a national policy.

innovations such as the breakthrough development of Meningitis B vaccine in 1989.

[10]By way of example, the estimated rate of urban literacy in 1959 was approximately 89% opposed to a mere 59% in rural areas (Rey 2021)

[11]It should be emphasized that the literacy campaign had not been the only policy instrument implemented to boost human capital formation. In terms of further example, access to higher education had been accomplished by a series of policies such as the introduction of gender quotas and several other affirmative action policies. The policies were best echoed by Che Guevarra's speech to the students and faculty at University of Las Villas stating that *"...days when education was a privilege of the white middle class" had ended. The university must paint itself black, mulatto, worker, and peasant. If it did not, the people were going to break down its doors "and paint the University the colors they like."* (Anderson 1997)

3 Data

3.1 Outcomes and samples

Our sample comprises 21 former European colonies in North and South America for the period 1870-2022. Using the average years of education as a rough proxy for human capital formation, we use Barro and Lee (2010) estimates of the mean years of schooling across the full spectrum of age cohorts which enables a straightforward between-country and within-country comparison. The trajectories of infant mortality and life expectancy for the period 1900-2022 are constructed in two-fold step. In first step, we use the outcome trajectories for the period 1950-2022 from UN World Population Prospects as the underlying benchmark. In the second step, we link the benchmark series to the pre-1950 period by making use of Mitchell (2010) estimates of infant mortality for the former European colonies in the Americas, extended back to 1900. In addition, the trajectories of life expectancy at birth from UN World Population Prospects for the period 1950-2022 are linked to the historical estimates of life expectancy for the same underlying sample for the period 1902-1950 based on the ground-breaking historical reconstruction of life expectancy estimates from Riley (2005), Barbieri et. al. (2015), Zijdeman et. al. (2015), Aburto et. al. (2020) among several others.

We consider three specific outcomes to evaluate the long-term effectiveness of the establishment of National Health Service on the trajectories of public health, namely, infant mortality, life expectancy at birth and average years of education. First, infant mortality rate is defined as the number of deaths of infants under one year per 1,000 live births in a given year and is widely considered as perhaps one of the most comprehensive indicators of the overall health system performance that is also reflective of the quality of healthcare, maternal health, quality of nutrition and level of socioeconomic conditions. One of the key limitations behind the reconstruction of the infant mortality trajectory arises from the incompleteness of death registries in Cuba in early 1900s, as shown by Díaz-Briquets (1983). Therefore, the only plausible alternative is to rely on census data to reconstruct a plausible historical trajectory of infant mortality. For the period 1900-1904, we combine Gonzalez Quinones and Ramos Pinol (1996) census-based estimates and Collver (1965) historical estimates for Latin American countries which provide a comprehensive and complete coverage of infant deaths from the birth registers. Since Cuba adopted the World Health Organization-based definitions of live birth, infant death and fetal deaths in 1965 (Catasus Cervera and Hernandez Castellon 1977, Rios Massabot 1983), the infant mortality, linking pre-1950 series with the post-1950 benchmark provides a plausible long-term trajectory of infant morality rate that can be compared both across space and time.

Second, the reconstruction of the historical life expectancy trajectory consists of several steps. In the first step, we use Astorga and FitzGerald (1998) estimates of life expectancy for the year 1900 which is around 32 years, as reported by McGuire and Frankel (2005), and is consistent with the level calculated by Cuban government in 1975 which estimated life expectancy at birth in 1900 at 33.2 years (Centro de Estudios Demograficos 1976, Farnos Morejon 1977). We construct the life expectancy series for the period 1900-1950 by combining year-specific estimates from the original tables of Velásquez and Toirac (1977) and Morejon (1977), Centros de Estudios Demograficos 1976, Mezquita (1970), and Debasa and González Quiñones (1971) as well as Astorga and Fitzgerald (1998) and Pérez Brignoli (1975). Our approach is simple and parsimonious as we compute the median estimate of life expectancy from the respective sources and exploit the implied annual rates of increase to construct a full and uninterrupted time series.

And third, to capture human capital investment, we rely on average years of schooling as a widely recognized indicator of human capital formation (Barro and Lee 2010). The average

years of schooling denote mean number of completed years of schooling that individuals in a given population have attained. By providing an easily tractable summary measure of the educational attainment of the population, it captures the level of education achieved over a period of time consistently. We use recently reconstructed historical estimates of the educational attainment (Wha-Lee and Lee 2016) to build a concise time-series of the average years of education for our period of estimation.

3.2 Auxilarly covariates

To adequately capture and reproduce the outcome trajectories prior to the establishment of National Health Service (i.e. Sistema Nacional de Salud), we build a compact vector of pre-reform outcome values in most recent pre-revolution years together with the auxiliary covariates as a separate class of predictors. To this end, we consider physical geographic characteristics to capture the salience of the disease-related and climatic environment such as latitude (in absolute terms), soil quality, fraction of land area in tropical zone, fraction of land area within 300 km coastline. Second, to capture demographic salience between Cuba and the Latin American country-level donor pool, we include population density as a separate predictor based on the historical reconstruction of density estimates (Klein Goldewijk et. al. 2018) linked to the United Nations estimates of population density in the post-1950 period. Furthermore, to establish salience in terms of colonial historical background, we use recently constructed estimates of the European share of the population during colonization from Easterly and Levine (2016) who demonstrate a strong association between colonial European settlement and present-day differences in economic development. And lastly, we also consider per capita GDP levels (Bolt and Van Zanden 2014) as an additional predictor to further capture salience and similarity between Cuba and Latin American countries in terms of the level of economic development. In consequence, the full set of predictors allows us to capture both the implicit and explicit attributes of Cuba's public health and human capital formation prior to the establishment of the universal health care. In the training stage, both past values of the outcomes and the full set of auxilarly covariates explain between 70% and 90% of the overall variation in infant mortality, life expectancy at birth and average years of education.

4 Identification Strategy

To estimate the causal impact of Cuba's National Health Service (NHS) established in 1960 on public health outcomes and human capital formation, we employ the Synthetic Control Method (SCM) developed by Abadie and Gardeazabal (2003) and formalized in Abadie et al. (2010) and Abadie et al. (2015). The SCM constructs a synthetic version of Cuba by creating a weighted combination of countries from a donor pool—former European colonies in the Western Hemisphere that did not implement comprehensive universal health care during the same period. This synthetic control closely mirrors Cuba's pre-intervention characteristics and outcome trends, providing a credible counterfactual for what would have occurred in the absence of the NHS.

