University of Pavia

Master's Thesis

Independence and Stature

– The Anthropometric Cost of Institutional Rupture in Africa and Asia –

Indipendenza e Statura

– Il Costo Antropometrico della Rottura Istituzionale in Africa e Asia –

Author:
Jeffrey Balsamo

Supervisor:
Prof. Maria Anna Leone

Student ID:
581237

A thesis submitted in fulfillment of the requirements for the

Master Programme in Economics, Development and Innovation

in the

Department of Economics and Management

Academic Year 2025 – 2026

Acknowledgements

This thesis marks a significant milestone in a long and exciting academic journey—one that is by no means a solitary achievement.

I would like to express my sincere gratitude to Prof. Maria Anna Leone for her valuable feedback and support throughout this project; her lectures on Difference-in-Differences methodologies inspired the empirical design of this work. I am also grateful to Prof. Jörg Baten, whose seminar first sparked my interest in using historical height data as a window into long-run economic development.

On a personal note, I would like to thank my family and friends. A special thanks goes to Johannes Bertram for his helpful feedback on parts of this thesis, and to Jonas Weitbrecht for the sports and cooking sessions that offered a perfect balance to the hours of writing. I am also deeply grateful to Hafez Kerim Aslan and Samuel Waidmann for that late-night discussion at Veggi Box about everything under the sun—which unexpectedly helped sharpen the institutional argument of this work. I am equally thankful to my sister, Alessandra Balsamo, and her partner, Justin Fulmer, for their thesis design feedback and for letting me use their living room as a research laboratory.

Finally, I owe an immense debt to my parents, Gyöngyi and Michele Balsamo. The core argument of this thesis is that past actions profoundly shape present outcomes—a claim that extends well beyond the realm of institutional economics. Without their foundational investments, sacrifices, and belief, I would never have reached this point. I am deeply grateful for their support throughout my entire journey.

Abstract

This study provides the first empirical evidence on the health consequences of decolonization by answering the following question:

What is the causal impact of attaining independence from colonial rule on human stature, and can this effect be explained by the nature of the institutional transition?

To this end, a theoretical framework is proposed that conceptualizes independence as a shock to an existing institutional equilibrium, where the severity of the shock depends fundamentally on the fragility of that equilibrium and the planning of the transition process. It is hypothesized that more severe institutional ruptures precipitate correspondingly larger negative impacts on terminal adult height—a widely used proxy for population health and net nutritional status across the life course.

To test this framework empirically, a Synthetic Difference-in-Differences (SDiD) event-study approach is applied to a macro-panel of adult male stature spanning 1850 to 1980. The analysis first compares 14 extractive colonies—characterized by fragile equilibria and profound transitional shocks—against two settler colonies, where independence merely formalized a stable, pre-existing equilibrium. The SDiD estimates confirm that in extractive colonies, independence caused a statistically significant, long-run collapse in human stature, yielding an average treatment effect on the treated ($\widehat{ATT}$) of -3.70 cm. In contrast, settler colonies experienced no such anthropometric decline.

To further isolate the institutional mechanism, within-group heterogeneity is evaluated across the extractive cohort by disaggregating the sample into planned and unplanned transfers of power. Consistent with the theoretical predictions, abrupt, unplanned transitions generated a more severe shock than planned, coordinated handovers. Collectively, these findings suggest that the health consequences of decolonization were far from uniform and depended fundamentally on the severity of the institutional rupture accompanying the transition to sovereignty.

Sintesi

Questo studio fornisce la prima evidenza empirica delle conseguenze sulla salute della decolonizzazione, rispondendo alla seguente domanda:

Qual è l'impatto causale dell'ottenimento dell'indipendenza dal dominio coloniale sulla statura umana, e tale effetto può essere spiegato dalla natura della transizione istituzionale?

A tal fine, viene proposto un quadro teorico che concettualizza l'indipendenza come uno shock a un equilibrio istituzionale esistente, in cui la gravità dello shock dipende fondamentalmente dalla fragilità di tale equilibrio e dalla pianificazione del processo di transizione. Si ipotizza che rotture istituzionali più gravi determinino impatti negativi proporzionalmente maggiori sull'altezza finale degli adulti—una proxy ampiamente utilizzata per la salute della popolazione e lo stato nutrizionale netto lungo l'intero arco della vita.

Per verificare empiricamente questo quadro, viene applicato un approccio di studio degli eventi basato sulla differenza sintetica nelle differenze (SDiD) a un macro-panel sulla statura degli uomini adulti che copre il periodo dal 1850 al 1980. L'analisi confronta innanzitutto 14 colonie estrattive—caratterizzate da equilibri fragili e profondi shock di transizione—con due colonie di insediamento, in cui l'indipendenza ha semplicemente formalizzato un equilibrio stabile e preesistente. Le stime SDiD confermano che nelle colonie estrattive l'indipendenza ha causato un crollo statisticamente significativo e di lungo periodo della statura umana, producendo un effetto medio del trattamento sui trattati ($\widehat{ATT}$) pari a -3,70 cm. Al contrario, le colonie di insediamento non hanno registrato alcun declino antropometrico di questo tipo.

Per isolare ulteriormente il meccanismo istituzionale, l'eterogeneità intragruppo viene valutata all'interno della coorte estrattiva, disaggregando il campione in trasferimenti di potere pianificati e non pianificati. In linea con le previsioni teoriche, le transizioni brusche e non pianificate hanno generato uno shock antropometrico più grave rispetto ai passaggi di potere pianificati e coordinati. Nel loro insieme, questi risultati suggeriscono che le conseguenze sulla salute della decolonizzazione sono state tutt'altro che uniformi e dipendevano fondamentalmente dalla gravità della rottura istituzionale che ha accompagnato la transizione alla sovranità.

Contents

List of Figures vi

List of Tables vii

1 Introduction 1

2 Theoretical Framework 7
- 2.1 Hypothesis 1: Extractive Colonies . . . 10
- 2.2 Hypothesis 2: Settler Colonies . . . 15
- 2.3 Hypothesis 3: Planned vs. Unplanned Transitions . . . 17

3 Variables and Data 19
- 3.1 Human Stature . . . 19
- 3.2 Colonial Status . . . 21
- 3.3 Independence . . . 22
  - 3.3.1 Data Harmonization . . . 22
  - 3.3.2 Independence Date Selection . . . 23
- 3.4 Extractive and Settler Colonies . . . 26
- 3.5 Planned and Unplanned Transitions . . . 27

4 Empirical Strategy 28
- 4.1 Identification Challenges . . . 28
- 4.2 Synthetic Difference-in-Differences . . . 30
  - 4.2.1 Econometric Specifications . . . 31
  - 4.2.2 Event-Study Extension . . . 33
- 4.3 Implications for Model Specification . . . 37
  - 4.3.1 Covariates . . . 37
  - 4.3.2 Heterogeneity Analysis . . . 38
- 4.4 Panel Construction . . . 38
  - 4.4.1 Temporal Boundaries . . . 39
  - 4.4.2 Unit-Level Inclusion Criteria . . . 41
  - 4.4.3 Donor Pool . . . 43

5 Main Results 45
- 5.1 Hypothesis 1: Extractive Colonies . . . 45
  - 5.1.1 SDiD Weights . . . 46
  - 5.1.2 Event Study Results . . . 49
- 5.2 Hypothesis 2: Settler Colonies . . . 52
  - 5.2.1 Event Study Results . . . 53
- 5.3 Hypothesis 3: Planned vs. Unplanned Transitions . . . 57
  - 5.3.1 Regional Comparison: Africa vs. Asia . . . 57

5.3.2 Transition Types within Africa 61

6 Conclusion **65**
6.1 Summary of Findings 65
6.2 Limitations 66
6.3 Avenues for Future Research 67
6.4 Policy Implications and Discussion 69

Bibliography **70**

A Data Properties and Baseline Calibration **81**
A.1 Master Sample 81
A.2 Descriptive Statistics 83
A.3 Supplementary SDiD Weight Distributions 84

B Robustness Checks **87**
B.1 Treatment Timing Sensitivity Check 87
B.2 Unit-by-Unit Analysis 88
B.3 Leave-One-Out Sensitivity Check 93
B.4 Sample Composition Sensitivity Check 97
B.5 In-Space Placebo Check 100
B.6 Error Structure and Inference 101

C Identification Assumptions Checks **105**
C.1 Treatment Anticipation 105
C.2 Confounding at Time of Treatment 107
C.3 Cross-Border Spillovers 108

D Econometric Specifications **111**
D.1 Formalization Weight Optimization 111
D.1.1 Unit Weights 111
D.1.2 Time Weights 114
D.2 Formalization Latent Factor Model 116

List of Figures

5.1 Unadjusted Anthropometric Trajectories: Extractive Cohort vs. Donor Pool 46
5.2 SDiD Event-Study Estimates: Extractive Cohort 50
5.3 Unadjusted Anthropometric Trajectories: Settler Cohort vs. Donor Pool 52
5.4 SDiD Event-Study Estimates: Settler Cohort 54
5.5 Distribution of Transition Types: Africa vs. Asia 57
5.6 SDiD Event-Study Estimates: Africa vs. Asia 59
5.7 SDiD Event-Study Estimates: Planned vs. Unplanned Transitions 62

B.1 Sensitivity to Treatment Timing: Mean vs. Max Date 87
B.2 Country-Specific $\widehat{ATT}$ Estimates: Mean Date 91
B.3 Country-Specific $\widehat{ATT}$ Estimates: Max Date 92
B.4 Leave-One-Out Estimates: Extractive Cohort 94
B.5 Leave-One-Out Estimates: Africa vs. Asia 95
B.6 Leave-One-Out Estimates: Planned vs. Unplanned Transitions 96
B.7 SDiD Event-Study Estimates: 1940 Cohort Excluded 98
B.8 Sensitivity to Sample Composition: 1940 Cohort Included vs. Excluded 99
B.9 Homoscedasticity Check of Residuals 104

List of Tables

4.1 Estimation Sample: Treated Units by Colony Type 42
4.2 Estimation Sample: Donor Pool and Matching Eligibility 44

5.1 Extractive Cohort: SDiD Unit Weights 48
5.2 Extractive Cohort: SDiD Time Weights 48
5.3 SDiD Results: Extractive Cohort — Average and Immediate Treatment Effects 51
5.4 SDiD Results: Settler Cohort — Average and Immediate Treatment Effects 55
5.5 SDiD Results Comparison: Africa vs. Asia 60
5.6 SDiD Results: Planned vs. Unplanned Transitions 63

A.1 Overview: Sample Composition, Independence Timing, and Transition Mechanisms . . 81
A.2 Summary Statistics: Treated Cohorts and Global Control Pools 83
A.3 Settler Cohort: SDiD Unit and Time Weights 85
A.4 Planned vs. Unplanned Transitions: SDiD Unit and Time Weights 86

1 Introduction

The past is never dead. It's not even past.

– William Faulkner, 1951

When Faulkner penned these words in *Requiem for a Nun*, he captured a profound truth that lies at the very core of modern political economy: deep historical events do not merely remain confined to the archives, but continue to exert a radiating influence over contemporary outcomes. A striking real-world manifestation of this historical persistence is colonialism. As established by a rich macro-historical literature, the institutional architectures imposed under colonial rule continue to shape the long-run economic trajectories of formerly colonized nations, thereby offering an explanation for why certain countries perform structurally better than others (Acemoglu et al., 2002, 2005; Engerman & Sokoloff, 2002). In doing so, this imperial legacy underscores one potential historical cause for the persistent economic inequality that separates developing nations from developed nations (Milanovic, 2016; Pomeranz, 2000; Pritchett, 1997). However, these cross-national disparities dictate more than just wealth differences; they influence global power asymmetries (Rodrik, 2011; Wallerstein, 2011), the capacity of states to insulate their populations from systemic shocks (Besley & Persson, 2011), and international migration pressures (Massey et al., 1993). Consequently, such historical explanations are essential not only for gaining deeper insights into the mechanics of global economic systems, but also for addressing the structural imbalances that define the modern international order.

This paper aims to further develop the understanding of these cross-national disparities by analyzing the following research question:

What is the causal impact of attaining independence from colonial rule on human stature, and can this effect be explained by the nature of the institutional transition?

To evaluate this question, the empirical strategy is divided into three sub-inquiries:

1) Did the transition to independence in extractive colonies trigger a collapse in human stature?

2) Did the presence of institutional continuity in settler colonies prevent or mitigate a similar biological penalty?

3) Within the extractive cohort, did structurally planned transitions lessen the severity of the biological shock compared to abrupt and disorderly ruptures?

Sub-inquiries 2 and 3 thus function as mechanism tests—they examine whether any effect identified in sub-inquiry 1 can be attributed to the nature of the institutional transition.

The analysis addressing these questions employs a Synthetic Difference-in-Differences (SDiD) event-study approach (Arkhangelsky et al., 2021; Ciccia, 2024; Porreca, 2022), where the timing of independence serves as the treatment event. The primary treatment group consists of 14 extractive colonies geographically clustered in Sub-Saharan Africa and Southeast Asia—which are further bifurcated into planned and unplanned transitions for the third sub-inquiry. A secondary treatment group comprises two settler colonies, while a set of 26 untreated sovereign nations constitutes the donor pool for constructing the synthetic counterfactuals. The underlying dataset is structured as a macro-panel tracking adult male stature from 1850 to 1980.

The SDiD estimates reveal that independence in extractive colonies caused a statistically significant, long-run collapse in human stature, yielding an overall average treatment effect on the treated (ATT) of −3.70 cm. Additionally, the two sub-analyses strongly suggest that this persistent penalty was driven by institutional rupture; while the settler cohort experienced no such decline, the within-group heterogeneity analysis of the extractive cohort indicates that abrupt, unplanned transitions induced a more severe shock than planned transfers of power.

This paper argues that the empirical findings can be theoretically conceptualized in the following way: In extractive colonies, independence acted not merely as a formal shift in legal sovereignty, but as a profound structural disruption to an existing institutional equilibrium. The resulting rupture frequently created an immediate power vacuum, triggering chaotic and often violently contested struggles to establish a new political baseline. This process, in turn, disrupted centralized resource allocation, public health networks, and the delivery of public goods. The collapse of such infrastructure degrades the local disease environment (Alsan & Goldin, 2019; Bleakley, 2007; Cutler & Miller, 2005), which subsequently inhibits nutrient absorption during childhood—a critical determinant of terminal adult stature (Baten & Blum, 2014; Deaton, 2007). Therefore, the initial institutional shock manifested as a persistent anthropometric penalty.

By analyzing these dynamics, this thesis contributes to several intersecting strands of literature, primarily spanning political economy, economic history, and anthropometric history.

The first contribution is theoretical. Classical institutional persistence studies, like the foundational colonial legacy framework pioneered by Acemoglu et al. (2001, 2002, 2005), either ignore the independence period entirely or treat it as a frictionless milestone during which sovereignty was seamlessly transferred. This study extends these frameworks by demonstrating that, in certain cases, the transition to sovereignty was an active, equilibrium-disrupting shock—with measurable and lasting consequences for human welfare. Consequently, it is argued that the process of decolonization must be explicitly incorporated into long-run development models.

Independence has been neglected not only theoretically, but also empirically. While the consequences of colonial rule have been examined extensively, the actual termination of imperial dominance has received comparatively limited scholarly attention (Lee & Paine, 2019; Ziltener & Künzler, 2017). Although a small body of literature has begun to examine the consequences of independence, these studies focus primarily on conventional economic outcomes, such as fiscal capacity (Dharmapala & Suesse, 2025), aggregate economic growth (Bertocchi & Canova, 2002; Lee & Paine, 2019), or international trade flows (Head et al., 2010). The anthropometric dimensions of independence, by contrast, have remained entirely unexamined. Thus, the empirical analysis of post-independence height effects constitutes the second contribution of this thesis.

Beyond these two contributions, this study engages directly with the global inequality debate discussed above through two distinct pathways.

First, an additional explanation for an observed divergence in global heights is offered. Baten and Blum (2012) demonstrate that global heights were relatively egalitarian until the 1880s; from that point forward—coinciding with the first wave of globalization and colonial expansion—a divergence occurred, wherein populations in the Global North grew taller at an accelerated rate relative to those in the Global South. Thus, the long-run divide between developed and developing nations manifested not only in macroeconomic terms but also in anthropometric outcomes. Given the timing of this divergence, an extensive literature has attempted to evaluate whether colonial rule itself was the root cause of these biological inequalities, yet this inquiry has yielded mixed and inconclusive empirical results (Baten & Maravall, 2021; Baten et al., 2013; Choi & Schwekendiek, 2009; Heyberger, 2022; Moradi et al., 2013). Currently, the prevailing view attributes this historical disparity primarily to the capacity of Western economies to implement investments in sanitation, public health networks, and urban infrastructure (Hatton & Bray, 2010; Steckel, 2009). However, this explanation accounts only for why Western populations grew taller—not for why developing nations failed to make equivalent investments and consequently fell behind. The persistent negative effect of independence identified in this study provides such a complementary explanation. It suggests that the institutional disruption triggered by decolonization prevented these public investments and thereby forestalled convergence toward the anthropometric baselines achieved by the developed world—at least through the 1980s, the limit of the empirical window employed in this study.

The second pathway through which this study contributes to the inequality debate is by explicitly examining a non-economic dimension of well-being. The vast majority of macro-historical studies investigating global disparities rely primarily on economic indicators, most notably GDP and measures of income or wealth distribution (e.g., Acemoglu et al., 2001, 2002, 2005; Chancel et al., 2022; Engerman & Sokoloff, 2002; Milanovic, 2016; Piketty & Goldhammer, 2014). However, as emphasized by Stiglitz et al. (2009), human well-being is inherently multidimensional, encompassing dimensions such as health, education, political participation, and personal security. Consequently, evaluating living standards solely through an economic lens fails to capture the full spectrum of human welfare. By utilizing human stature as the dependent variable—a metric widely accepted as a robust proxy for net nutritional status and population health (Baten & Blum, 2014; Deaton, 2007; Steckel, 1995)—this paper directly analyzes such a non-economic dimension.

The limitations of economic indicators become apparent when examining their relationship with these health outcomes. While a strong correlation exists between long-run economic development and health (Baten & Blum, 2012, 2014; Jones & Klenow, 2016; Steckel, 2009), the translation of material gains into health improvements is not always straightforward, and deviations between the two measures can be large (Jones & Klenow, 2016). For instance, Hatton and Bray (2010) observes that male stature in Western Europe increased most rapidly during specific historical intervals characterized by relatively slow economic growth, driven instead by coordinated public health interventions. This systemic decoupling becomes especially salient in the context of decolonization: while Bertocchi and Canova (2002) find that economic growth accelerated in African colonies post-independence, and Lee and Paine (2019) report a neutral growth effect, this study identifies a severe and persistent penalty in human stature. Relying exclusively on macroeconomic metrics would therefore have obscured this welfare cost entirely.

Finally, this paper makes a methodological contribution to the quantitative economic history literature. Although recent studies have deployed staggered Difference-in-Differences designs and Synthetic Control (SC) methods to capture post-independence effects (e.g., Dharmapala & Suesse, 2025), the SDiD estimator has not yet been applied to such a macro-historical, cross-country panel. This application serves as an empirical test of the utility and performance of SDiD within such a setting.

The remainder of this paper proceeds as follows. Chapter 2 develops the theoretical framework linking decolonization, institutional transition, and human stature. Chapter 3 introduces the data and variables employed in the analysis. Chapter 4 outlines the empirical strategy and describes the construction of the estimation sample. Chapter 5 presents the main empirical findings. Chapter 6 summarizes these findings, discusses limitations, and outlines implications for future research and policy.

2 Theoretical Framework

To understand the potential effect of independence on a nation's anthropometric trajectory, it is first necessary to examine the theoretical determinants of long-run development. Although the frameworks discussed below were primarily conceptualized to explain variations in macroeconomic growth, they serve as proxies for analyzing stature trajectories in this study. This approach is justified by the previously established strong correlation between aggregate economic output and average stature (Baten & Blum, 2012, 2014; Jones & Klenow, 2016; Steckel, 2009).

While scholars discuss various sources for differences in economic development, such as geography or culture, it is generally agreed that institutions are a fundamental driver of long-run economic outcomes (Acemoglu et al., 2005).

North (1990) defines institutions as "the rules of the game in a society or, more formally […] the humanly devised constraints that shape human interaction." By establishing these "rules of the game," institutions provide the necessary structure for human coexistence. They create a predictable equilibrium that—while not necessarily efficient—dictates how resources are managed, conflicts are resolved, and public goods are distributed. Ultimately, the efficacy with which institutions execute these functions dictates a nation's macroeconomic performance and, by extension, its long-run anthropometric development.

Following the framework established by Acemoglu et al. (2005), economic institutions are inherently endogenous, emerging from the collective choices of a society. However, because alternative institutional arrangements yield asymmetric distributions of resources, societal subgroups harbor conflicting preferences regarding which structures should prevail. Consequently, a conflict emerges over the question of who determines which institutional architecture is ultimately implemented. The authors posit that these prevailing "equilibrium institutions" are dictated by whichever group wields the greatest political power.

Within this framework, they distinguish between two forms of political power: *de jure* and *de facto*. *De jure* political power originates directly from the formal political institutions of a society (such as the form of government), representing power that is legally assigned, such as holding an electoral mandate. Conversely, *de facto* political power is held by groups or individuals who, despite lacking formal authority within the political institutions, can still exert significant influence over political outcomes through their accumulated resources. This includes the capacity to mount armed resistance, mobilize mass protests, or shape legislation through lobbying.

Acemoglu et al. (2002) demonstrate that colonialism itself functioned as an "institutional equilibrium." This equilibrium was established because colonizing nations arrived possessing overwhelming *de facto* political power—primarily through military superiority over the indigenous populations. They subsequently leveraged this advantage to restructure existing political institutions, thereby securing *de jure* political power. Because political power dictates the choice of economic institutions, the colonizing nations ultimately "chose" the institutions for the colonized nations.

However, colonizers did not implement a uniform set of institutions across all territories (Acemoglu et al., 2001). In regions where the historical disease environment prevented large-scale European settlement, imperial powers created "extractive states" (e.g., the Belgian colonization of the Congo). These extractive institutions were not designed to protect the property rights of the broader population or to introduce robust governance structures; rather, the colonizers established only the minimum institutional capacity required to efficiently extract resources and transfer them to the metropole. In contrast, in territories where the disease environment permitted European settlement (such as Australia or Canada), colonizers engaged in genuine state-building, replicating European institutions characterized by strong property rights and inclusive governance structures, designed to protect and serve the settlers (Acemoglu et al., 2001).

At this point, it must be acknowledged that the specific empirical strategy employed by Acemoglu et al. (2001)—namely, the reliance on historical settler mortality rates to identify these regimes—has been subject to methodological critique (e.g., Albouy, 2012). Nevertheless, while the identification strategy remains debated, the broader conceptual approach of proxying institutional differences by differentiating between extractive and settler colonialism is widely accepted within the literature (e.g., Acemoglu et al., 2002; Easterly & Levine, 2016; Engerman & Sokoloff, 2002; Nunn, 2009). Thus, distinguishing between these divergent colonial strategies remains a valid conceptual choice.

Viewed through this theoretical lens, this thesis proposes the following mechanism: Although colonialism represented an institutional equilibrium in both scenarios, the fundamental nature of these equilibria differed vastly. In extractive colonies, the implementation of shallow, purely exploitative institutions created a "fragile" equilibrium, sustained only by the coercive force of an outside power. In contrast, in settler colonies, the implementation of robust, inclusive institutions established a self-sustaining equilibrium that did not rely on external coercion.

However, if colonial rule represented such an institutional equilibrium, then independence should not be understood merely as the legal withdrawal of colonial administration or the restoration of sovereignty. Rather, it constituted an institutional disruption with bifurcated consequences. In extractive colonies, independence acted as a shock-like disruption by removing the external force sustaining the existing institutions, thereby triggering a search for a new equilibrium. In contrast, in settler colonies, where robust institutions rooted in domestic elites were already established, independence did not represent a rupture but rather a formalization of an existing equilibrium.

2.1 Hypothesis 1: Extractive Colonies

The literature supports this conceptualization of independence precipitating a search for a new equilibrium in extractive colonies. Wimmer and Min (2006) identify the transition from imperial rule to independent nation-states as one of the two most profound institutional transformations of the modern era—alongside the initial rise of the empires themselves. Reinforcing this view of a transitional phase, Collins (2017) notes that post-colonial political action was "fundamentally a problem of structure and agency."

This conceptualization raises the critical question of how the search for a new equilibrium manifested in practice. In post-colonial Africa, the sudden absence of a dominant political authority triggered intense struggles among competing elites who were unable to fully consolidate control, resulting in states devoid of hegemony (Collins, 2017; Fadakinte, 2020). A similar dynamic has been documented across post-colonial Pakistan and India: because imperial powers systematically excluded indigenous populations from formal administrative roles, the newly independent nations faced a severe deficit of bureaucratic expertise required to govern (Subrahmanyam, 2006). Furthermore, colonizing nations actively hindered the formation of domestic elites by excluding indigenous populations from universities, thereby limiting the accumulation of human capital (Tödt, 2021). By the time the Congo gained independence in 1960, for instance, there were only 16 indigenous university graduates (Reno, 2006).

Thus, the literature suggests that in extractive colonies, the search for a new equilibrium was a chaotic process that created a profound institutional vacuum. These missing institutional structures in post-colonial states have been linked to a drop in state capacity — understood here as the administrative and institutional capability of a government to execute core functions, maintain essential infrastructure, and deliver public goods and services (Cappelen & Sorens, 2018). Although cross-country panel data quantitatively measuring this drop remain scarce, detailed historical case studies starkly illustrate the magnitude of this collapse.

Vander Eycken and Vander Vorst (1967) demonstrate this for the Belgian Congo: following independence, the number of Belgian civil servants plummeted from 9,620 in 1959 to just 674 in 1965. To counter this reduction, the number of Congolese civil servants was rapidly expanded from 72,000 in 1958 to 122,000 in 1962. However, because they were historically barred from university education, very few of these new civil servants possessed the necessary training to occupy leadership positions. Consequently, the total number of personnel capable of adequately fulfilling a technical or leadership role fell to 1,716, representing an 80% drop compared to the colonial baseline. Simultaneously, available public capital evaporated, leading to a reduction in state investment expenditures from 11.6% of GNP in 1958 to a mere 0.3% in 1964.

This severe deficit in qualified personnel and capital triggered a rapid degradation of public services across multiple sectors (Vander Eycken & Vander Vorst, 1967). In public health, the containment of endemic diseases ceased to be adequately managed, sparking an immediate resurgence of illnesses. Simultaneously, the state lost its monopoly on violence, precipitating localized rebellions and widespread banditry. This institutional collapse extended directly to physical infrastructure; within the years immediately following independence, 80% of the road network became entirely unusable. Domestic rail and river traffic plummeted by 40% between 1959 and 1963, driven by a compounding combination of regional insecurity, the mass departure of foreign technicians, and a structural inability to import critical replacement equipment (Vander Eycken & Vander Vorst, 1967).

Nonetheless, in certain cases, this collapse in state capacity was not simply the result of extractive institutions hindering proper state-building combined with a poorly managed transition like in Congo; rather, it was actively engineered by the departing hegemon. The transition of Guinea in 1958 serves as the prime example of this deliberate destruction of state capacity. As Meredith (2013) details, Guinea's democratic vote for independence was met with a punitive withdrawal by the French administration. An overnight exodus of roughly 3,000 administrators, engineers, and teachers instantly paralyzed the state's logistical machinery. Furthermore, departing officials engaged in systematic sabotage: government archives were burned, infrastructure was vandalized down to the removal of lightbulbs and telephones, and police officers actively destroyed their own barracks. Additionally, departing military physicians confiscated the country's medical supplies, actively disabling the civilian health apparatus. Ultimately, the new government was forced to rebuild its state capacity from an intentionally devastated baseline, relying on a remnant of just 150 French employees who voluntarily remained.

Beyond the collapse of administrative capacity, independence frequently precipitated severe violent conflict. Extractive colonial regimes frequently relied on 'divide and rule' strategies that intentionally fragmented indigenous societies along ethnic or regional lines (Subrahmanyam, 2006). Consequently, as the imperial hegemon withdrew, competing domestic factions fiercely clashed to fill the ensuing power vacuum, effectively paralyzing any peaceful transition (Subrahmanyam, 2006). This dynamic is empirically supported by Wimmer and Min (2006), who identify institutional transformations—such as the formation of independent nation-states from former empires—as important catalysts for war. Thus, in addition to the structural decay of the state apparatus, decolonization was frequently compounded by the outbreak of anti-colonial wars or devastating domestic conflicts.

In conclusion, the literature demonstrates that the search for a new equilibrium following decolonization was frequently a violent and chaotic process in extractive colonies that precipitated a severe collapse in state capacity. While this establishes the political and institutional reality of the transition, a crucial question remains: how does this collapse directly impact anthropometric development?

As suggested by the political economy literature (e.g., Acemoglu, 2005), state capacity serves as the fundamental prerequisite for a government to provide essential public goods, including clean water, sanitation networks, and healthcare infrastructure. When this capacity evaporates—as observed during the independence processes—the provision of these critical public goods ceases (Acemoglu, 2005). In economic anthropometrics, the absence of such public health infrastructure is widely recognized as an important driver of a deteriorating terminal adult height (Iyer, 2010; Komlos, 1998; Thomas & Strauss, 1992).

The mechanism connecting public health infrastructure and stature unfolds as follows: state-provided public health goods are essential for controlling and improving the local disease environment (Alsan & Goldin, 2019; Bleakley, 2007; Cutler & Miller, 2005). The disease environment, in turn, is a paramount determinant of human stature (Baten & Blum, 2014). Biologically, this occurs because high exposure to childhood infections and endemic diseases severely inhibits nutrient absorption, resulting in stunted physical growth (Deaton, 2007). Consequently, the abrupt loss of public health infrastructure during chaotic decolonization transitions functions as a severe biological shock, worsening the disease environment and manifesting as a permanent reduction in the affected population's final adult height.

The capacity of institutional transformations to trigger this exact causal chain finds empirical precedent in the literature. For instance, Adserà et al. (2021) demonstrate that the post-socialist transition from planned to market economies in former communist nations functioned as an institutional shock that ultimately reduced terminal adult height by approximately 1 cm among the affected birth cohorts due to the sudden degradation of local healthcare delivery systems.

However, as discussed previously, independence did not only lead to a drop in state capacity but was also frequently accompanied by severe wars and domestic conflict. The empirical literature demonstrates that such violence exerts a significant negative impact on anthropometric development. Children who grew up during the Biafran War, for instance, suffered a permanent reduction in terminal adult height (Akresh et al., 2012). Similar anthropometric consequences have been documented for the 1993–2005 civil war in Burundi (Bundervoet et al., 2009), while the 2002–2007 civil conflict in Côte d'Ivoire has been shown to have severely deteriorated general childhood health (Minoiu & Shemyakina, 2014).

Thus, the literature strongly suggests that the transition to independence negatively impacted the stature of populations in extractive colonies, driven by a collapse in state capacity, the outbreak of violent conflict, or a compounding of both shocks. This theoretical expectation of a post-colonial biological shock is supported by descriptive anthropometric data. Moradi (2010) demonstrates that while height trends in Sub-Saharan Africa displayed some heterogeneity, they were largely characterized by stagnation or decline following the wave of African decolonization in the 1960s. These negative post-colonial trends for African countries are confirmed by Baten and Blum (2012). Additionally, they observe parallel anthropometric declines in South Asia and stagnation in Southeast Asia during the 1940s and 1950s—the exact periods corresponding to their respective independence movements. During this same timeframe (1940–1960), regions that were never colonized exhibited continuous increases in cohort height.

Therefore, based on the established theoretical framework, this thesis posits the following primary hypothesis:

Hypothesis 1: *Because the search for a new institutional equilibrium resulted in a collapse of state capacity, the transition to independence in extractive colonies had a statistically significant negative impact on the terminal adult stature of the affected cohorts.*

2.2 Hypothesis 2: Settler Colonies

The causal chain laid out for extractive colonies does not hold true for settler colonies. As discussed previously, because the disease environment permitted large-scale European migration, the colonizing powers engaged in genuine state-building, establishing inclusive institutions—such as property rights, public education, and representative assemblies (Acemoglu et al., 2001). Consequently, when these colonies transitioned to independence, the event did not constitute a shock pushing the institutions out of equilibrium. Rather, it represented a legal transfer of sovereignty to an already established domestic elite.