The validity of the SCM relies on the assumption that, in the absence of the intervention, the outcome trajectories of Cuba and its synthetic counterpart would have followed similar paths. By comparing the post-intervention outcomes of Cuba with those of the synthetic control, we attribute any significant deviations to the effect of the NHS.

To enhance the robustness of our estimates, we incorporate several complementary methods. First, we employ the Augmented Synthetic Control Method (ASCM) proposed by Ben-Michael

et al. (2021). The ASCM improves upon the traditional SCM by "augmenting" the SCM estimate with an outcome model and adding a ridge penalty to estimate the weights, therefore adjusting for any discrepancies in pre-intervention outcomes between the treated unit and the synthetic control, effectively reducing bias in the estimated treatment effect. This method is particularly useful when the fit between the treated unit and the synthetic control is imperfect during the pre-intervention period.

Second, we employ the Synthetic Difference-in-Differences (Synthetic DiD) approach developed by Arkhangelsky et al. (2021) as an additional robustness check. The Synthetic DiD combines the strengths of the difference-in-differences estimator with the SCM. This method estimates not only unit specific weights but also give different weights to each pre-intervention period. This allows for time-varying unobserved confounders, thus relaxing the parallel trends assumption, and non perfect fit between the treated and its synthetic counterpart therefore overcoming limitation of the SCM.

Third, we implement the interactive fixed effects (IFE) estimator developed by Xu (2017), which models the counterfactual using a factor model framework. Unlike SCM, which relies on a weighted average of control units, IFE allows for time-varying unobserved heterogeneity by estimating latent factors that capture common trends across units. This approach is particularly valuable when the parallel trends assumption may be violated, as it can accommodate unit-specific responses to unobserved common shocks.

Finally, we apply the matrix completion method proposed by Athey et al. (2021), which treats the missing counterfactual outcomes as entries in a partially observed matrix. This approach leverages regularized principal component analysis to impute the missing values, exploiting both cross-sectional and time-series patterns in the data. By using nuclear norm regularization with varying penalty parameters (λ), the method provides a flexible framework for estimating treatment effects that is robust to model misspecification and can handle complex patterns in the data structure.period.

By utilizing these complementary methodologies, we ensure that our findings are robust to various identification strategies and that the estimated effects are credibly attributed to the introduction of the NHS in Cuba. These methods collectively address potential confounding factors and strengthen the causal interpretation of our results.

5 Results

5.1 Synthetic Control Weights and Covariates Balance

Table 1 reports the composition of the synthetic control groups that best reproduce Cuba's outcome trajectories prior to the introduction of universal health care in 1961. Pre-1961 trajectory of infant mortality is best synthesized as a convex combination El Salvador (62%) , United States (27%), and of Venezuela (11%). In a similar vein, the dynamics of Cuba's life expectancy at birth prior to 1961 is best reproduced by the convex combination of Paraguay (69%) and Uruguay (30%). Finally, the synthetic version of Cuba's average years of schooling trajectory consists of Costa Rica (44%), Panama (37%), and Dominican Republic (19%).

The synthetic control analysis presented in Table 1 demonstrates good overall balance across key predictors and outcomes for each variable analyzed. Particularly noteworthy is the strong alignment observed in pre-treatment outcomes between Cuba and its synthetic counterpart for the infant mortality, life expectancy, and school enrollment indicators.

Table 1: Synthetic Control Weights and Balance by Outcome

Outcome	Variable/Country	Cuba	Synthetic Cuba	Donor Sample / Weight
Infant Mortality	**Weights**			
	El Salvador			0.62
	United States of America			0.27
	Venezuela			0.11
	Balance			
	Infant Mortality (t-1)	4.78	5.13	5.14
	Infant Mortality (t-2)	4.82	5.14	5.16
	Infant Mortality (t-3)	4.85	5.16	5.18
	Infant Mortality (t-4)	4.88	5.17	5.20
	Absolute Latitude	21.63	21.62	22.92
	Population Density	0.04	0.05	0.04
	% European Descent	63.00	57.46	53.40
	GDP per capita	2250.31	4844.00	4703.94
	Infant Mortality	5.73	5.71	5.66
	Longitude	-79.02	-92.87	-83.79
	Near Coast (Dummy)	100.00	68.48	40.35
	Soil Quality Index	45.95	54.61	38.42
	% Tropical Area	100.00	72.82	49.15
Life Expectancy	**Weights**			
	Paraguay			0.69
	Uruguay			0.30
	Balance			
	Life Expectancy (t-1)	62.08	61.22	56.36
	Life Expectancy (t-2)	60.11	60.50	53.86
	Life Expectancy (t-4)	45.00	46.71	44.97
	Absolute Latitude	21.63	25.86	22.50
	Population Density	0.04	0.00	0.01
	% European Descent	63.00	64.45	60.14
	GDP per capita (log)	7.73	8.01	8.16
	Life Expectancy	46.06	45.92	43.75
	Longitude	-79.02	-57.88	-73.40
	Near Coast (Dummy)	100.00	11.58	29.28
	Soil Quality Index	45.95	40.71	34.89
	% Tropical Area	100.00	40.50	47.21
Education	**Weights**			
	Costa Rica			0.44
	Panama			0.37
	Dominican Republic			0.19
	Balance			
	School Enrollment (t-1)	4.08	4.04	3.82
	School Enrollment (t-2)	4.00	3.98	3.76
	School Enrollment (t-3)	3.91	3.93	3.70
	School Enrollment (t-4)	3.83	3.87	3.64
	Absolute Latitude	21.63	11.33	20.36
	Population Density	0.04	0.02	0.02
	% European Descent	63.00	53.35	52.68
	GDP per capita (log)	7.55	7.67	7.79
	Longitude	-79.02	-80.15	-77.13
	Near Coast (Dummy)	100.00	99.57	45.60
	Population in 1400	95479.00	65576.68	627850.25
	School Enrollment	1.90	1.91	1.93
	Soil Quality Index	45.95	46.16	37.31
	% Tropical Area	100.00	94.66	61.54

While minor discrepancies exist, such as the GDP per capita for the infant mortality outcome and some geographical variables (longitude, percent tropical area, and proximity to the coast), these variations are relatively modest. To further strengthen our analysis and address any remaining balance concerns, we will employ robust methods such as the augmented Synthetic Control Method and Synthetic Difference-in-Differences. These methods are specifically designed to mitigate potential imbalances and enhance the credibility and robustness of the estimated treatment effects.