The literature supports this theoretical analysis. Crucially, the academic discussion of state capacity during the independence period is almost entirely absent regarding settler colonies—a gap that itself hints at the unremarkable nature of these transitions. Where evidence does exist, it points toward institutional continuity rather than an institutional vacuum. By the time the United States declared independence in 1776, for instance, it had already established multiple degree-granting universities, enabling the training of a domestic bureaucratic elite (Thelin, 2004). This suggests that the colonies had developed sufficient administrative autonomy to avoid dependence on the British imperial administration, thus preventing the emergence of an administrative vacuum following independence.

Furthermore, because infrastructure investment serves as a proxy for state capacity and the ability to provision public goods (Hendrix, 2010), it should be expected that these investments were not negatively affected by independence if the theoretical framework holds. Historical data on railway expansion in Canada is consistent with this prediction—following Confederation, a rapid expansion of the network occurred, with total track length multiplying several-fold between 1867 and the early twentieth century (Parkin, 1909).

Because state capacity did not collapse, the provision of essential public goods remained stable (Acemoglu, 2005). Through the theoretical lens connecting public goods to the disease environment and subsequently to human stature (Baten & Blum, 2014; Cutler & Miller, 2005), it should therefore be expected that population stature in settler colonies was not negatively affected by the transition to independence.

Therefore, building upon the contrast with extractive colonies, this thesis posits the following secondary hypothesis:

Hypothesis 2: *Because institutional continuity prevented a collapse in state capacity, the transition to independence in settler colonies did not negatively impact the terminal adult stature of the affected cohorts.*

Importantly, this framework abstracts away from intra-national heterogeneity—a distinction particularly crucial in the context of settler colonies. The arguments regarding institutional stability refer specifically to the settler populations that controlled the state apparatus. In contrast, indigenous populations often experienced severe marginalization and deeply extractive policies within those same borders (Wolfe, 2006). While acknowledging this historical reality, measuring this sub-national divergence falls beyond the scope of this paper due to the scarcity of ethnically or regionally disaggregated historical anthropometric data.

2.3 Hypothesis 3: Planned vs. Unplanned Transitions

While the preceding sections used a dichotomous approach to differentiate between settler and extractive regimes, Acemoglu et al. (2001) originally conceptualized these colonial strategies as a continuum. Consequently, macro-level heterogeneity exists within both the extractive and settler colony categories; not every extractive colony experienced an identical institutional vacuum, nor did every settler colony achieve perfect institutional stability. Thus, the discussed framework should be interpreted as capturing aggregate historical trends rather than offering an exact historical description for each individual nation.

However, rather than a limitation, this within-group variation presents a valuable econometric opportunity. Analyzing the variation within groups helps to hold the broad institutional environment constant. This reduces unobservable between-group confounding variables, allowing for a cleaner isolation of the specific mechanisms driving the post-independence state capacity collapse.

To operationalize this heterogeneity, this analysis leverages the historical framework established by Ziltener and Künzler (2017). While their original research identifies three pathways of decolonization in extractive colonies, for the purpose of empirical tractability, this thesis aggregates these trajectories into a binary classification: The first category comprises planned, coordinated, and gradual transfers of power (e.g., Botswana), where colonial administrations actively trained local civil servants and established a domestic administrative elite prior to departure. The second category encompasses rapid, disruptive, and sometimes violent transitions (e.g., Benin or the Democratic Republic of Congo). In these cases, colonizers departed swiftly without adequately training local bureaucrats, often dismantling existing state capacity and precipitating an immediate administrative rupture.

Following the framework established in this section, the subsequent line of reasoning emerges: while both pathways are rooted in extractive institutional structures—implying a baseline decline in the terminal height of the affected cohort—the magnitude of the resulting institutional vacuum is expected to vary systematically with the severity of the political transition. In planned transitions, the deliberate, phased transfer of administrative capacity buffers the newly independent state, mitigating the disruptive shock of decolonization. In contrast, abrupt transitions combine a sudden bureaucratic void with—in some instances—conflict-induced structural destruction, thereby precipitating a more severe collapse in state capacity. This divergence yields the tertiary hypothesis of this thesis:

Hypothesis 3: *Within extractive colonies, the magnitude of the post-independence anthropometric penalty depends upon the nature of the institutional transition. Specifically, abrupt or unplanned transitions induce a more profound biological shock compared to planned transitions.*

As discussed previously, quantitative cross-country evidence directly measuring the causal link from independence to state capacity collapse is scarce. Consequently, the theoretical framework relies on triangulating case study evidence from geographically and historically diverse contexts to establish the generalized trend. While deep qualitative historical analysis is adequate for establishing robust causal mechanisms, the illustrative cases presented in this section are limited in both number and depth due to the scope of this paper. Because this brief selection excludes a comprehensive historical proof, the framework should be interpreted as a suggestive baseline. This limitation directly motivates the subsequent empirical analysis, which aims to complement these qualitative insights with an econometric approach to support the suggested causal mechanism.

3 Variables and Data

To translate the preceding theoretical framework into an empirical design, this chapter motivates the selection of the variables, details their operationalization, and discusses the properties of the underlying data. A comprehensive overview of the master sample, incorporating all variable definitions and coding decisions introduced in this chapter, is provided in Table A.1.[1]

3.1 Human Stature

Employing human stature as the dependent variable fulfills a core contribution outlined in the introduction: capturing a non-economic dimension of long-run well-being. Terminal adult height measures the cumulative net nutritional, epidemiological, and environmental experiences of childhood and adolescence (Baten & Blum, 2014; Deaton, 2003). Consequently, fluctuations in average cohort height directly reflect structural changes in public health, nutrition, and the broader socioeconomic environment—making stature a robust proxy for a specific non-economic dimension of welfare, namely health (Baten et al., 2013). Stature is preferred over alternative indicators because those alternatives are either unavailable for the historical period under study—the Human Development Index, for instance, was only implemented in 1990—or are characterized by severe data discontinuities, as is the case for infant mortality (Baten & Blum, 2015) and life expectancy (Zijdeman, 2015). Height therefore represents the most reliable and consistent historical indicator for capturing non-economic welfare trends over the long run (Baten et al., 2013).

The anthropometric data are sourced from the Clio Infra database (Baten & Blum, 2012, 2015),

[1] The rationale behind the country selection for this sample is discussed in detail in Chapter 4.4.

which provides decadal birth cohort estimates of average male height (in centimeters) for 165 countries spanning the period 1810 to 1989. To interpret this metric accurately, it is necessary to clarify what this structural format implies. Because the observations are organized into decadal birth cohorts, each data point (e.g., 1960) represents an aggregate of all individuals born within that ten-year interval (1960–1969). The authors adopted this approach because historical anthropometric data are scarce at the annual level; by aggregating into decadal intervals, individual year outliers are smoothed out, and the reliability of the estimates is improved.

A further interpretive point concerns the temporal dimension of this metric. Because the recorded stature represents terminal adult height, a specific birth cohort serves as a biological record of the institutional and environmental conditions prevailing during the decades following birth (Baten & Blum, 2012, 2015). Thus, the 1960 cohort does not reflect the *status quo* of that specific year, but rather the cumulative net nutritional investment made in that generation from infancy through adolescence.

The selection of the Clio Infra database is motivated by two reasons. First, the database has been applied in similar macro-historical settings and demonstrated robustness (e.g., Baten & Blum, 2012, 2014; OECD, 2014; Roser et al., 2021). Second, this selection is dictated by the historical scope of the research design. Anthropometric data for long historical periods are scarce, with Clio Infra and the NCD Risk Factor Collaboration (NCD-RisC) (2016) representing the only two comprehensive global panels available. However, while NCD-RisC offers highly granular annual data, its underlying methodology presents a significant constraint for this analysis. NCD-RisC relies on a Bayesian hierarchical model that estimates missing country-year observations by borrowing strength across temporal and geographic dimensions. This estimation procedure systematically pulls smaller sample sizes and localized variations back toward a global or regional group mean, introducing a risk of artificially smoothing out sharp structural breaks (Veenman et al., 2024). Because the empirical objective of this thesis is to isolate the institutional shocks precipitated by the transition to independence, data that systematically suppress such sharp deviations are methodologically unsuitable. Consequently, because Clio Infra constructs its decadal cohorts directly from documented historical measurements—such as military, prison, and survey records—rather than relying on geographic interpolation, it represents the most appropriate panel available to capture the pre- and post-independence anthropometric trends required for this analysis.

3.2 Colonial Status

To determine both a country's colonial status and its year of independence, this analysis utilizes the Colonial Dates Dataset (COLDAT) (Becker, 2023). COLDAT was selected because it systematically aggregates the most prominent colonial datasets in the literature (Correlates of War Project (CoW), 2025; Lange et al., 2006; Olsson, 2009; Wimmer & Min, 2006). By synthesizing these sources, COLDAT effectively minimizes the measurement error and definitional fluctuations that often arise when relying on a single, isolated dataset. Consequently, this thesis adopts the definition of colonialism established by Becker (2023):

> *"A colony is a territory whose domestic and or foreign affairs are dominated by a European nation state, and whose population is constructed as inferior to the colonizer. A European colony is a territory outside of Europe colonized by a European power. [...] A territory is usually regarded as a colony once external control is established over a significant part of its territory. A formal declaration is not necessary."*

Based on these criteria, the sample is stratified into a colonized and never-colonized group, where the colonized nations constitute the treated group and the never-colonized nations serve as the primary control group.

3.3 Independence

Following the definition employed in the COLDAT framework, the timing of independence is understood as the year in which "control by a European power vanished" (Becker, 2023). Operationalizing this definition presents two methodological challenges. First, while the Clio-Infra height data are structured as decadal birth cohorts, the COLDAT independence dates are reported on an annual basis, requiring harmonization between the two datasets. Second, because COLDAT aggregates four underlying data sources that do not always agree on the timing of independence, the dataset provides two possible dates: a mean date and a maximum date. The selection of one of these dates for the primary model requires clear justification. The following subsections address these two issues.

3.3.1 Data Harmonization

To combine the independence dates with the Clio-Infra height data, each country's independence year is rounded to the nearest decadal observation. For instance, a country gaining independence in 1956 is assigned to the 1960 cohort, whereas a country gaining independence in 1954 is assigned to 1950.

This rounding strategy is chosen to minimize the trade-off between two competing forms of measurement error. Flooring the treatment date (rounding down) captures the immediate onset of the institutional rupture but introduces attenuation bias, as the designated treatment cohort would inevitably include several years of unaffected pre-shock births. Conversely, ceiling the treatment date (rounding up) guarantees a purely post-shock treated cohort but introduces control-group contamination, since the untreated baseline would then contain individuals who spent part of their childhood under the post-independence institutional environment. Standard mathematical rounding represents a compromise between these opposing biases and ensures that the majority of individuals within a given birth cohort experienced the institutional framework that dominates that cohort's time period.

It must be noted that measurement error cannot be eliminated completely under this approach. However, as shown in Table A.1, within the estimation sample used in this study, six of the fourteen extractive colonies possess independence dates that already coincide with a full decade, rendering rounding unnecessary. Among the remaining countries, the direction of rounding is relatively balanced, with three cases rounding upward and five cases rounding downward. Consequently, the two sources of bias largely offset one another. Nevertheless, the slight predominance of downward rounding implies a small attenuation bias, suggesting that the estimated treatment effects may understate the true post-independence shock and should therefore be interpreted as lower-bound estimates. Among the two settler colonies, the balance is exact, with one case rounding upward and one downward. Overall, measurement error arising from temporal aggregation is therefore expected to exert only a minor influence on the estimation results.

3.3.2 Independence Date Selection

As noted above, COLDAT provides two alternative chronological estimates of independence: an unweighted mean of all dates recorded across its underlying sources and a maximum date corresponding to the latest year reported by any single source. For example, in the case of Canada, Lange et al. (2006) and Olsson (2009) record the independence date as 1867, Wimmer and Min (2006) report 1866, while CoW (2025) records 1919. The mean approach calculates a simple average of these four dates, rounding to the nearest integer (yielding 1877), while the maximum date is recorded as 1919.

The availability of both estimates is a distinct advantage of COLDAT, as the divergence between the mean and maximum date serves as a diagnostic indicator of source congruence. Where the two metrics are identical or nearly identical, it signals broad agreement across the four datasets, thereby increasing confidence in the documented date. For instance, both the mean and maximum date for Burkina Faso are recorded as 1960, indicating that all sources agree on the same independence year. Such congruence, however, is not universal. In some cases, competing definitions of formal sovereignty generate substantial disagreement among the underlying datasets (Becker, 2023). Under these conditions, calculating a simple unweighted average may generate a historically vacant "phantom date" that does not correspond to any actual institutional event.

Canada illustrates this issue particularly clearly. As discussed previously, a majority cluster of three out of four underlying sources places independence between 1866 and 1867. However, the standard unweighted mean across all sources equals 1877 because it is artificially skewed upward by the 1919 CoW observation. Consequently, the mean date becomes a purely arithmetic artifact that corresponds neither to the timeline identified by the majority of sources nor to the maximum date itself.

The final choice of independence dates is therefore based on three decision pathways determined by the degree of source congruence.

High Congruence: The first and most common pathway applies to thirteen of the sixteen colonies included in the sample. In these cases, independence dates are sufficiently clustered across the underlying sources that the mean and maximum measures either coincide or round to the same decadal cohort, rendering the distinction empirically irrelevant. For example, the mean date for Tanzania is 1961 and the maximum date is 1963. Although these dates differ, standard rounding assigns both to the 1960 cohort, implying that the discrepancy has no effect on the analysis. In such cases, the standard COLDAT mean is adopted directly.

Moderate Divergence: The second pathway applies to Indonesia and Australia, where discrepancies between the mean and maximum dates are sufficiently large that the two measures round to different decades. In these cases, the underlying source-level dates are examined to determine whether the COLDAT mean reflects a genuine historical consensus or merely a phantom date. For Indonesia, the mean date is 1952, which rounds to 1950, while the maximum date is 1962, which rounds to 1960. Inspection of the underlying sources reveals that Lange et al. (2006) and Olsson (2009) record 1945, Wimmer and Min (2006) records 1950, and CoW (2025) records 1962. Consequently, three of the four sources round to the 1950 cohort, confirming that the COLDAT mean reflects a genuine historical consensus rather than an arithmetic artifact. The mean is therefore retained in the baseline specification. Additionally, to assess sensitivity to alternative historical timelines, robustness checks are conducted using the maximum recorded dates (Appendix B.1 and B.2).

Severe Divergence: The third pathway applies exclusively to Canada, where disagreement among the underlying datasets is sufficiently pronounced to generate a phantom mean. As illustrated above, the standard COLDAT mean is 1877, rounding to the 1880 cohort. Conversely, the majority cluster centers around 1867 (the mean of the dates documented by Lange et al. (2006), Wimmer and Min (2006), and Olsson (2009)), which rounds to 1870. Finally, the maximum date of 1919 rounds to the 1920 cohort. Consequently, the unweighted mean deviates substantially from any historically documented institutional event. To eliminate this arithmetic artifact, the standard COLDAT mean is replaced by the majority-cluster date of 1870 in the baseline specification. Following the protocol established for the previous pathway, a robustness check using the maximum date (1920) is also performed (Appendix B.1 and B.2).

The independence date established by these criteria serves as the start of treatment (T_0).

3.4 Extractive and Settler Colonies

Analyzing Hypotheses 1 and 2 requires a further step of stratification. Building on the initial colonized–never-colonized dimension, the treated group is divided into extractive and settler colonies using the framework established by Acemoglu et al. (2001). As discussed previously, the empirical identification strategy employed in that study—particularly the use of settler mortality as an instrument—has been subject to critique (Albouy, 2012). However, this criticism mainly concerns the reliability of the historical mortality data when used as a continuous measure of institutional development, rather than the underlying conceptual distinction between the different colonial strategies themselves.

For the purposes of this analysis, a binary classification of colonial regimes is sufficient. In this context, the conceptual framework remains highly robust, especially given that subsequent contributions have employed alternative measures to capture similar historical distinctions. For instance, Easterly and Levine (2016) construct a dataset based on the European share of the population during colonization, classifying countries with a settler share above a given threshold as settler colonies. Their findings are broadly consistent with the structural patterns identified by Acemoglu et al. (2001), lending further empirical support to the validity of this classification.

Following this rationale, the "Neo-Europes" in the sample—Canada and Australia—are classified as settler colonies. In contrast, the colonized countries within Sub-Saharan Africa, as well as South and Southeast Asia, are classified as extractive colonies.