5.2 Main Results

Table 2 summarizes our main findings. The synthetic control method reveals heterogeneous effects across outcomes. For infant mortality, we find an average post-intervention reduction of 0.84 deaths per 1,000 live births (15%), with the effect strengthening from 0.58 deaths per 1,000 in 1970 to 1.10 deaths per 1,000 by 1980, before stabilizing around 0.83-0.93 deaths per 1,000 from 2000 onward. Life expectancy shows a more modest average improvement of 2.93 years (6%), peaking at 6.22 years in 1980 before declining to 2.83 years by 2020. Education exhibits the most dramatic effect, with an average increase of 1.84 years of schooling (96%), rising from 0.92 years in 1970 to a peak of 2.20 years by 2010, then slightly moderating to 2.01 years in 2020.

Table 2: Summary of Treatment Effects

	Outcome Variable		
	Infant Mortality	Life Expectancy	Years of Education
	(1)	(2)	(3)
Average Effect (Post-Treatment)	-0.84	+2.93	+1.84
	(-15%)	(+6%)	(+96%)
Effect vs. Last Pre-Treatment	+17%	+4.7%	+45%
Dynamic Effects by Decade:			
1970	-0.583	+3.85	+0.92
1980	-1.096	+6.22	+1.56
1990	-1.038	+4.50	+1.84
2000	-0.923	+4.82	+1.62
2010	-0.863	+4.26	+2.20
2020	-0.827	+2.83	+2.01

5.2.1 Infant Mortality

Figure 1 presents the main comparison between Cuba's observed infant mortality rate and the synthetic control constructed from the donor pool of Latin American countries and the United States. The synthetic control closely tracks Cuba's pre-intervention trajectory from 1900 until the intervention point (indicated by the dashed vertical line), demonstrating good pre-treatment fit. Post-intervention, Cuba's infant mortality rate diverges below the synthetic control, suggesting a beneficial treatment effect with an average reduction of 0.84 infant deaths per 1,000 live births, representing a 15% improvement.

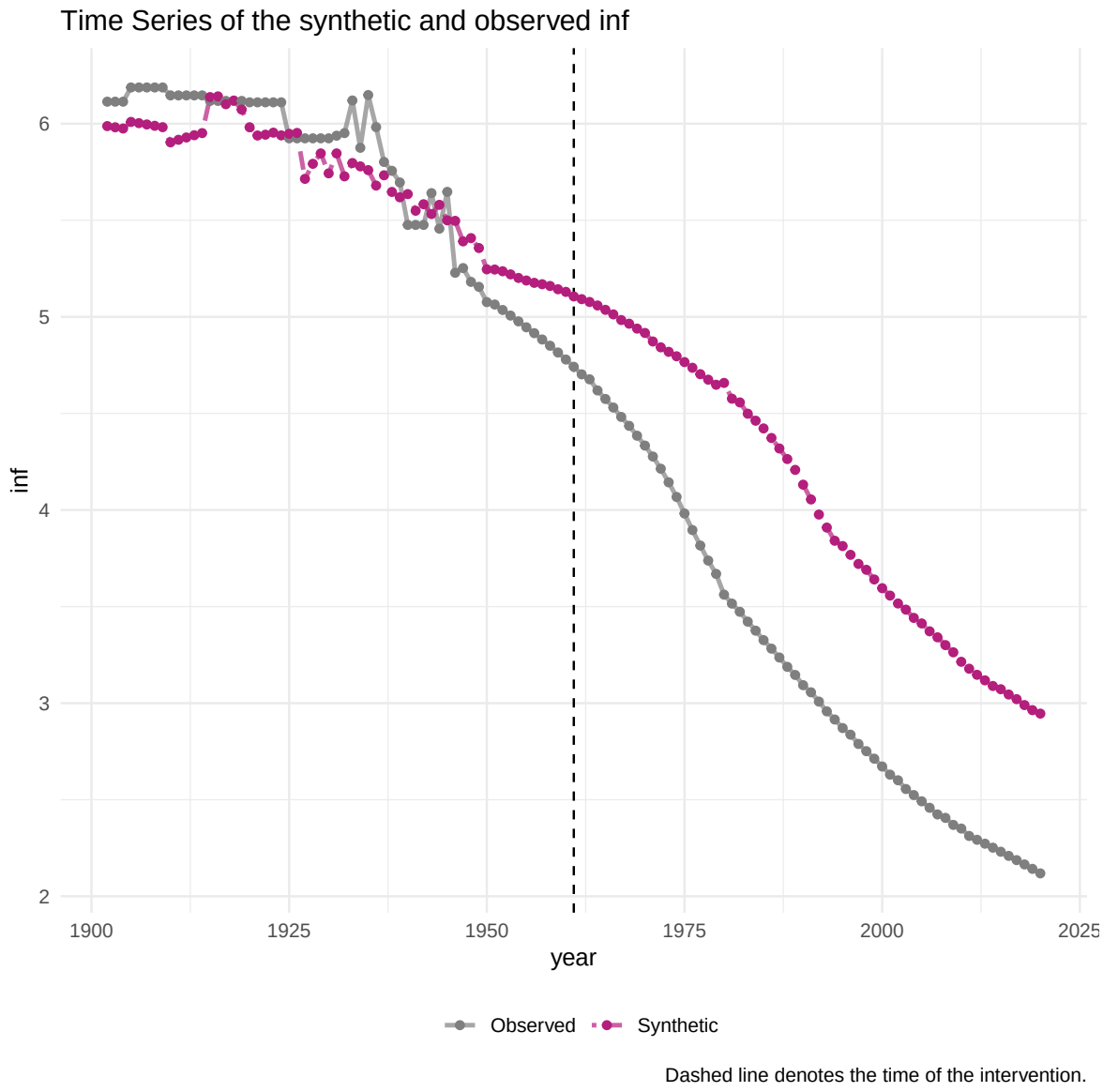


Figure 1: Cuba vs Synthetic Cuba Infant Mortality Trajectories.

5.2.2 Life Expectancy

Figure 2 presents the primary comparison between Cuba's observed life expectancy and its synthetic counterpart. The pre-intervention period shows excellent fit, with the synthetic control closely tracking Cuba's life expectancy trajectory from 1900 until the intervention. Post-intervention, Cuba's life expectancy exceeds the synthetic control, suggesting an average treatment effect of +2.93 years, representing a 6% improvement.

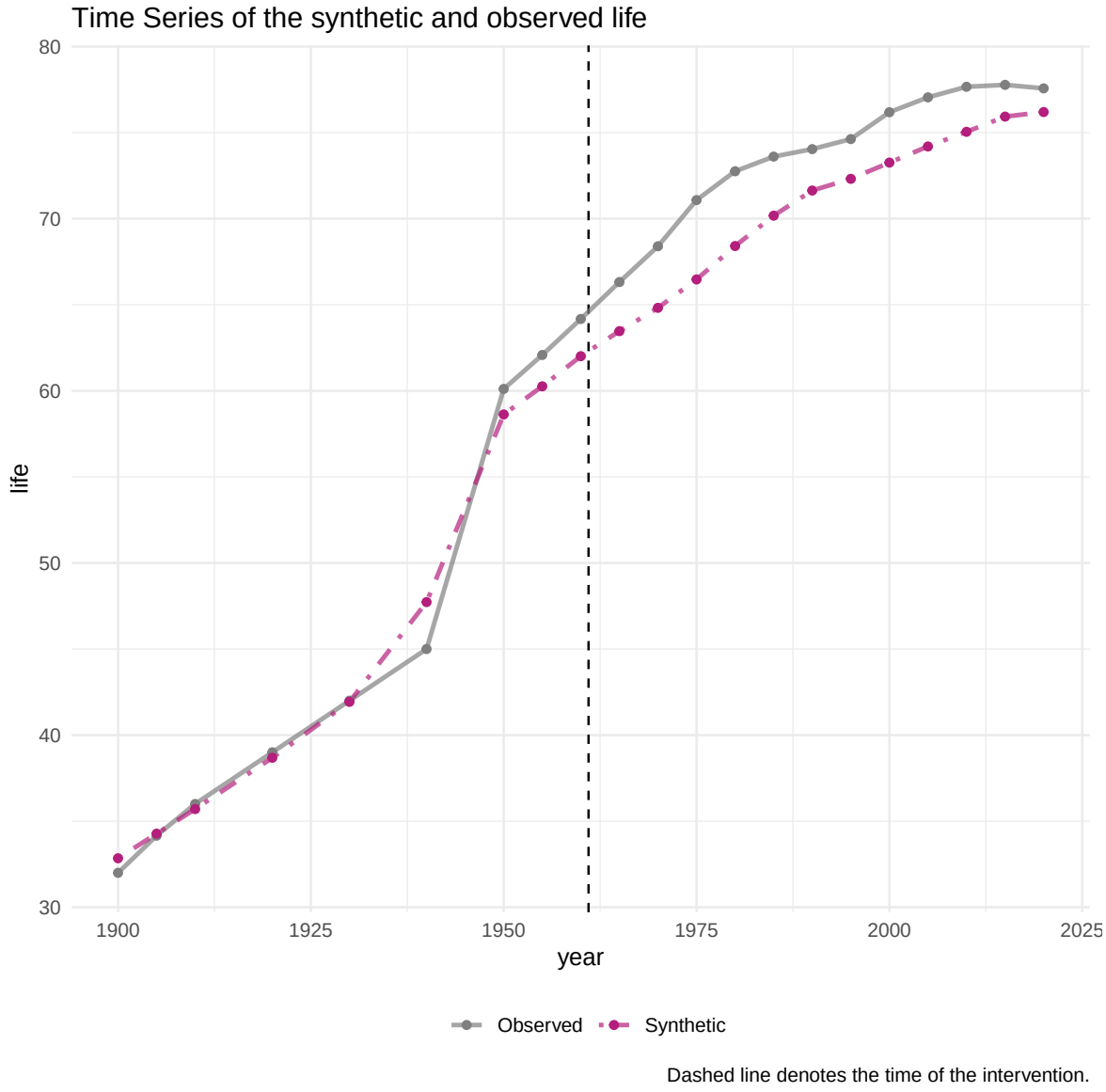


Figure 2: Cuba vs Synthetic Cuba Life Expectancy Trajectories.

5.2.3 Education

Figure 3 displays the comparison between Cuba's observed average years of education and the synthetic control. The pre-intervention fit is excellent, with the synthetic control precisely tracking Cuba's educational trajectory from 1900 until the intervention. Post-intervention, Cuba's educational attainment diverges sharply upward from the synthetic control, with an average treatment effect of +1.84 years of schooling—nearly doubling (96% increase) what would have been expected in the absence of intervention. This represents the most substantial proportional effect among all outcomes analyzed.