3.5 Planned and Unplanned Transitions

Finally, to evaluate Hypothesis 3, the extractive cohort is further subdivided based on variations in power transfers. To this end, this thesis draws on the dataset constructed by Ziltener and Künzler (2017), which encompasses 83 countries across Africa and Asia. The dataset was chosen as it provides, to the best of the author's knowledge, the only available systematic cross-country codification of decolonization pathways, enabling the assessment of transition heterogeneity necessary for this paper. As outlined in the theoretical framework, this analysis collapses the three categories of Ziltener and Künzler (2017) into two distinct pathways of decolonization: planned and unplanned.

4 Empirical Strategy

With the data and key variables established, the analysis now turns to the empirical strategy. Selecting the appropriate econometric framework requires briefly revisiting the primary objective and theoretical setting of the study. The objective is to causally isolate the effect of independence on human stature. To this end, independence is conceptualized as the treatment event, meaning that treatment timing (T_0) is defined by each country's year of independence. Countries that have achieved independence are therefore considered treated units, while the control group consists of countries that have not experienced independence at a given point in time, including both never-colonized countries and not-yet-independent territories.

4.1 Identification Challenges

A naive approach to addressing the research question would be to compare average height levels within a country before and after independence. However, such a comparison is confounded by global time trends, making it impossible to disentangle the causal effect of independence from broader improvements in health and nutrition. Alternatively, one might consider a cross-sectional comparison between the treatment and control group. Yet, this approach suffers from selection bias, as these groups differ systematically in geographic, historical, and economic characteristics. Consequently, both approaches would yield biased estimates of the treatment effect.

The empirical structure of this study naturally lends itself to a Difference-in-Differences (DiD) framework, which addresses these challenges. By differencing over time and across groups, the canonical DiD estimator eliminates both common global trends and time-invariant unobserved heterogeneity. However, unlike in the canonical 2×2 DiD setup, decolonization occurred in staggered waves over time. This requires an estimator capable of handling variation in treatment timing. Recent advances in the econometric literature propose robust solutions for such staggered adoption settings (e.g., Callaway & Sant'Anna, 2021; Sun & Abraham, 2021).

Nonetheless, these approaches still rely on another DiD requirement: the parallel trends assumption. This condition demands that, in the absence of independence, the average stature of the colonies would have evolved perfectly in tandem with the never-colonized control group. Given the profound structural, geographic, and economic disparities between the predominantly European control group and the colonized world, unconditional parallel trends (as required by Sun & Abraham, 2021) are highly implausible.[1]

Callaway and Sant'Anna (2021) relax this restriction by allowing for conditional parallel trends—meaning trends must only be parallel after adjusting for observed pre-treatment covariates. Nevertheless, this solution is also unsuited for the specific context of this paper. First, reliable pre-treatment covariates for 19th- and early 20th-century colonies are scarce. Second, the structural divergence between the two groups is so pronounced that conditioning on observable covariates is unlikely to resolve the fundamental baseline imbalances. Finally, their estimator relies on sufficiently large samples and overlapping treatment cohorts to ensure reliable estimation and inference. Table A.1 shows that these conditions are not satisfied in the present setting due to the small number of treated units.

Another potential approach to address violations of the parallel trends assumption is the Synthetic Control (SC) method (Abadie & Gardeazabal, 2003; Abadie et al., 2010). This estimator constructs a weighted combination of control units to match the pre-treatment trajectory of

[1] Chapter 5.1 and 5.2 formally demonstrate that this assumption is violated.

the treated unit, thereby relaxing the need for strict parallel trends. However, because the SC method does not explicitly account for unit fixed effects (an additive intercept shift), any level differences between treated and control units must be captured entirely through pre-treatment matching. Because the structural and economic divides between the treated and control groups are so pronounced, constructing a credible synthetic control based purely on pre-treatment fit is highly improbable. Consequently, the applicability of SC in this context is limited.

4.2 Synthetic Difference-in-Differences

Introduced by Arkhangelsky et al. (2021), the Synthetic Difference-in-Differences (SDiD) estimator provides a suitable solution to the challenges identified above. While the canonical SDiD framework is formulated for block treatment adoption, recent extensions allow for staggered treatment timing (e.g., Ciccia, 2024; Porreca, 2022), making the method applicable to the decolonization setting considered in this study.

More importantly, the estimator is specifically designed for applications in which the standard parallel trends assumption is unlikely to hold. In contrast to approaches such as Callaway and Sant'Anna (2021), it does not rely on conditioning on observed pre-treatment covariates to achieve comparability. Instead, SDiD synthesizes the strengths of both DiD and SC methods by applying a double-weighting procedure. It constructs unit and time weights that reweight both the control units and the pre-treatment periods to optimally match the pre-intervention trajectory of the treated units. This generates a data-driven counterfactual that approximates the dependent variable in the absence of treatment. Furthermore, by incorporating fixed effects alongside this weighting procedure, SDiD overcomes the limitations of standard SC by allowing for absolute level differences between the treated and synthetic control units. This provides the necessary flexibility to construct a credible counterfactual for the present setting.

Finally, because SDiD allows for a placebo inference approach that remains asymptotically consistent even for treatment groups as small as $N = 1$ (Arkhangelsky et al., 2021; Ciccia, 2024), it is uniquely suited to the small sample sizes analyzed in this study.[2]

4.2.1 Econometric Specifications

Arkhangelsky et al. (2021) and Clarke et al. (2023) show that the SDiD procedure described above can be mathematically expressed as follows:

$$(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \arg \min_{\tau,\mu,\alpha,\beta} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \hat{\omega}_i^{sdid} \hat{\lambda}_t^{sdid} \right\} \tag{4.1}$$

To operationalize this estimator, a balanced panel consisting of N units observed over T time periods is required.[3] For every unit i (i.e., each country) in each period t (i.e., each decade), an outcome Y_{it} is observed, which in this context corresponds to the country's decadal mean male height. The variable W_{it} serves as a binary treatment indicator, where $W_{it} = 1$ if unit i is treated in period t, and $W_{it} = 0$ if the unit remains untreated. As an illustration, a country gaining independence in 1953 is assigned a T_0 value of 1950 (following the rounding procedure discussed above); consequently, its treatment indicator switches to $W_{it} = 1$ from the 1950 cohort onward. Conversely, a country that was never colonized retains $W_{it} = 0$ throughout the entire sample period. The parameter μ represents the global intercept. Additionally, the equation incorporates unit fixed effects (α_i) and time fixed effects (β_t). As previously established, these parameters capture time-invariant baseline disparities between countries—such as underlying geographic differences in human stature—as well as common global time shocks, such as worldwide improvements in nutrition or health technologies.

[2]The specific statistical conditions required for this consistency, along with the empirical validation of their fulfillment in this paper, are detailed in Appendix B.6.

[3]The empirical implications of this requirement are discussed in detail in Section 4.4.

Ultimately, the equation demonstrates that SDiD is a weighted Two-Way Fixed Effects (TWFE) optimization problem. The arg min operator seeks to find the specific causal parameter $\hat{\tau}^{sdid}$ and fixed effects that minimize the sum of squared residuals between the observed average height and the model's predicted outcome. However, unlike standard Ordinary Least Squares (OLS), which assigns equal importance to all observations, the penalty for these residual errors is scaled by two sets of generated weights:

$\hat{\omega}_i^{sdid}$ **(Unit Weights)**: The unit weights determine the proportional weight assigned to each country within the control group (also referred to as the donor pool). They are optimally chosen to ensure the synthetic control group mirrors the pre-treatment trajectory of the treated units as closely as possible.

$\hat{\lambda}_t^{sdid}$ **(Time Weights)**: The time weights are selected by matching pre-treatment periods to post-treatment outcomes within the control group. By anchoring to the post-treatment average of these control units, the algorithm assigns higher weights to the historical periods that best predict this post-shock reality. This ensures that the subsequent divergence of the treated group is evaluated against a baseline that optimally captures the potential counterfactual trajectory in the absence of the shock.

As formally demonstrated in Appendix D.1, the configuration of these weights is based on an optimization procedure. Here, both sets of weights are constrained to be non-negative and sum to one. By restricting the weights to the simplex, the estimator ensures that the synthetic counterfactual is strictly a convex combination of the observed data. In practical terms, this means the counterfactual is purely a weighted average of the actual countries in the donor pool. Therefore, the synthetic unit's pre-treatment height can never be taller than the tallest country in the pool, nor shorter than the shortest, keeping the estimation strictly within the bounds of reality.

Furthermore, this optimization process incorporates a ridge penalty—a regularization technique applied to stabilize the estimation. For the unit weights, this penalty prevents the algorithm from placing excessive weight on a single donor country, thereby mitigating the risk of overfitting.

Importantly, the calculation of these weights is explicitly anchored to the unweighted average of the treated units. Therefore, the resulting causal parameter $\hat{\tau}^{sdid}$ identifies the Average Treatment Effect on the Treated ($\widehat{ATT}$). [4] As such, the estimator does not recover the effect of decolonization for a randomly selected country (ATE), but rather the historical impact on the specific cohort of colonies that experienced independence.

4.2.2 Event-Study Extension

Equation 4.1 formalizes the classical SDiD estimator under the assumption of simultaneous treatment adoption. While instructive for understanding the underlying mechanics, the model is unsuitable for this thesis due to the staggered timing of independence. To accurately capture this dynamic, the methodological extension developed by Ciccia (2024) is employed. He demonstrates that the SDiD estimator can be re-parameterized into a dynamic event-study framework. By realigning the varying calendar years of independence into a relative event-time scale (where the timing of independence is normalized to $t = 0$), the complexities associated with staggered adoption are neutralized. Mathematically, this is achieved by disaggregating $\hat{\tau}^{sdid}$ as follows:

$$\hat{\tau}^{sdid}_{a,\ell} = \underbrace{\frac{1}{N^a_{tr}} \sum_{i \in I^a} Y_{i,a-1+\ell}}_{\substack{\text{Observed} \\ \text{Treated Outcome}}} - \underbrace{\sum_{i=1}^{N_{co}} \omega_i Y_{i,a-1+\ell}}_{\substack{\text{Synthetic} \\ \text{Counterfactual}}} - \underbrace{\sum_{t=1}^{a-1} \left(\frac{1}{N^a_{tr}} \sum_{i \in I^a} \lambda_t Y_{i,t} - \sum_{i=1}^{N_{co}} \omega_i \lambda_t Y_{i,t} \right)}_{\substack{\text{Weighted} \\ \text{Pre-Treatment Gap}}} \tag{4.2}$$

[4] Throughout this thesis, ATT denotes the theoretical population parameter, while $\widehat{ATT}$ represents the empirical SDiD estimator.

At its core, this equation executes a classical DiD estimation. However, it extends the traditional model by realigning the data into a relative event-study framework and applying optimal synthetic weights to construct the counterfactual. To understand these mechanics, the formula can be broken down term by term.

Starting with the foundational notation, index i denotes an individual country. Building on this, a represents a specific adoption cohort (i.e., the calendar year a group of colonies achieved independence), and ℓ denotes the relative time period passed since that adoption. For any given cohort, I^a denotes the specific set of treated units (the newly independent colonies), and N_{tr}^a represents the total number of units within that treated pool. Accordingly, the first term calculates the simple average of the outcome Y (the mean height) for these colonies. To synchronize the relative event timeline with the absolute calendar grid, the subscript utilizes $a - 1$ (the last pre-independence decade immediately preceding cohort a's independence) as a permanent historical anchor. By adding the relative time ℓ to this anchor year, the formula pulls the empirical data exactly ℓ years into the independence era. Consequently, this term represents the observed post-treatment reality.

Importantly, because the event-study specification accommodates varying adoption cohorts a, the composition of the donor pool differs from that in a simultaneous-adoption framework. Under simultaneous treatment, the control group consists exclusively of never-treated units (i.e., never-colonized countries). In the staggered setting, however, colonies that have not yet achieved independence may also serve as valid controls. This inclusion is subject to an important restriction: a country can only remain in the donor pool if it stays untreated throughout the entire post-treatment horizon being evaluated.

The second term uses these control units to construct the unobserved counterfactual.

Here, N_{co} represents the total number of units in the donor pool. By applying the optimal unit weights (ω_i), the term then calculates a weighted average of the control units' outcomes in the exact same relative period ($a - 1 + \ell$) as the treated cohort. The difference between the first and second terms represents the raw gap between the actual colonies and the synthetic control at time ℓ. This calculation inherently neutralizes global time trends, thus representing the "first difference" in a classical DiD setting.

While the first difference removes time trends, the raw post-treatment gap calculated above may still be contaminated by selection bias if the treated cohort and the donor countries possess inherently different structural baselines. The final term represents the "second difference" and neutralizes exactly that bias. It evaluates the entire pre-treatment history up to $a - 1$ and calculates the historical gap between the actual colonies and the donor pool for each period t. The outcome data for both groups in these prior periods is multiplied by the optimal time weights λ_t. Consequently, the subtracted baseline correction is driven exclusively by the historically relevant periods, neutralizing pre-treatment disparities without absorbing irrelevant historical noise. Finally, by subtracting the weighted pre-existing baseline (the second difference) from the post-treatment gap (the first difference), the resulting difference of these differences isolates the unbiased, causal effect of independence for cohort a.

While the cohort-specific estimator ($\hat{\tau}_{a,\ell}^{sdid}$) identifies the causal effect for individual adoption groups, estimating the overarching structural impact of a treatment requires aggregating these varying timelines. Ciccia (2024) demonstrates how this can be achieved for each relative period ℓ:

$$\hat{\tau}_{\ell}^{sdid} = \sum_{a \in A_{\ell}} \frac{N_{tr}^{a}}{N_{tr}^{\ell}} \hat{\tau}_{a,\ell}^{sdid} \tag{4.3}$$

The summation evaluates A_ℓ, which represents the subset of all treated cohorts that actually possess observable data for relative period ℓ. Additionally, rather than taking a simple average of the cohort effects, the estimator relies on a weighted sum. It multiplies each cohort's effect ($\hat{\tau}^{sdid}_{a,\ell}$) by its proportional size ($\frac{N^a_{tr}}{N^\ell_{tr}}$), where N^ℓ_{tr} is the total number of treated units across all cohorts observable at time ℓ. This ensures that larger independence cohorts proportionately drive the aggregate event-study trajectory.

In a final step, the global $\widehat{ATT}$ can be recovered directly from the aggregated event-study estimates by performing the following calculation:

$$\widehat{ATT} = \frac{1}{T_{post}} \sum_{\ell=1}^{T_{tr}} N^\ell_{tr} \hat{\tau}^{sdid}_\ell \tag{4.4}$$

This equation aggregates all point estimates ($\hat{\tau}^{sdid}_\ell$) across the post-treatment relative periods, summing from $\ell = 1$ up to T_{tr}, which represents the maximum number of post-treatment periods observed across the sample. Each point estimate is multiplied by the number of treated units observable in a specific relative period N^ℓ_{tr}. This ensures that relative years with robust sample sizes exert a proportionally stronger influence on the final average. Finally, the entire summation is divided by T_{post} (the total number of treated post-treatment observations). This shows that the event-study specification and the classical SDiD are mathematically unified; the global $\widehat{ATT}$ is simply the weighted average of the post-treatment event-study timeline.

Importantly, for this ATT to represent a truly unbiased causal effect, the identification assumptions of classical DiD must be satisfied: the no-anticipation assumption, the absence of time-varying confounders at the time of treatment, and the Stable Unit Treatment Value Assumption (SUTVA) (Arkhangelsky & Samkov, 2025). Appendix C demonstrates that the empirical design is robust against severe violations of these assumptions.

4.3 Implications for Model Specification

The event-study framework formalized in Equations 4.2 through 4.4 represents the final econometric specification employed to analyze the research question. However, the double-weighting procedure embedded within this estimator has important implications for the broader empirical setup. Specifically, it dictates a departure from classical TWFE approaches regarding both the inclusion of control variables and the use of interaction terms to evaluate heterogeneity.

4.3.1 Covariates

Arkhangelsky et al. (2021) demonstrate that the inclusion of additional control variables in the SDiD framework is not necessary. In a conventional TWFE model, only additive unit (α_i) and time fixed effects (β_t) are included. While these account for permanent differences between countries and common shocks affecting all units, they cannot capture unobserved factors that influence countries differently over time. Consequently, time-varying control variables are often introduced to proxy for such dynamics.

As formally shown in Appendix D.2, SDiD addresses this issue through a latent interactive fixed effects structure that allows common shocks to affect countries with different intensities. Importantly, these latent effects should not be confused with the standard additive fixed effects (α_i and β_t) appearing in Equation 4.1. They are never specified explicitly as parameters in the estimation equation; rather, the algorithm's double-weighting procedure implicitly accounts for them during the construction of the synthetic counterfactual. As a result, provided that a satisfactory pre-treatment match is achieved, the inclusion of additional control variables becomes mathematically redundant. Consequently, this study does not employ time-varying control variables.

4.3.2 Heterogeneity Analysis

Furthermore, the nature of the SDiD matching process alters how treatment heterogeneity must be evaluated (Arkhangelsky et al., 2021). In classical TWFE models, heterogeneity is typically assessed by pooling all observations and introducing a binary interaction term (e.g., $W_{it} \times$ Extractive). However, applying this pooled approach in an SDiD framework undermines the estimator's core mechanical strength: the targeted construction of a synthetic counterfactual. If structurally divergent groups—such as settler and extractive colonies—are pooled into a single treatment cohort, the algorithm is forced to optimize a single set of unit and time weights that match the average pre-treatment trajectory of this blended sample. This would result in a compromised synthetic control group that accurately reflects the historical baseline of neither group. To avoid this bias, the recommended approach in an SDiD setting is to stratify the sample along the dimensions where heterogeneity is expected (Arkhangelsky & Samkov, 2025). By estimating the event-study model separately for each group, the algorithm is permitted to construct counterfactuals calibrated exclusively to the specific historical trajectories of those subsets.

Therefore, this study relies on sample splitting to evaluate Hypothesis 2 (Extractive vs. Settler) and Hypothesis 3 (Planned vs. Unplanned).

4.4 Panel Construction

Having established the empirical setup, the final methodological step is to address the SDiD data requirements. For the algorithm to function, a strictly balanced panel is necessary, where each treated unit possesses a minimum of two pre-treatment observations and one post-treatment observation (Arkhangelsky et al., 2021; Clarke et al., 2023). However, given the historical scope of this study, the height data are inherently unbalanced, with missing observations across various country-year cells. Consequently, constructing a balanced panel required systematic sample restrictions. This implies that while certain countries possess sufficient isolated data for unit-by-unit analysis, they had to be excluded from the main pooled regression.

In general, the construction of the balanced panel was treated as an optimization problem, balancing the trade-off between statistical power (sample size) and the structural integrity of the identification assumptions. This optimization was executed in two distinct steps: defining the temporal boundaries of the panel first, and then establishing strict unit-level inclusion criteria.

4.4.1 Temporal Boundaries

The estimation window for the extractive cohort is fixed between 1910 and 1980. The lower bound of 1910 is determined by two considerations. First, preliminary data exploration indicated that the nineteenth century represents a distinct structural environment. Extending the panel backward captured divergent pre-treatment trends during those earlier periods that could not be reconciled via unit-weighting. Because the credibility of the SDiD counterfactual relies heavily on structural alignment in the decades immediately preceding the institutional shock, forcing the algorithm to match on distant, misaligned historical periods yields no analytical benefit. Thus, truncating the lower bound at 1910 restricts the algorithm to the most critical pre-independence decades, where parallel trends hold. Second, the availability of continuous anthropometric data for this group declines markedly before 1910, meaning that the inclusion of earlier periods would have required a substantial reduction in the number of treated units.

The upper bound is set at 1980 because the Clio-Infra dataset provides observations only until that point for most countries. Although extending the analysis closer to the present would be possible using alternative data sources (i.e., NCD Risk Factor Collaboration (NCD-RisC), 2016), doing so would introduce composition bias arising from differences in sampling methodology between Clio-Infra and other datasets. Moreover, while observations several decades removed from the institutional shock are valuable for mapping general, long-term trends, they carry a higher risk of being confounded by subsequent, unrelated macroeconomic or political shocks. The potential benefits of extending the panel were therefore judged to be smaller than the risk of introducing additional bias.

Furthermore, a central principle of choosing the estimation window was the prioritization of pre-treatment trend alignment over the simple maximization of the treated sample size (N_{tr}). For example, excluding 1940 would have increased the extractive cohort from 14 to 19 units, but this specification was rejected because it worsened the synthetic pre-treatment match. Consequently, the baseline specification retains 1940 and favors stronger identification over a larger sample.[5]

Because the settler cohort generally achieved independence earlier than the extractive cohort, a shifted estimation window was required to capture their transitions. Initial data exploration indicated that for the settler colonies, data continuity was stronger and pre-treatment trends were more stable in earlier periods than for the extractive colonies, allowing such a shift. To maintain methodological symmetry with the extractive specification, a corresponding eight-decade window from 1850 to 1920 was selected.

[5]The full analysis of the alternative specification excluding 1940 is detailed in Appendix B.4.

4.4.2 Unit-Level Inclusion Criteria

Establishing the estimation windows constitutes only the first phase of panel construction. To be retained in the final sample, a country must also satisfy the following unit-level conditions within its respective time frame:

Pre-Treatment Periods ($T_{pre} \geq 2$): At least two pre-treatment observations are available, since multiple data points are necessary to establish a baseline trajectory and compute the unit and time weights that align the donor pool with the treated unit (Arkhangelsky et al., 2021).

Pre-Treatment Anchoring ($t = a - 1$): A valid observation immediately prior to independence exists. This rule is implemented because the immediate pre-treatment period serves as a fundamental reference point for the SDiD baseline correction (Arkhangelsky et al., 2021). Missing data at this critical juncture would force the estimator to extrapolate across a structural gap on the eve of the institutional transition, significantly weakening the reliability of the synthetic match.

Post-Treatment Periods ($T_{post} \geq 1$): At least one post-treatment observation is provided, since without a realized post-event outcome, it is impossible to calculate the divergence from the synthetic counterfactual.

Beyond these temporal requirements, the sample was subject to a final constraint regarding biological plausibility. The absolute maximum population-level growth rate recorded in modern history is approximately 2 centimeters per decade (NCD Risk Factor Collaboration (NCD-RisC), 2016), demonstrating that human stature evolves gradually across generations. Consequently, extreme short-term volatility in decadal averages typically indicates underlying measurement error or sampling bias within the primary sources. Ultimately, this led to the exclusion of Malaysia, as the underlying data reported an anomalous 10-centimeter increase between 1970 and 1980.

As Table 4.1 illustrates, the systematic exclusion results in a distinct geographic concentration. The final extractive cohort ($N_{tr} = 14$) is composed entirely of nations located in Sub-Saharan Africa ($N_{tr} = 8$) and South and Southeast Asia ($N_{tr} = 6$). Consequently, the treatment effects ($\widehat{ATT}$) should not be interpreted as a universal global average, but rather as effects identified within the specific institutional context of these regions. Furthermore, the geographic concentration is accompanied by a distinct temporal clustering in treatment timing. The South and Southeast Asian nations cluster around the 1950 decadal observation, whereas the Sub-Saharan African transitions are concentrated in 1960.

Table 4.1: Estimation Sample: Treated Units by Colony Type

Country	**Region**	T_0 **(Rounded)**	**Transition**
Panel A: Extractive Colonies ($N = 14$)			
Burkina Faso	Sub-Saharan Africa	1960	Unplanned
Ghana	Sub-Saharan Africa	1960	Planned
Ivory Coast	Sub-Saharan Africa	1960	Planned
Mali	Sub-Saharan Africa	1960	Unplanned
Niger	Sub-Saharan Africa	1960	Unplanned
Nigeria	Sub-Saharan Africa	1960	Planned
Senegal	Sub-Saharan Africa	1960	Unplanned
Tanzania	Sub-Saharan Africa	1960	Planned
Cambodia	Southeast Asia	1950	Unplanned
India	South Asia	1950	Planned
Indonesia	Southeast Asia	1950	Unplanned
Myanmar	Southeast Asia	1950	Unplanned
Laos	Southeast Asia	1950	Unplanned
Vietnam	Southeast Asia	1950	Unplanned
Panel B: Settler Colonies ($N = 2$)			
Canada	North America	1870	N/A
Australia	Oceania	1900	N/A

Notes: This table presents the final treated estimation sample stratified by institutional regime (Panel A and Panel B). T_0 denotes the rounded independence dates (start of treatment). *Planned* refers to transitions characterized by gradual administrative handovers; *Unplanned* indicates transitions defined by abrupt political ruptures or violent conflict. *N/A* denotes units for which this categorization is not applicable.

In contrast, the settler cohort is reduced to only two units ($N_{tr} = 2$): Canada and Australia. This restricted sample size is driven by three constraints. First, as a matter of historical reality, purely settler colonies were considerably less common than extractive regimes. Second, the early independence of some settler societies—most notably the United States in 1776—precludes their inclusion, as continuous anthropometric data for a viable untreated donor pool do not extend sufficiently far back to construct a credible synthetic match. Third, unlike the extractive cohort, which experienced a concentrated wave of decolonization in the mid-twentieth century, settler transitions were temporally dispersed. This lack of chronological clustering makes it difficult to capture multiple settler transitions within a single continuous estimation window. While this limited sample constrains statistical power, it still allows the settler analysis to serve as a suggestive consistency check for the mechanism proposed in Hypothesis 2.

4.4.3 Donor Pool

Based on the exact same criteria as outlined for the treated group, the universal donor pool ($N_{co} = 25$, depicted in Table 4.2) is constructed. For each treatment cohort, the estimator selects an optimally weighted combination of these untreated nations to approximate the counterfactual anthropometric trajectory that cohort would have followed in the absence of decolonization.

However, as indicated by the "Donor Suitability" column, the eligibility of a specific donor depends on the temporal boundaries of the cohort being analyzed. This results in $N_{co} = 16$ potential donors for the extractive cohort and $N_{co} = 21$ for the settler cohort. Units restricted exclusively to the extractive donor pool (e.g., Japan, Finland) typically lack continuous historical data extending back to the 1850 baseline required to match the settler cohort. Conversely, units restricted exclusively to the settler donor pool fall into two distinct categories. First, some possess valid nineteenth-century data but suffer from missing observations during the twentieth-century extractive window. Second, several nations (such as Myanmar, Indonesia, Cambodia, and Vietnam) are actually members of the treated extractive cohort. Because their institutional transitions did not occur until the 1950s, they remain completely unaffected by decolonization during the 1850–1920 period. Consequently, the algorithm can leverage them as valid, untreated counterfactuals for the earlier settler transitions, maximizing the available control units.

Table 4.2: Estimation Sample: Donor Pool and Matching Eligibility

Country	**Region**	T_0 **(Rounded)**	**Donor Suitability**
Austria	Central Europe	N/A	Settler
Cambodia	Southeast Asia	1950	Settler
China	East Asia	N/A	Extractive & Settler
Czech Republic	Central Europe	N/A	Settler
Denmark	Northern Europe	N/A	Extractive & Settler
Ethiopia	Sub-Saharan Africa	N/A	Extractive
Finland	Northern Europe	N/A	Extractive
France	Western Europe	N/A	Extractive & Settler
Germany	Central Europe	N/A	Extractive & Settler
Indonesia	Southeast Asia	1950	Settler
Italy	Southern Europe	N/A	Extractive & Settler
Japan	East Asia	N/A	Extractive
Laos	Southeast Asia	1950	Settler
Malaysia *	Southeast Asia	1960	Settler
Myanmar	Southeast Asia	1950	Settler
Netherlands	Western Europe	N/A	Extractive & Settler
Norway	Northern Europe	N/A	Extractive & Settler
Portugal	Southern Europe	N/A	Extractive & Settler
Russia	Eastern Europe	N/A	Extractive & Settler
Spain	Southern Europe	N/A	Extractive & Settler
Sweden	Northern Europe	N/A	Extractive & Settler
Thailand	Southeast Asia	N/A	Extractive
Turkey	Western Asia	N/A	Settler
United Kingdom	Western Europe	N/A	Extractive & Settler
Vietnam	Southeast Asia	1950	Settler

Notes: This table presents the final donor pool. T_0 denotes the rounded independence dates (start of treatment). Units marked N/A were never formally colonized. *Donor Suitability* indicates that a country possesses continuous data and remained strictly untreated during the cohort-specific estimation window (Extractive: 1910–1980; Settler: 1850–1920). * Malaysia is included as a settler donor; the anthropometric anomaly outlined (see Section 4.4.2) is confined to the mid-twentieth century (post-1960s data) and does not compromise the integrity of its nineteenth-century observations used for the settler-period counterfactual.

5 Main Results

With the final estimation samples established, it is now possible to examine the hypotheses deduced in Chapter 2.

5.1 Hypothesis 1: Extractive Colonies

Hypothesis 1: *Because the search for a new institutional equilibrium resulted in a collapse of state capacity, the transition to independence in extractive colonies had a statistically significant negative impact on the terminal adult stature of the affected cohorts.*

Consequently, this hypothesis anticipates that following independence, the extractive cohort should exhibit either an absolute decline or a relative deceleration in average height compared to the control group.

Figure 5.1 provides an initial descriptive overview of the raw, unweighted mean height data to visually assess whether this anticipated trend is observable prior to applying any estimator.[1]

The unadjusted trajectories suggest that such a divergence did indeed occur. Following the wave of decolonization (indicated by the shaded independence period between 1950 and 1960), the anthropometric trajectory of the treated extractive cohort visibly begins to decline. In contrast, while the control pool also experiences a slight deceleration in its growth rate during this same period, its absolute trajectory continues steadily upward.

However, a visual inspection of the pre-independence period (pre-1950) reveals that the raw trends between the two groups are not strictly parallel.

[1] A complete summary of the descriptive statistics is provided in Appendix A.2.

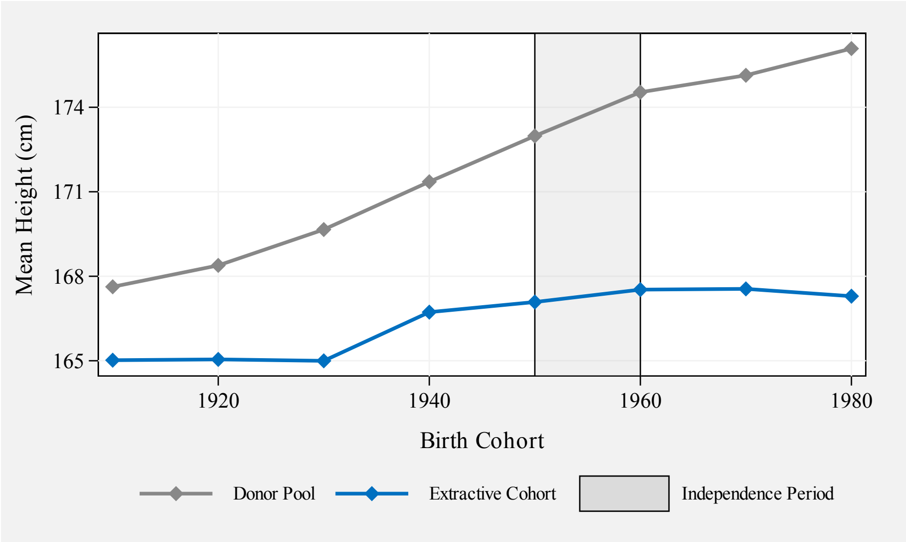


Figure 5.1: Unadjusted Anthropometric Trajectories: Extractive Cohort vs. Donor Pool

Notes: This plot illustrates the decadal mean heights for the extractive cohort ($N_{tr} = 14$) compared to the donor pool ($N_{co} = 16$). The shaded area represents the primary window during which these colonies gained independence (1950–1960).

Furthermore, because both cohorts experience a degree of post-1960 deceleration, it remains possible that the stagnation in the extractive cohort is driven by unobserved global shocks rather than the institutional transition itself. This lack of parallel pre-trends and the presence of shared global decelerations underscore the necessity of employing the SDiD approach, which systematically constructs a synthetic counterfactual to isolate the true causal effect of independence from these confounding factors.

5.1.1 SDiD Weights

Tables 5.1 and 5.2 demonstrate how these synthetic counterfactuals were constructed by systematically reweighting the identified donor countries and the available pre-treatment decades.

The third column of Table 5.1—representing the pooled unit weights ($\widehat{\omega}_i^{sdid}$) for the entire extractive sample—demonstrates that the estimator functioned as intended. The weights are evenly distributed across multiple donor countries without over-relying on a single unit, confirming that the ridge penalty successfully mitigated the risk of overfitting. Furthermore, the cohort-specific distributions in columns 1 and 2 reveal the algorithm's sensitivity to regional and structural variations. As previously noted, the 1950 cohort consists predominantly of Southeast Asian colonies; thus, the estimator assigned its highest weight to Thailand (0.101), leveraging it to capture geographically proximate characteristics. Conversely, Ethiopia is assigned a weight of zero for this Asian cohort. However, when constructing the counterfactual for the 1960 cohort—which consists entirely of African nations—Ethiopia is included as a donor (0.070). In total, 14 of the 16 primary donor countries were utilized for the 1950 cohort, while 12 were selected for the 1960 cohort.