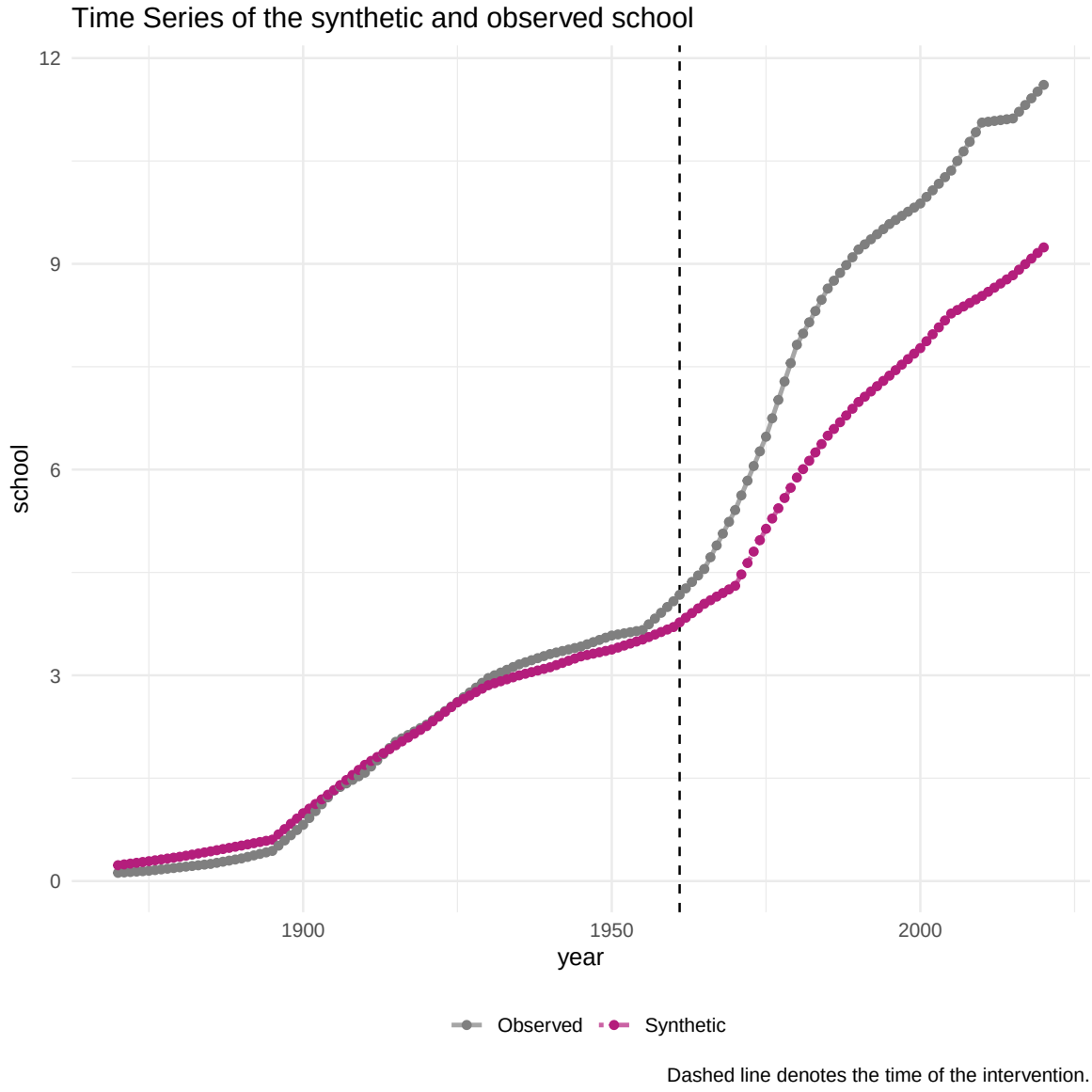
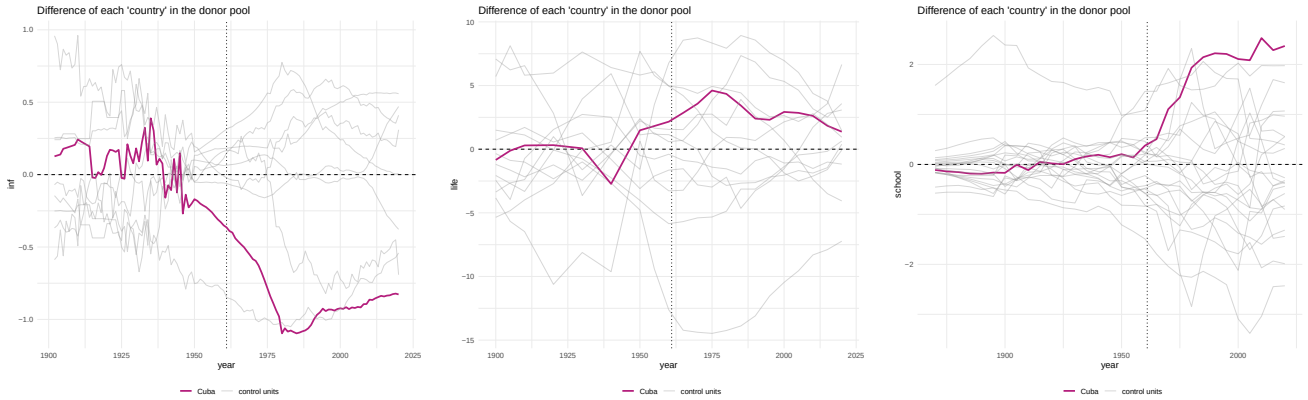


Figure 3: Cuba vs Synthetic Cuba Average Years of Education Trajectories.

5.3 Inference

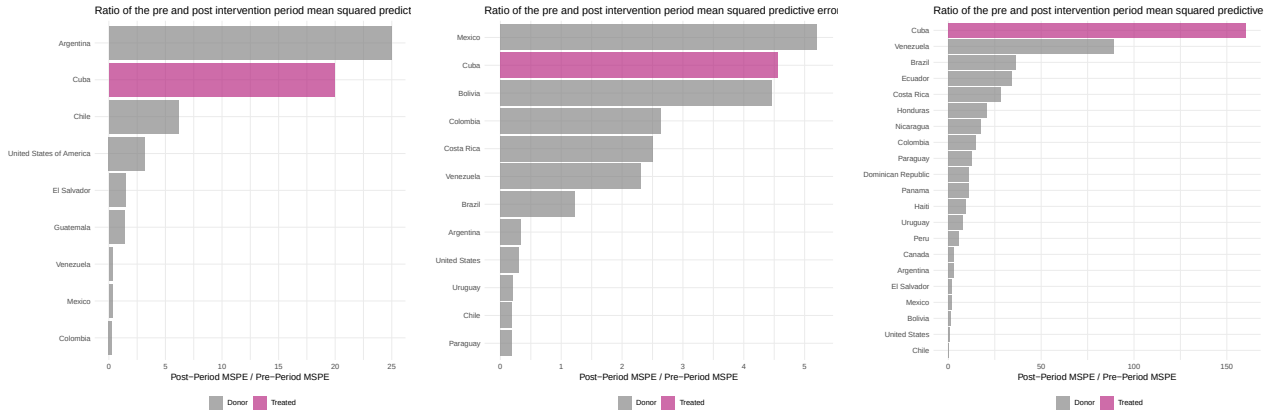
We assess statistical significance through placebo tests in space. For each donor country, we apply the synthetic control method as if it had received treatment, then compare the distribution of placebo effects to Cuba's actual effect.



(a) Infant Mortality (b) Life Expectancy (c) Education

Figure 4: Placebo Tests in Space

Figure 4 reports the placebo trajectories and Figure 5 reports the ratio of post- to pre-intervention mean squared prediction error (MSPE) for all countries. For infant mortality and education, Cuba ranks among the highest ratios, suggesting effects unlikely to occur by chance (p ¡ 0.10). For life expectancy, Cuba's ratio falls in the middle of the distribution, indicating weaker statistical evidence.