Table 5.2 details the corresponding time weights ($\widehat{\lambda}_t^{sdid}$). Column 1 shows that the estimator anchored the synthetic control for the 1950 cohort heavily on the year 1910 ($\ell = -4$ decades prior to treatment), assigning it a weight of 0.695. This weight allocation suggests that the structural anthropometric trajectories of this early period provided the most robust predictive power for the post-treatment counterfactual. In contrast, Column 2 reveals a different temporal strategy for the 1960 cohort: the algorithm placed the majority of the weight (0.666) on 1950 ($\ell = -1$ decade). This indicates that for the African subset, the immediate pre-treatment decade offered the most stable comparative baseline. Ultimately, the weighted averages in Column 3 confirm that 1910 (0.478) and 1950 (0.381) served as the primary temporal anchors, providing the foundational comparative baseline for the pooled event-study model.

Table 5.1: Extractive Cohort: SDiD Unit Weights

Cohort 1950 ($N = 6$)			**Cohort 1960 ($N = 8$)**			**Weighted Average (Total)**		
R.	**Country**	$\widehat{\omega}_i^{sdid}$	**R.**	**Country**	$\widehat{\omega}_i^{sdid}$	**R.**	**Country**	$\widehat{\omega}_i^{sdid}$
1	Thailand	0.101	1	Portugal	0.121	1	Russia	0.108
2	Russia	0.098	2	Russia	0.116	2	Portugal	0.100
3	Spain	0.083	3	Japan	0.097	3	Spain	0.082
4	Netherlands	0.074	4	Norway	0.088	4	Japan	0.082
5	Portugal	0.071	5	Spain	0.082	5	Thailand	0.078
6	Finland	0.065	6	Italy	0.077	6	Italy	0.071
7	Italy	0.064	7	Ethiopia	0.070	7	China	0.066
8	Sweden	0.063	8	China	0.069	8	France	0.058
9	Japan	0.063	9	Thailand	0.060	9	Germany	0.050
10	China	0.063	10	France	0.055	10	Norway	0.050
11	France	0.062	11	Germany	0.054	11	United Kingdom	0.048
12	United Kingdom	0.049	12	United Kingdom	0.048	12	Ethiopia	0.040
13	Germany	0.044	N/A	Netherlands	–	13	Netherlands	0.032
14	Denmark	0.042	N/A	Finland	–	14	Finland	0.028
N/A	Norway	–	N/A	Sweden	–	15	Sweden	0.027
N/A	Ethiopia	–	N/A	Denmark	–	16	Denmark	0.018

Notes: This table presents the assigned SDiD unit weights ($\widehat{\omega}_i^{sdid}$) for the donor countries, ranked by their magnitude (R. = Rank). Because the staggered event-study design comprises two distinct treatment timings (1950 & 1960), the algorithm naturally generates separate weight distributions for the 1950 cohort, represented in column 1, and the 1960 cohort, represented in column 2. Column 3 computes a weighted average of these distributions based on the respective cohort sizes. Because the overall pooled event-study $\widehat{ATT}$ is weighted by the number of treated units observable in a given relative period (N_{tr}^{ℓ}), applying this proportional weighting to the donor pool provides the most accurate representation of the aggregate synthetic counterfactual. Missing values (–) indicate that the donor unit received a weight of zero for that specific cohort's synthetic control.

Table 5.2: Extractive Cohort: SDiD Time Weights

Cohort 1950 (Treatment $\ell = 0$)			**Cohort 1960 (Treatment $\ell = 0$)**			**Weighted Average (Total)**		
R.	**Year**	$\widehat{\lambda}_t^{sdid}$	**R.**	**Year**	$\widehat{\lambda}_t^{sdid}$	**R.**	**Year**	$\widehat{\lambda}_t^{sdid}$
1	1910 ($\ell = -4$)	0.695	1	1950 ($\ell = -1$)	0.666	1	1910	0.478
2	1920 ($\ell = -3$)	0.161	2	1910 ($\ell = -5$)	0.316	2	1950	0.381
3	1940 ($\ell = -1$)	0.144	N/A	1920 ($\ell = -4$)	–	3	1920	0.069
N/A	1950 ($\ell = 0$)	–	N/A	1940 ($\ell = -2$)	–	4	1940	0.062

Notes: This table presents the assigned SDiD time weights $\widehat{\lambda}_t^{sdid}$ for the pre-treatment years, ranked by their magnitude (R. = Rank). The same column logic as for the unit weights applies. Relative time (ℓ) denotes the distance in decades to the respective year of independence. For the 1950 cohort, 1910 represents $\ell = -4$, while for the 1960 cohort, it represents $\ell = -5$. The weighted average column reflects the chronological importance across the entire pooled sample. Missing values (–) indicate that the year received a weight of zero.

5.1.2 Event Study Results

The previous section demonstrated that the construction of the synthetic counterfactual was structurally sound. Consequently, this section evaluates the efficacy of the matching procedure in satisfying the parallel pre-treatment trends assumption and traces the path of the post-treatment anthropometric trajectory.

Figure 5.2 presents the SDiD event-study plot. The visual evidence suggests that the weighting algorithm succeeded: in the decades prior to independence, the point estimates ($\widehat{\tau}_{\ell}^{sdid}$) fluctuate tightly around the zero line. Although a slight downward drift is observable 40 to 50 years prior to treatment, the trajectory remains notably flat in the three decades immediately preceding independence ($\ell = -3$ to $\ell = -1$). Moreover, the 95% confidence intervals for all pre-treatment point estimates consistently encompass the zero line. This alignment confirms that the pre-treatment trajectories of the extractive cohort and its synthetic counterpart are statistically indistinguishable, thereby validating the identifying parallel trends assumption and granting structural credibility to the post-treatment estimates.

Following the transition to independence ($\ell \geq 0$), a divergence occurs. The post-treatment event-study estimates, along with their 95% confidence intervals, fall persistently below the zero-line, indicating a statistically significant reduction in terminal adult height attributable to decolonization. Furthermore, the magnitude of this divergence appears to compound over time. This enduring trajectory suggests that the institutional transition did not merely induce a transitory disruption, but rather inflicted a persistent, long-term biological deficit on the affected cohorts.

However, the point estimates must be interpreted with an appropriate recognition of their statistical precision due to the compact sample size ($N_{tr} = 14$). Throughout the estimation window, the confidence intervals are relatively wide, reflecting the underlying variance of the estimation. This baseline uncertainty is further exacerbated in the later post-treatment decades (especially $\ell = 3$), where the intervals widen.

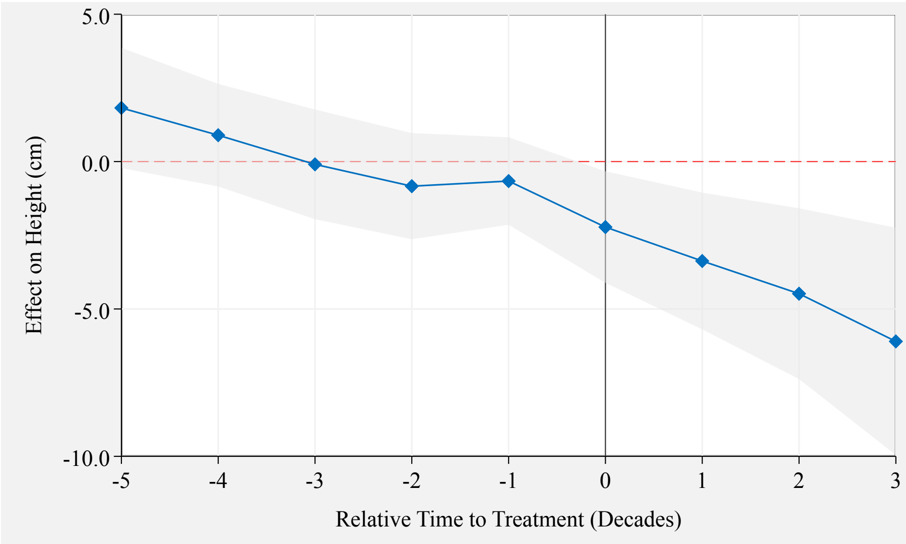


Figure 5.2: SDiD Event-Study Estimates: Extractive Cohort

Notes: This figure presents the SDiD event-study estimates for the extractive cohort. The solid markers represent the point estimates ($\hat{\tau}_{\ell}^{sdid}$) of the difference in mean terminal adult stature (in cm) between the treated cohort and its synthetic counterfactual, normalized relative to the temporal point of independence (T_0). The shaded regions represent the 95% confidence intervals.

This expansion is a direct consequence of the natural attrition of observable treated units over time. Because the historical dataset concludes in a fixed calendar year (1980), nations that transitioned to independence later in the sample inherently possess fewer observable post-treatment periods. Consequently, as the relative-time window moves further from the initial shock ($\ell = 0$), these later cohorts progressively drop out of the estimation matrix. This contraction of the effective sample size inflates the standard errors and expands the variance of the long-run estimates.

Despite this long-term variance, the formal regression results presented in Table 5.3 confirm a substantial, statistically significant collapse in adult stature. The overall $\widehat{ATT}$ reveals a height penalty of -3.70 cm ($p < 0.01$). Yet, because this aggregate parameter incorporates anthropometric data spanning up to three decades post-independence—thereby increasing the risk of capturing subsequent, unrelated macroeconomic or epidemiological shocks—the immediate treatment effect offers a more rigorously identified causal estimate. Focusing exclusively on the first birth cohort born immediately following the transition to independence ($\ell = 0$), the estimator still records a statistically significant reduction in height of −2.22 cm ($p < 0.05$).

Table 5.3: SDiD Results: Extractive Cohort — Average and Immediate Treatment Effects

Outcome: Height (cm)	**Extractive Cohort**
$\widehat{ATT}$	−3.702***
	(1.034)
Immediate Impact ($\ell = 0$)	−2.222**
	(0.976)
Sample Size (N_{tr})	14
Donor Pool Units (N_{co})	16

Notes: This table presents the estimated SDiD parameters for the extractive cohort. Row 1 reports the global $\widehat{ATT}$ averaged across all post-treatment periods. Row 2 isolates the immediate causal impact on the birth cohort born in the exact decade of transition to independence ($\ell = 0$). Standard errors are constructed using the placebo distribution derived from the untreated donor pool and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

In conclusion, the empirical results demonstrate that Hypothesis 1 cannot be rejected: the transition to independence in extractive colonies precipitated a substantial, statistically significant collapse in mean adult height. However, while these baseline estimates establish a clear link between decolonization and the decline in adult stature, they cannot definitively isolate the underlying mechanism—specifically, whether this anthropometric contraction was driven by the hypothesized institutional vacuum.

5.2 Hypothesis 2: Settler Colonies

To isolate this institutional mechanism, the subsequent step examines whether a more stable transitional environment yields a different anthropometric outcome. Thus, the analysis now turns to the second proposition formulated in the theoretical framework:

Hypothesis 2: *Because institutional continuity prevented a collapse in state capacity, the transition to independence in settler colonies did not negatively impact the terminal adult stature of the affected cohort.*

If Hypothesis 2 holds true, the settler cohort should exhibit an anthropometric trajectory that either parallels or positively diverges from the donor pool post-independence. Consistent with the analytical approach applied to the extractive cohort, Figure 5.3 provides an initial descriptive overview of the raw, unweighted mean height data to visually assess these trajectories prior to formal estimation.

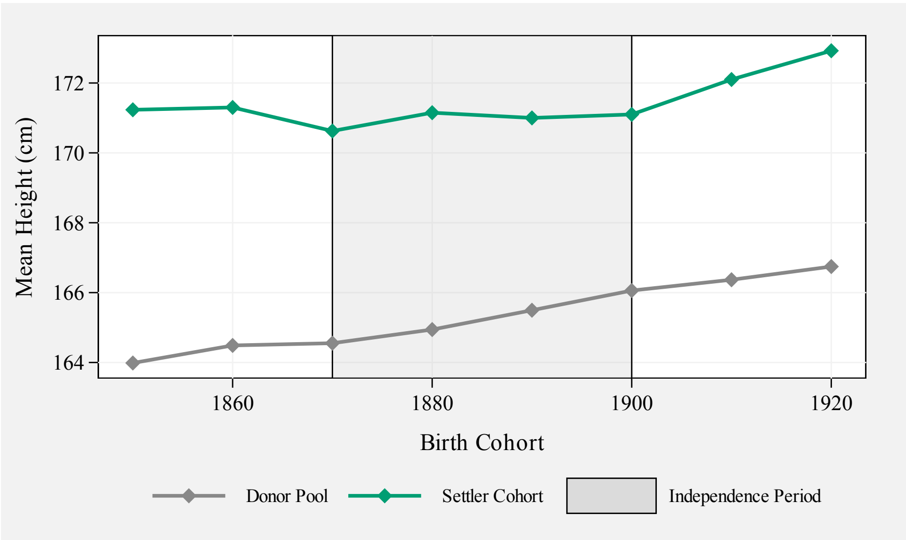


Figure 5.3: Unadjusted Anthropometric Trajectories: Settler Cohort vs. Donor Pool

Notes: The plot illustrates the decadal mean heights for the settler cohort ($N_{tr} = 2$) compared to the donor pool ($N_{co} = 21$). The shaded area represents the primary window during which these colonies gained independence (1870–1900).

The data suggest that following Canada's independence in 1870, the mean height trajectory of the settler cohort shifted upward, with growth rates exceeding those of the donor pool. Following this initial surge, a period of stagnation is observable until the independence of Australia in 1900. After this juncture, the trajectory increases again and, once more, the growth rate exceeds that of the donor pool.

In contrast, the control units exhibit no sharp structural breaks following these independence dates, maintaining a steady, secular growth trajectory throughout the entire temporal window. Descriptively, this post-independence acceleration within the settler cohort, coupled with the absence of a similar effect in the donor pool, aligns with the intuition of Hypothesis 2.

However, the plot also confirms that the raw pre-treatment trajectories between the settler nations and the donor pool are non-parallel. This baseline discrepancy underscores the continued necessity of the SDiD estimator to appropriately weight the sample and isolate the true causal effect.

5.2.1 Event Study Results

As with the extractive cohort, an analysis of the synthetic weight assignments was conducted to ensure the validity of the matching procedure. This analysis yielded no problematic irregularities, confirming the stability and reliability of the SDiD algorithm for this sub-sample. The full distribution of these synthetic weights is provided in Appendix A.3.

Figure 5.4 presents the event-study plot for the settler colonies. In the pre-treatment window, the point estimates fluctuate tightly around the zero-line, and their 95% confidence intervals consistently encompass it, confirming robust parallel trends. Notably, the pre-treatment fit for the settler cohort is substantially cleaner than that of the extractive baseline. This suggests that the settler colonies and their matched donor pool exhibit a high degree of structural equivalence, allowing the algorithm to seamlessly construct the counterfactual.

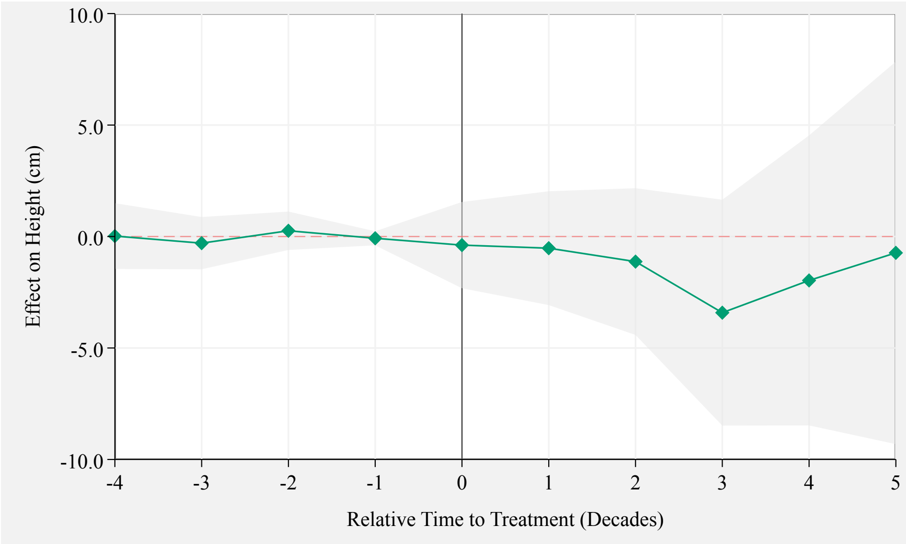


Figure 5.4: SDiD Event-Study Estimates: Settler Cohort

Notes: This figure presents the SDiD event-study estimates for the settler cohort. The solid markers represent the point estimates ($\hat{\tau}_{\ell}^{sdid}$) of the difference in mean terminal adult stature (in cm) between the treated cohort and its synthetic counterfactual, normalized relative to the temporal point of independence (T_0). The shaded regions represent the 95% confidence intervals.