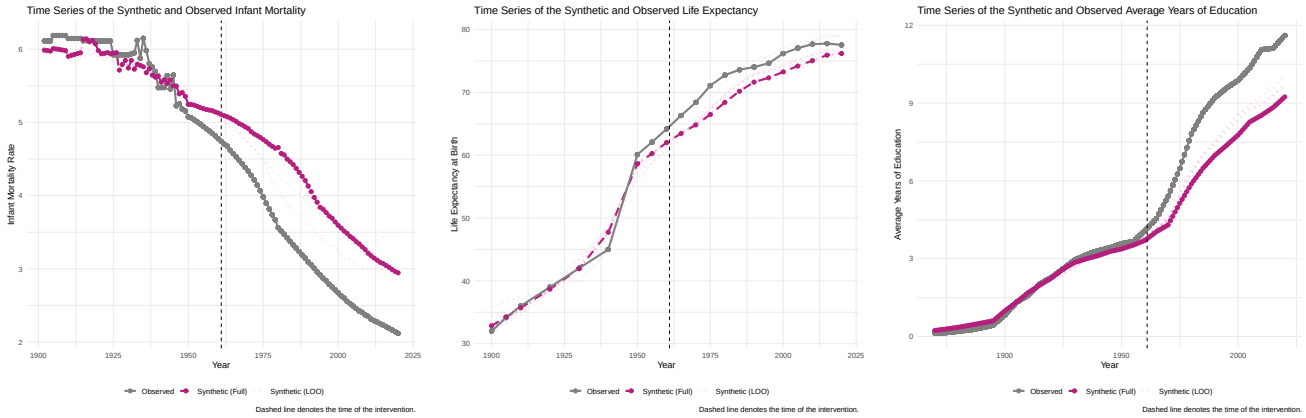
(a) Infant Mortality (b) Life Expectancy (c) Education

Figure 5: MSPE Ratios: Post/Pre-Intervention Period

5.4 Robustness Checks

5.4.1 Leave-One-Out Analysis

To assess sensitivity to donor pool composition, we iteratively exclude each donor country and re-estimate the synthetic control. The results, reported in Figure 6, remain stable across specifications. Notably, excluding the United States—despite its substantial economic differences from Latin American countries—does not materially alter our estimates, confirming that results reflect regional patterns rather than comparison with a developed economy.

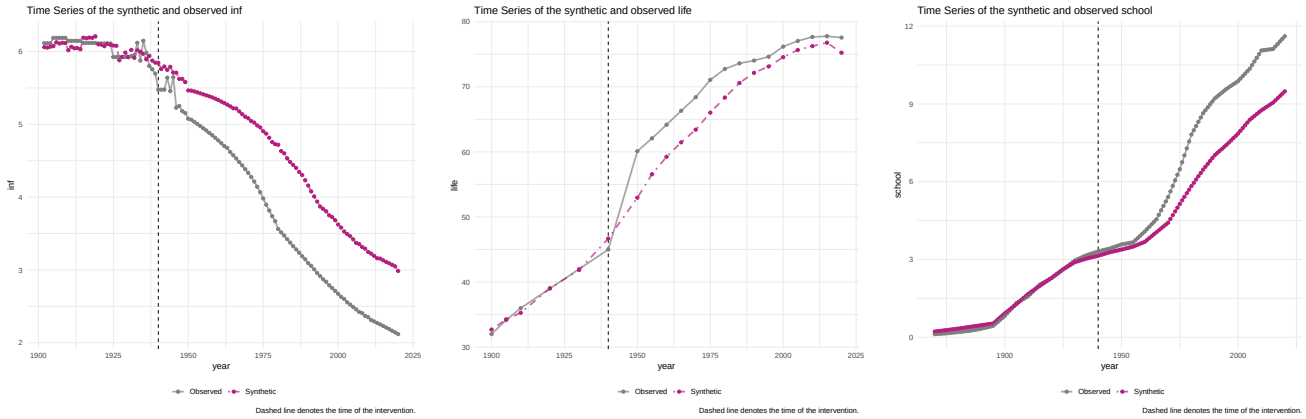


(a) Infant Mortality (b) Life Expectancy (c) Education

Figure 6: Leave-One-Out Analysis

5.4.2 Alternative Intervention Timing

As a placebo test in time, we re-estimate the model using 1940 as a hypothetical intervention date. The results are reported in Figure 7. The synthetic control maintains reasonable pre-1940 fit, with no significant divergence between observed and synthetic trajectories before the actual intervention, supporting our main specification.



(a) Infant Mortality (b) Life Expectancy (c) Education

Figure 7: Backdating Analysis: 1940 as Alternative Treatment Date

5.5 Alternative Estimators

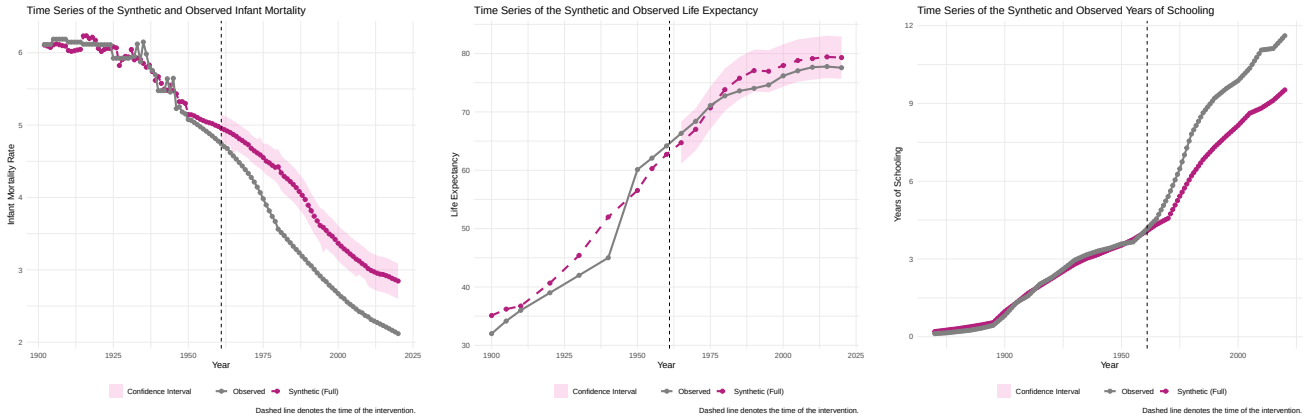
5.5.1 Augmented Synthetic Control Method

Table 3 and Figure 8 present results using the augmented synthetic control method (Ben-Michael et al., 2021), which relaxes the convexity constraint and provides formal confidence intervals. The augmented SCM yields a slightly smaller but still substantial effect for infant mortality (-0.64 deaths per 1,000; 11% reduction), with confidence intervals excluding zero for most of the post-period. Life expectancy shows a negative effect (-1.14 years; -2.4%), contrasting with standard SCM results. Education results remain robust (+1.51 years; 80% increase) with confidence intervals consistently excluding zero.