Following independence, the post-treatment trajectory reveals no statistically significant decline in adult stature. The point estimates remain anchored near zero, with their 95% confidence intervals continuing to span the zero line. While the absence of a structural break is visually evident, the highly restricted sample size ($N_{tr} = 2$) inevitably causes the confidence intervals to widen following treatment. This expansion becomes substantially more pronounced from $\ell = 3$ onward; at this juncture, Australia is dropped from the estimation due to its later independence date, leaving only Canada in the treated sample. Consequently, since the immediate $\ell = 0$ to $\ell = 2$ windows still include both countries, they provide suggestive evidence of biological stability. However, because the point estimates in the later post-treatment periods reflect only a single country and thus suffer from severe variance, they are excluded from the interpretation of the results. This truncation does not compromise the empirical design, as the settler cohort serves exclusively as a mechanism test to evaluate the immediate institutional shock at independence.

The visual null effect is corroborated by the formal estimates presented in Table 5.4. The aggregate $\widehat{ATT}$ for the post-treatment window is -1.132 cm, but it is not statistically significant at any conventional level. However, because this aggregate $\widehat{ATT}$ incorporates the highly volatile later periods, applying the analytical logic utilized for the extractive cohort is essential. Isolating the immediate $\ell = 0$ impact yields a point estimate of -0.385 cm. Although this immediate estimate features diminished variance, it remains statistically indistinguishable from zero.

Table 5.4: SDiD Results: Settler Cohort — Average and Immediate Treatment Effects

Outcome: Height (cm)	**Settler Cohort**
$\widehat{ATT}$	-1.132
	(1.419)
Immediate Impact ($\ell = 0$)	-0.385
	(0.805)
Sample Size (N_{tr})	2
Donor Pool Units (N_{co})	21

Notes: This table presents the estimated SDiD parameters for the settler cohort. Row 1 reports the global $\widehat{ATT}$ averaged across all post-treatment periods. Row 2 isolates the immediate causal impact on the birth cohort born in the exact decade of transition to independence ($\ell = 0$). Standard errors are constructed using the placebo distribution derived from the untreated donor pool and reported in parentheses. Both parameters are statistically insignificant at conventional levels. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Based on these findings, Hypothesis 2 cannot be rejected: in settler colonies, the transition to independence had no detrimental impact on terminal adult stature. While a treated sample size of two precludes comprehensive global generalizations, it still provides evidence that the nature of the transition—gradual institutional continuity versus abrupt rupture—played a role in mediating the developmental consequences of decolonization.

However, relying solely on this aggregate comparison between settler and extractive colonies presents two analytical challenges.

First, it introduces the risk of spatial and ecological confounding; because the settler colonies (Australia and Canada) occupy vastly different geographic environments than the extractive cohort (Africa and Asia), disparities in climate and baseline epidemiological burdens could bias the estimates. Second, treating the extractive cohort as a homogeneous group risks masking internal dynamics. This pooling leaves open the critique that the observed decline might be driven by the idiosyncratic trajectory of a single continent rather than being a universal consequence of extractive decolonization.

To address these limitations and isolate the institutional mechanism, it is necessary to examine heterogeneity strictly within the extractive cohort. If the collapse of state capacity is indeed the primary driver of the anthropometric decline, the severity of the penalty should scale directly with the chaotic nature of the handover. This leads to the final proposition derived in Chapter 2:

Hypothesis 3: *Within extractive colonies, the magnitude of the post-independence anthropometric penalty depends upon the nature of the institutional transition. Specifically, abrupt or unplanned transitions induce a more profound biological shock compared to planned transitions.*

5.3 Hypothesis 3: Planned vs. Unplanned Transitions

5.3.1 Regional Comparison: Africa vs. Asia

As discussed in Section 4.4.2, extractive colonies exhibit distinct regional clustering, being primarily concentrated in South/Southeast Asia and Sub-Saharan Africa. This geographic divide offers an empirical setting to test Hypothesis 3. As Figure 5.5 demonstrates, the distribution of transition types differs significantly between these two regions. Within the African sub-sample, transitions were evenly split (50% planned, 50% unplanned). Conversely, the Asian sub-sample was characterized predominantly by unplanned power transfers (83% unplanned, 17% planned). Consequently, if the institutional rupture framework holds, the post-independence biological penalty should be more pronounced in Asia.

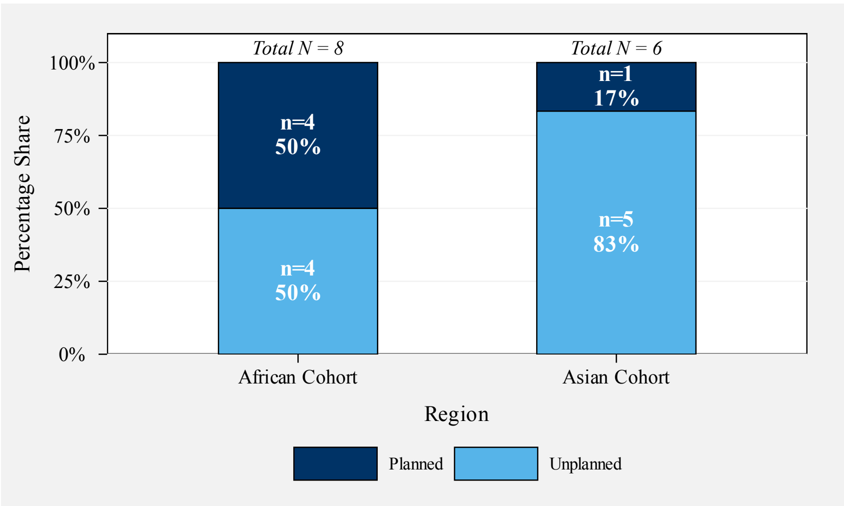


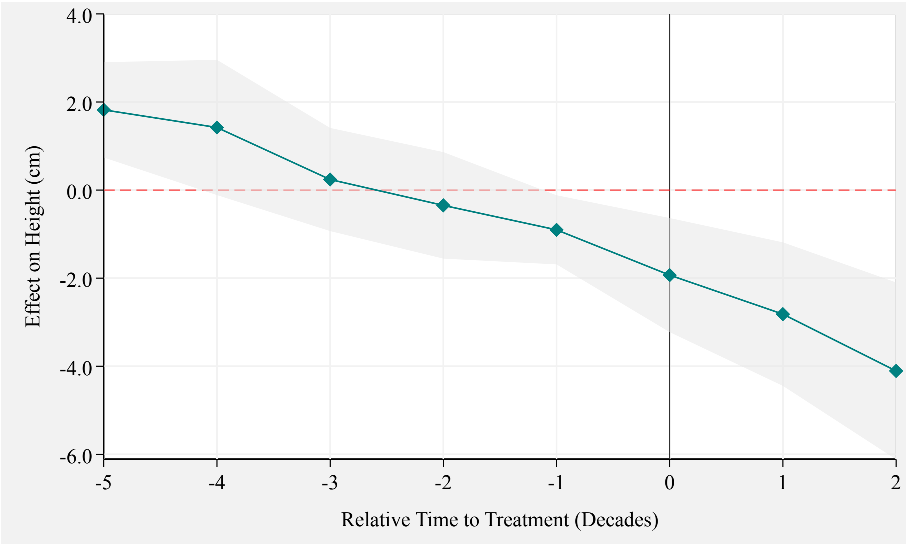
Figure 5.5: Distribution of Transition Types: Africa vs. Asia

Notes: This figure illustrates the proportion of planned vs.unplanned transitions to independence within the extractive cohort, disaggregated by geographic region (Africa and Asia). The absolute number of countries (N) and the relative percentage share are displayed within each bar. Total observations per region are indicated above the bars.

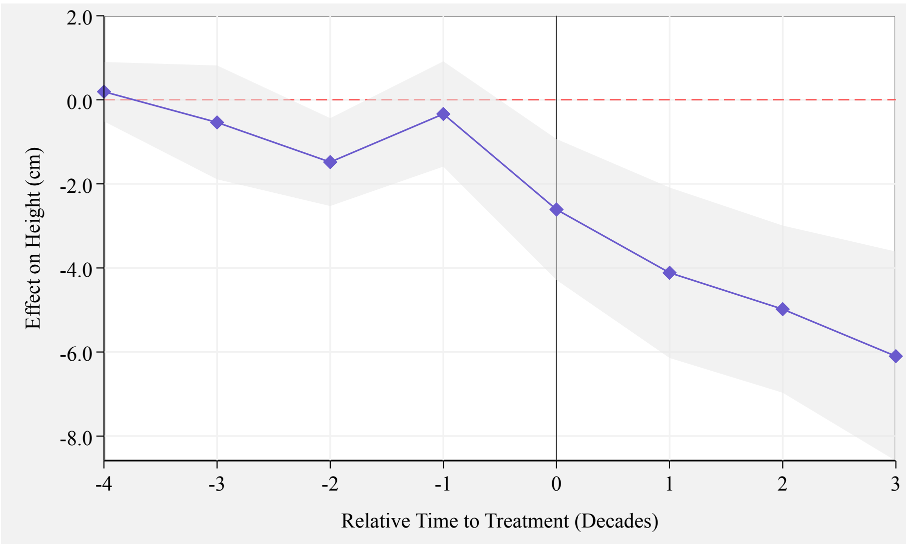
Figure 5.6 evaluates this prediction by comparing the event study plots for the African and Asian sub-samples. The validity of the synthetic matching for these geographic splits is corroborated by the 1950 and 1960 donor pool distributions presented in Tables 5.1 and 5.2, which demonstrate the stability of the assigned weights. However, visual inspection of the pre-treatment fit dictates careful interpretation. In the Asian sub-sample, the parallel trends assumption generally holds, with only a minor deviation at $\ell = -2$. In contrast, the SDiD algorithm struggled to construct a perfectly parallel counterfactual for the African sub-sample, where several pre-treatment estimates deviate significantly from zero. While earlier noise ($\ell = -5$, $\ell = -4$) is less problematic, the significant deviation at $\ell = -1$, coupled with a general downward pre-treatment trajectory, implies that the independence coefficient for Africa may partially capture a pre-existing structural decline rather than an isolated institutional shock. Consequently, the post-treatment estimates for Africa must be interpreted as an upper bound on the magnitude of the true biological penalty.

Following the shock, the coefficients for both regions either become or remain statistically negative, confirming the presence of an anthropometric penalty. Crucially, this simultaneous drop demonstrates that the decline observed in the globally pooled model was not merely an artifact driven by a single continent's idiosyncratic trajectory. Furthermore, the compounding magnitude of these negative estimates over subsequent decades suggests that—across both regions—the institutional rupture inflicted a persistent deficit rather than a transitory disruption. However, as established in the aggregate analysis, this long-term persistence must be viewed with caution; over extended post-treatment horizons, the point estimates become increasingly vulnerable to unobserved macroeconomic or epidemiological confounders.

Moving beyond visual inspection, the estimates in Table 5.5 demonstrate that the empirical results are in line with Hypothesis 3. Both the aggregate $\widehat{ATT}$ and the immediate impact in the first decade ($\ell = 0$) are markedly larger in Asia than in Africa, with all regional coefficients remaining statistically significant at conventional levels.



(a) African Cohort



(b) Asian Cohort

Figure 5.6: SDiD Event-Study Estimates: Africa vs. Asia

Notes: This figure presents the SDiD event-study estimates for the extractive cohort, disaggregated by geographic region. Panel (a) illustrates the trajectories for African colonies, while Panel (b) displays the trajectories for Asian colonies. The solid markers represent the point estimates ($\widehat{\tau}_{\ell}^{sdid}$) of the difference in mean terminal adult stature (in cm) between the treated cohort and its synthetic counterfactual, normalized relative to the temporal point of independence (T_0). The shaded regions represent the 95% confidence intervals.

Furthermore, because the African $\widehat{ATT}$ represents an upper bound on the magnitude of the true penalty, the actual biological decline in Africa was likely smaller. Crucially, this implies that the true divergence between the Asian and African sub-samples is likely even wider than these estimates suggest.

To formally assess whether the difference between the $\widehat{ATT}$ coefficients is significant, a two-sample Z-test was conducted, yielding a p-value of 0.17 ($Z = 1.37$). While this test fails to reach conventional significance thresholds, this result is mathematically expected given the lack of statistical power inherent in the restricted sample sizes ($N_{tr} = 6$ and $N_{tr} = 8$). When this lack of power is coupled with the pre-treatment noise in the African sub-sample, a cautious interpretation is required. Consequently, despite the substantial absolute divergence of 1.5 cm between the point estimates, this cross-regional difference provides suggestive, rather than statistically conclusive, evidence for Hypothesis 3.

Table 5.5: SDiD Results Comparison: Africa vs. Asia

Outcome: Height (cm)	**African Cohort**	**Asian Cohort**
$\widehat{ATT}$ (Overall)	-2.952^{**}	-4.451^{**}
	(0.667)	(0.863)
Immediate Impact ($\ell = 0$)	-1.932^{**}	-2.609^{**}
	(0.548)	(0.804)
Sample Size (N_{tr})	8	6
Donor Pool Units (N_{co})	12	14

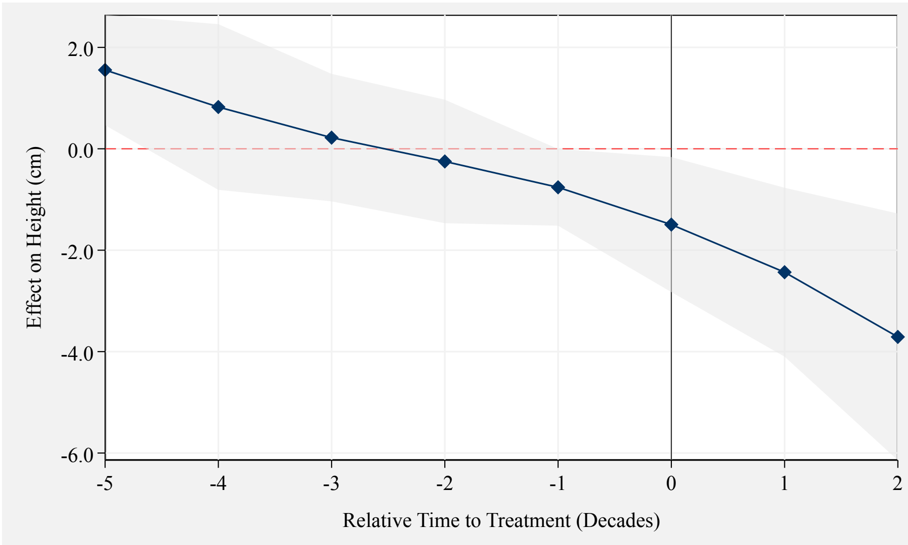
Notes: This table presents a regional decomposition of the estimated SDiD parameters to evaluate Hypothesis 3. Column 1 reports the estimates for the Sub-Saharan African sub-sample, while Column 2 isolates the South/Southeast Asian sub-sample. Row 1 reports the global $\widehat{ATT}$ averaged across all post-treatment periods, while Row 2 isolates the immediate causal impact in the exact decade of transition to independence ($\ell = 0$). Standard errors are constructed using the placebo distribution derived from the untreated donor pool and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

To conclude, this analysis suggests that the biological penalty observed in the pooled extractive model was not driven by a single continent, thereby helping to narrow down the mechanisms of anthropometric decline. Nevertheless, this cross-regional comparison remains susceptible not only to spatial and ecological confounders, but also to historical ones. For instance, Asian nations were historically characterized by a deeper legacy of centralized state control than their African counterparts (Herbst, 2000). This raises the possibility that the severity of the drop in Asia is biased by a higher baseline reliance on state infrastructure, rather than driven solely by the higher prevalence of unplanned transitions. As a result, disentangling the true effect of the transition mechanism requires holding these macro-regional baselines constant. This is achieved by examining the administrative heterogeneity strictly within the Sub-Saharan African sample. Because this sample features a perfectly balanced 50/50 split between planned and unplanned handovers (Figure 5.5), it provides an optimal empirical setting to finally isolate the effect of the transition mechanism itself.