Table 3: Augmented Synthetic Control Results

	Infant Mortality (1)	Life Expectancy (2)	Years of Education (3)
Augmented SCM Estimate	-0.64	-1.14	+1.51
	(-11%)	(-2.4%)	(+80%)
95% Confidence Interval	[-0.95, -0.33]	[-3.42, 1.14]	[1.12, 1.90]

Note: Confidence intervals constructed following Chernozhukov and Wüthrich (2021).



(a) Infant Mortality (b) Life Expectancy (c) Education

Figure 8: Augmented Synthetic Control Method Results

Note: Confidence intervals are constructed using Chernozhukov et al. (2021)

5.5.2 Synthetic Difference-in-Differences

Table 3 reports synthetic difference-in-differences estimates (Arkhangelsky et al., 2021), which combine features of both synthetic control and difference-in-differences methods. The SDID estimates confirm significant reductions in infant mortality (18.1%, p = 0.081) and improvements in education (1.53 years, p = 0.003), while life expectancy effects remain statistically insignificant (1.24 years, p = 0.770).

Table 4: Synthetic Difference-in-Differences Results

	Infant Mortality (1)	**Life Expectancy** (2)	**Years of Education** (3)
SDID Estimate	-0.181	+1.248	+1.534
Standard Error	(0.104)	(4.277)	(0.522)
95% Confidence Interval	[-0.386, 0.022]	[-9.361, 7.134]	[0.509, 2.558]
P-value	0.081	0.770	0.003

Note: Treatment effects computed for 1870-2022. Standard errors adjusted for serial correlation and heteroskedasticity using 1,000 bootstrap iterations.

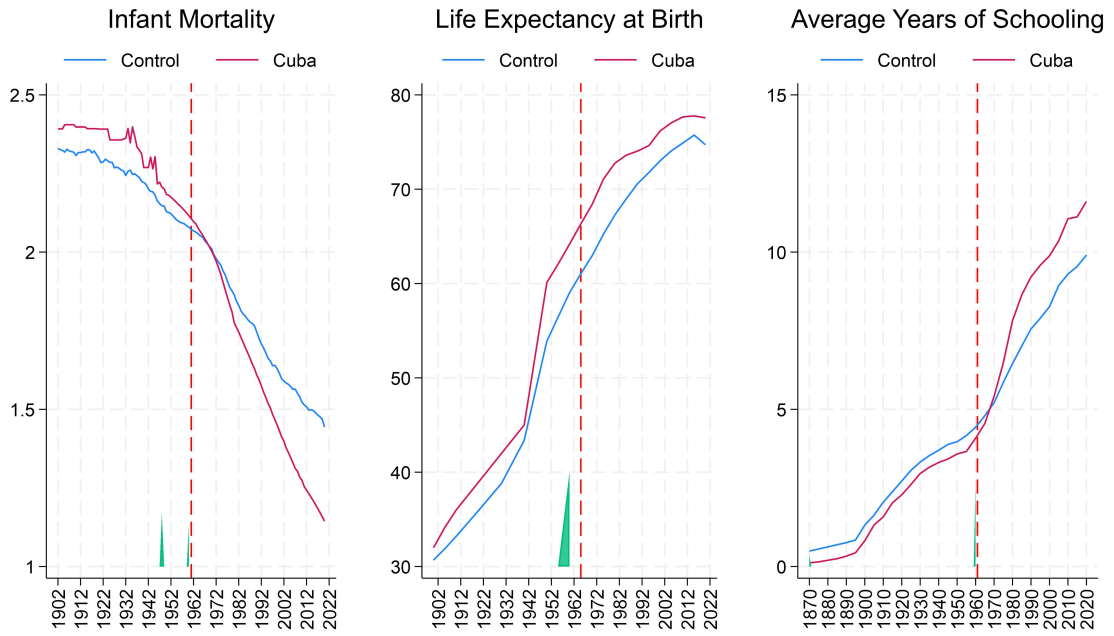


Figure 9: Synthetic Difference-in-Differences Treatment Effects Over Time

5.5.3 Interactive Fixed Effects and Matrix Completion

Table 5 presents results using interactive fixed effects (Xu, 2017) and matrix completion methods (Athey et al., 2021). Both approaches confirm significant reductions in infant mortality (25-29

Table 5: Alternative Estimators: Interactive Fixed Effects and Matrix Completion

	Infant Mortality (1)	**Life Expectancy** (2)	**Years of Education** (3)
Panel A: Interactive Fixed Effects			
Estimate	-0.247	+1.390	+2.006
Standard Error	(0.066)	(3.056)	(0.357)
P-value	0.000	0.681	0.000
Panel B: Matrix Completion ($\lambda = 0.003$)			
Estimate	-0.294	+1.447	+1.532
Standard Error	(0.039)	(1.898)	(0.117)
P-value	0.000	0.446	0.000
Panel C: Matrix Completion ($\lambda = 0.01$)			
Estimate	-0.292	+1.145	+1.799
Standard Error	(0.040)	(1.360)	(0.307)
P-value	0.000	0.401	0.000

5.5.4 Comparison with U.S. States

To expand the donor pool and test external validity, we include U.S. states alongside Latin American countries. Due to data avalability this is only possible for infant mortality and life expectancy. Despite institutional differences, neither U.S. states nor Latin American countries implemented comparable universal healthcare systems, preserving the validity of our identification strategy. Using data from 1902-2022 for 44 U.S. states and 5 Latin American countries,

we find an average reduction in infant mortality of 22% (p = 0.021) and improvement in life expectancy of 4.8 years (p = 0.065). Figure 10 report the results.

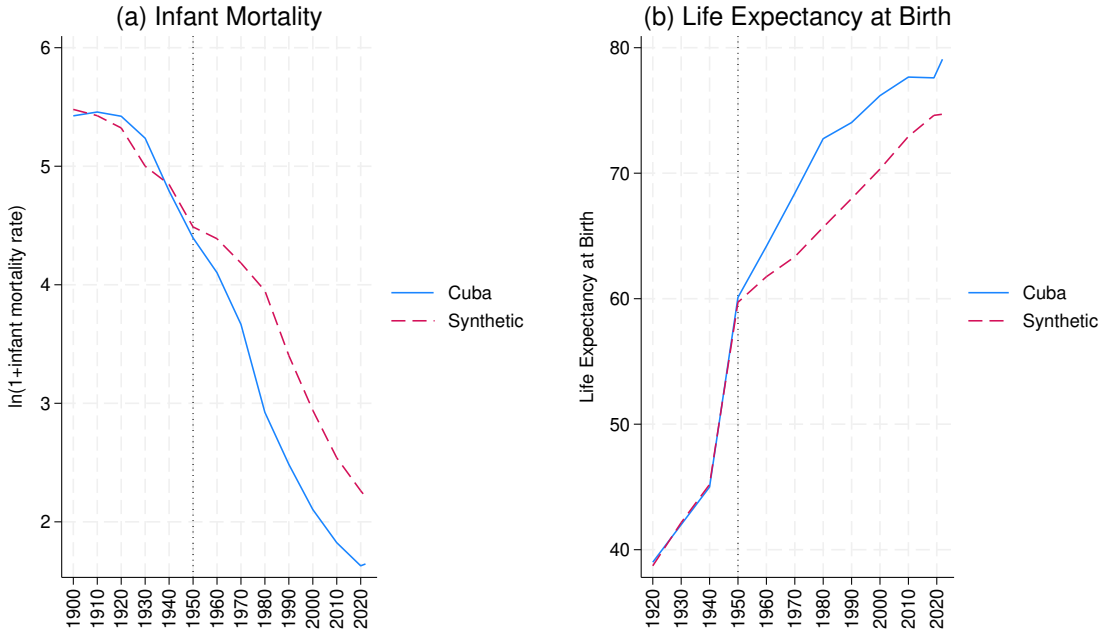


Figure 10: Long-term effect of the comprehensive national health care on life expectancy and human capital using a donor pool of US states and Latin American countries, 1902-2022

5.6 Summary

The evidence consistently points to substantial improvements in infant mortality and education following Cuba's healthcare intervention, with more modest and uncertain effects on life expectancy. The education results prove most robust across all specifications, while life expectancy effects show sensitivity to methodological choices. These patterns suggest that universal healthcare's impact may vary considerably across different health and human capital outcomes.

6 Conclusion

In this paper, we examine the contribution of universal health care to public health and human capital investment. To this end, we exploit the introduction of the comprehensive and large-scale universal health care in Cuba in 1961 as a quasi-experimental source of variation in health and human capital outcomes. Utilizing novel historical trajectories of infant mortality, life expectancy at birth and average years of education across a sample of 12 former European colonies in the Americas for the period 1870-2020, we compare the outcomes in Cuba with the donor pool of countries with such audacious and large-scale public health reform, and estimate the missing counterfactual scenario to evaluate the long-term effect of the universal health on the immediate outcomes (i.e. infant mortality and life expectancy) as well as the spillover effect on human capital formation.

By making use of the plain vanilla and more advanced variants of the synthetic control method, we find evidence of large-scale improvement in public health outcomes and human capital investment vis-á-vis the synthetic counterfactual. We show that former European colonies

in the Americas, including the United States, provide an excellent and almost noise-free quality of the fit in the underlying trajectories in the pre-intervention period. Our results suggest that Cuba's infant mortality decreased by around 35 percent in response to the introduction of the universal health care. The estimated effect is both large, statistically significant at conventional level, and appears to be permanent. We also find evidence of the large but somewhat more temporary effect of the universal health care on life expectancy at birth. By contrast, the estimated effect of the universal health care on average years of schooling is large, statistically significant and also appears to be permanent with the increasing effect magnitude over time.

The estimated magnitude of improvement in public health and human capital formation appears to be robust to distinctive variants and alternations of the identification strategy. For instance, the infant mortality reduction estimated using matrix completions, interactive fixed-effects, synthetic difference-in-differences, and augmented synthetic control estimators appears to be permanent and comparable in terms of size and significance to our baseline estimates. By contrast, our estimates imply that the improvement of life expectancy vis-á-vis the synthetic counterfactual seems to be somewhat more temporary. On the contrary, the improvement of schooling trajectory is both large and permanent and tends to increase substantially over time.

Our results show that the introduction of universal health care with complete coverage in 1961 led to widespread and encompassing improvement of both public health and human capital formation. The improvement has been particularly strongly concentrated on the reduction of infant mortality and average years of education. Elevated life expectancy trajectory appears to be somewhat less permanent and tends to weaken in the beginning of the 1980s in the midst of the collapse of both Soviet Union and world sugar prices. Despite these difficulties, Cuba's life expectancy, infant mortality and human capital formation trajectories up to the present day still outperform its synthetic American counterfactual.

References

Abadie, A., A. Diamond, and J. Hainmueller (2010): "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program," *Journal of the American Statistical Association*, 105, 493–505.

——— (2015): "Comparative politics and the synthetic control method," *American Journal of Political Science*, 59, 495–510.

Abadie, A. and J. Gardeazabal (2003): "The economic costs of conflict: A case study of the Basque Country," *American Economic Review*, 93, 113–132.

Arkhangelsky, D., S. Athey, D. A. Hirshberg, G. W. Imbens, and S. Wager (2021): "Synthetic difference in differences," *American Economic Review*, 111, 4088–4118.

Athey, S., M. Bayati, N. Doudchenko, G. Imbens, and K. Khosravi (2021): "Matrix completion methods for causal panel data models," *Journal of the American Statistical Association*, 116, 1716–1730.

Ben-Michael, E., A. Feller, and J. Rothstein (2021): "The augmented synthetic control method," *Journal of the American Statistical Association*, 116, 1789–1803.

Bleakley, H. (2010): "Health, human capital, and development," *Annual Review of Economics*, 2, 283–310.

Chernozhukov, V., K. Wüthrich, and Y. Zhu (2021): "An Exact and Robust Conformal Inference Method for Counterfactual and Synthetic Controls," *Journal of the American Statistical Association*, 116, 1849–1864.

Feinsilver, J. M. (2023): *Healing the masses: Cuban health politics at home and abroad*, Irvine, CA: University of California Press.

Grossman, M. (1972): "On the concept of health capital and the demand for health," *Journal of Political Economy*, 80, 223–255.

Keck, C. W. and G. A. Reed (2012): "The curious case of Cuba," *American Journal of Public Health*, 102, e13–e22.

Nikelly, A. G. (1988): "Health care in Cuba," *Public Health*, 102, 19–25.

Xu, Y. (2017): "Generalized synthetic control method: Causal inference with interactive fixed effects models," *Political Analysis*, 25, 57–76.