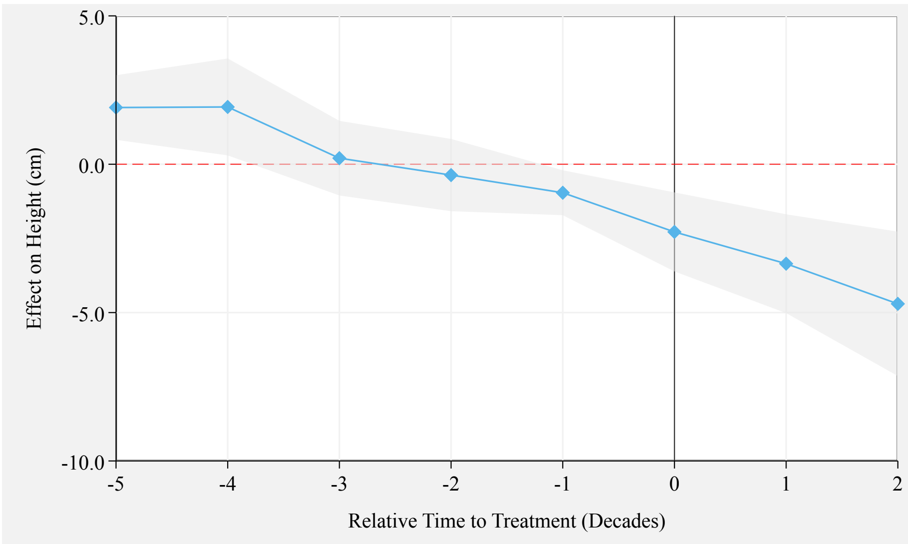
5.3.2 Transition Types within Africa

Figure 5.7 presents the event study plots for this within-Africa comparison. An analysis of the synthetic weight assignments yielded no problematic irregularities, with the full distributions provided in Appendix A.3. However, visual inspection of the plots indicates that the SDiD algorithm faced challenges with pre-treatment noise.

Both cohorts exhibit a general downward pre-treatment trajectory. In the planned group, the estimate at $\ell = -5$ deviates significantly from zero, while the unplanned group exhibits noise at $\ell = -5$, $\ell = -4$, and $\ell = -1$. Consistent with the broader African cohort analysis, this implies that the independence coefficients may partially capture a pre-existing structural decline rather than an isolated shock. Therefore, these estimates must serve as an upper bound of the true independence effect. Nevertheless, the coefficients for both groups either become or remain statistically negative post-independence, indicating an anthropometric penalty.



(a) Planned Transitions



(b) Unplanned Transitions

Figure 5.7: SDiD Event-Study Estimates: Planned vs. Unplanned Transitions

Notes: This figure presents the SDiD event-study estimates for the African extractive sub-sample, stratified by the nature of the institutional transition. Panel (a) illustrates the trajectories for planned transitions, while Panel (b) displays the trajectories for unplanned transitions. The solid markers represent the point estimates ($\widehat{\tau}_{\ell}^{sdid}$) of the difference in mean terminal adult stature (in cm) between the treated cohort and its synthetic counterfactual, normalized relative to the temporal point of independence (T_0). The shaded regions represent the 95% confidence intervals.

As Table 5.6 demonstrates, the pattern identified in the cross-regional analysis also holds true within the African sub-sample. Both the overall $\widehat{ATT}$ and the immediate impact ($\ell = 0$) are substantially larger in the unplanned group compared to the planned group, with all coefficients remaining statistically significant at conventional levels. To formally evaluate the regional magnitude of these effects, a two-sample Z-test for the difference between the $\widehat{ATT}$ coefficients yields a p-value of 0.44 ($Z = 0.77$). Also in this case, this lack of formal statistical significance is expected due to the constrained sample sizes ($N_{tr} = 4$ per group). Consequently, the evidence should be interpreted as directional, rather than statistically conclusive.

Table 5.6: SDiD Results: Planned vs. Unplanned Transitions

Outcome: Height (cm)	**Planned Cohort**	**Unplanned Cohort**
$\widehat{ATT}$ (Overall)	-2.546^{***}	-3.446^{***}
	(0.822)	(0.822)
Immediate Impact ($\ell = 0$)	-1.495^{**}	-2.281^{***}
	(0.679)	(0.679)
Sample Size (N_{tr})	4	4
Donor Pool Units (N_{co})	11	10

Notes: This table presents a decomposition of the estimated SDiD parameters for the African sub-sample to evaluate Hypothesis 3 based on the nature of the power transfer. Column 1 reports the estimates for planned transitions, while Column 2 isolates unplanned transitions. Row 1 reports the global $\widehat{ATT}$ averaged across all post-treatment periods, while Row 2 isolates the immediate causal impact in the exact decade of transition to independence ($\ell = 0$). Standard errors are constructed using the placebo distribution derived from the untreated donor pool and reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

However, despite the outlined methodological constraints, the divergence of nearly a centimeter between the two transition types provides theoretically consistent evidence for the institutional rupture mechanism. Therefore Hypothesis 3 cannot be rejected: the nature of the institutional transition within extractive colonies mattered, with unplanned transitions inducing a more severe shock than planned handovers.

Finally, to ensure that the estimated $\widehat{ATT}$ values presented in this chapter capture reliable effects rather than methodological artifacts, a series of statistical robustness checks was conducted. As detailed in Appendix B, the baseline estimates remain robust to variations in treatment timing, unit-by-unit disaggregation, “leave-one-out” checks, and alternative sample compositions. Furthermore, an “in-space” placebo test confirms that the SDiD estimator does not mechanically generate structural breaks in the absence of an actual institutional shock.

6 Conclusion

6.1 Summary of Findings

The objective of this thesis was to answer the following research question:

What is the causal impact of attaining independence from colonial rule on human stature, and can this effect be explained by the nature of the institutional transition?

To answer this, a theoretical framework was established positing that extractive colonialism represented an unstable institutional equilibrium, maintained by the coercive force of an external power. Consequently, independence constituted a severe shock to this equilibrium, triggering a collapse in state capacity and public health provision that ultimately degraded the terminal adult height of the affected cohorts. Conversely, in settler colonies characterized by inclusive institutions, the institutional equilibrium was stable; therefore, independence represented a mere formalization of this equilibrium rather than a structural break. Furthermore, the framework hypothesized that within the extractive cohort, the magnitude of the shock depended on the extent of the institutional rupture: colonies with planned transitions (e.g., through the prior integration of local bureaucrats) would experience a smaller disruption, resulting in a less severe biological penalty compared to those with unplanned, chaotic handovers.

The empirical results yielded a cohesive narrative consistent with these theoretical predictions. In the extractive colonies of Sub-Saharan Africa and Southeast Asia, decolonization precipitated a statistically significant and persistent decline in terminal adult stature. In stark contrast, the settler colonial cohort exhibited no negative anthropometric impact. Further heterogeneity analyses within the extractive cohort provided support for the institutional mechanism: the Asian sub-sample, which featured a higher concentration of unplanned transitions, suffered a steeper anthropometric decline than its African counterpart. Crucially, this pattern persisted even when holding geographic and historical confounders constant; within the Sub-Saharan African sample, unplanned transitions yielded more severe height penalties than planned ones.

Ultimately, these findings offer a compelling answer to the research question: the impact of independence on human stature was far from uniform. Instead, the anthropometric cost of decolonization depended fundamentally on the severity of the institutional rupture triggered during the transition process. In countries where the structural break was profound, a long-lasting biological penalty was incurred; however, in nations where this disruption was minimal or non-existent, the anthropometric cost was either substantially mitigated or avoided entirely.

6.2 Limitations

Nevertheless, these findings must be interpreted within the context of specific empirical constraints.

First, while the baseline extractive results ($N_{tr} = 14$) are robust—bolstered by the directional consistency of the disaggregated unit-by-unit analysis—the findings regarding the settler cohort and the intra-cohort heterogeneity should be interpreted with greater caution due to their restricted sample sizes and pre-trend uncertainties. Consequently, while the combination of the empirical evidence points decisively toward the hypothesized institutional mechanism, these specific sub-group dynamics should be understood as suggestive rather than conclusive.

Second, the temporal aggregation of the dependent variable into decadal birth cohorts and spillovers between nations (SUTVA violations) introduce a measurement constraint—namely, attenuation bias. While the methodological discussions demonstrate that this bias yields conservative, lower-bound estimates, it ultimately precludes the calculation of a strictly unbiased $\widehat{ATT}$.

Third, the external validity of these findings is bounded by the geographical scope of the available data. Because the extractive cohort is exclusively comprised of nations within Sub-Saharan Africa and South and Southeast Asia, the conclusions speak most directly to the post-1945 decolonization wave under British, French, and Belgian administration. Whether the documented anthropometric penalty generalizes to other colonial contexts—such as Latin America, where decolonization preceded this period by over a century, or the Middle East, where mandate structures produced different institutional baselines—remains unanswered within this analysis.

Finally, the reliance on aggregate national data obscures intra-national heterogeneity, particularly within settler colonies. As established in the theoretical framework, while settler populations benefited from institutional stability, indigenous groups within those same borders frequently endured localized extraction and severe marginalization. Empirically capturing this sub-national biological divergence remains unfeasible due to the scarcity of ethnically disaggregated historical anthropometric records.

6.3 Avenues for Future Research

Future research could build upon the findings and limitations of this thesis along three distinct avenues. First, as the primary empirical constraints of this analysis stem from historical data scarcity, subsequent research should prioritize the expansion of anthropometric datasets, both cross-nationally and at the sub-national, ethnically disaggregated level.

Beyond data collection, a second critical avenue is methodological. Given the inherent difficulty of assembling additional historical data, future studies could transition from SDiD to Matrix Completion estimators (e.g., Athey et al., 2021). Rather than geometrically aligning complete historical trajectories as in SDiD, these estimators recover latent factor structures from the global outcome matrix and use them to impute missing observations. Because Matrix Completion conceptualizes causal inference as an imputation problem, it does not require the balanced panels imposed by SDiD. Consequently, this methodological shift would allow researchers to exploit a larger share of the available historical record, thereby expanding the sample size and enabling the inclusion of regions beyond Sub-Saharan Africa and Asia. However, this flexibility introduces a trade-off in interpretability. Because the latent factors recovered by the algorithm are abstract mathematical constructs rather than historically observable entities, Matrix Completion operates largely as an algorithmic "black box." It therefore represents an ideal complement to the more transparent and interpretable SDiD framework employed in this thesis. If estimates derived from a broader, unbalanced dataset using Matrix Completion were to converge with the findings presented here, it would strengthen the robustness of the institutional interpretation advanced in this thesis. Nevertheless, the validity of Matrix Completion estimators depends critically on the availability of a sufficiently large and dense matrix of observed outcomes to permit reliable recovery of the latent factor structure. Future research must therefore first assess whether the currently available historical records satisfy these underlying data requirements before adopting this approach.

Finally, to triangulate the quantitative findings of this thesis from a qualitative perspective, a third promising avenue would involve conducting detailed historical case studies regarding the institutional mechanism—similar to the narrative framework established in Chapter 2—across a broader array of extractive and settler colonies.

6.4 Policy Implications and Discussion

Although the historical context of decolonization differs from modern political crises, the results of this thesis bear direct relevance for contemporary episodes of institutional rupture, such as state failure or regime collapse. The findings suggest that such systemic shocks may similarly precipitate severe and persistent deteriorations in human capital. This risk underscores the critical importance of state-building and development policies that prioritize the preservation of administrative capacity and public health infrastructure during episodes of institutional transition. Failure to do so risks compounding short-run institutional collapse into long-run welfare losses.

Furthermore, while seminal contributions in economic history have extensively documented how extractive colonial institutions shaped persistent divergence in long-run development outcomes, the anthropometric penalties identified in this analysis reveal an additional layer of historical disadvantage. Because the empirical results indicate a persistent effect rather than a transient shock, the chaotic search for a new institutional equilibrium in extractive colonies functioned as a compounding "tax" on human development. Consequently, fully understanding contemporary disparities between the Global North and the Global South requires not only examining the institutions established under colonial rule, but also accounting for the manner in which political authority was transferred at empire's end.

Ultimately, this reaffirms Faulkner's opening observation: "*The past is never dead. It's not even past.*" The institutional ruptures of decolonization are not merely historical events confined to the archives. They are biological realities, permanently inscribed in the stature of generations left in their wake.

Bibliography

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, *105*(490), 493–505. https://doi.org/10.1198/jasa.2009.ap08746

Abadie, A., & Gardeazabal, J. (2003). The Economic Costs of Conflict: A Case Study of the Basque Country. *American Economic Review*, *93*(1), 113–132. https://doi.org/10.1257/000282803321455188

Acemoglu, D. (2005). Politics and economics in weak and strong states. *Journal of Monetary Economics*, *52*(7), 1199–1226. https://doi.org/10.1016/j.jmoneco.2005.05.001

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, *91*(5), 1369–1401. https://doi.org/10.1257/aer.91.5.1369

Acemoglu, D., Johnson, S., & Robinson, J. A. (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution*. *The Quarterly Journal of Economics*, *117*(4), 1231–1294. https://doi.org/10.1162/003355302320935025

Acemoglu, D., Johnson, S., & Robinson, J. A. (2005, January 1). Chapter 6 Institutions as a Fundamental Cause of Long-Run Growth. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of Economic Growth* (pp. 385–472, Vol. 1). Elsevier. https://doi.org/10.1016/S1574-0684(05)01006-3

Adserà, A., Dalla Pozza, F., Guriev, S., Kleine-Rueschkamp, L., & Nikolova, E. (2021). Height and well-being during the transition from plan to market. *Economic Policy*, *36*(105), 77–120. https://doi.org/10.1093/epolic/eiaa030

Akresh, R., Bhalotra, S., Leone, M., & Osili, U. O. (2012). War and Stature: Growing Up during the Nigerian Civil War. *American Economic Review*, *102*(3), 273–277. https://doi.org/10.1257/aer.102.3.273

Albouy, D. Y. (2012). The Colonial Origins of Comparative Development: An Empirical Investigation: Comment. *The American Economic Review*, *102*(6), 3059–3076. Retrieved April 26, 2026, from https://www.jstor.org/stable/41724681

Alsan, M., & Goldin, C. (2019). Watersheds in Child Mortality: The Role of Effective Water and Sewerage Infrastructure, 1880–1920. *Journal of Political Economy*, *127*(2), 586–638. Retrieved April 26, 2026, from https://www.journals.uchicago.edu/doi/10.1086/700766

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic Difference-in-Differences. *American Economic Review*, *111*(12), 4088–4118. https://doi.org/10.1257/aer.20190159

Arkhangelsky, D., & Samkov, A. (2025, June 18). *Sequential Synthetic Difference in Differences*. arXiv: 2404.00164 [econ.EM]. https://doi.org/10.48550/arXiv.2404.00164

Athey, S., Bayati, M., Doudchenko, N., Imbens, G., & Khosravi, K. (2021). Matrix Completion Methods for Causal Panel Data Models. *Journal of the American Statistical Association*, *116*(536), 1716–1730. https://doi.org/10.1080/01621459.2021.1891924

Bai, J. (2009). Panel Data Models With Interactive Fixed Effects. *Econometrica*, *77*(4), 1229–1279. https://doi.org/10.3982/ECTA6135

Baten, J., & Blum, M. (2012). Growing Tall but Unequal: New Findings and New Background Evidence on Anthropometric Welfare in 156 Countries, 1810–1989. *Economic History of Developing Regions*, *27*, S66–S85. https://doi.org/10.1080/20780389.2012.657489

Baten, J., & Blum, M. (2014). Why are you tall while others are short? Agricultural production and other proximate determinants of global heights. *European Review of Economic History*, *18*(2), 144–165. https://doi.org/10.1093/ereh/heu003

Baten, J., & Blum, M. (2015). *Infant Mortality*. Retrieved June 21, 2026, from https://clio-infra.eu/Indicators/InfantMortality.html

Baten, J., & Maravall, L. (2021). The influence of colonialism on Africa's welfare: An anthropometric study. *Journal of Comparative Economics*, *49*(3), 751–775. https://doi.org/10.1016/j.jce.2021.01.001

Baten, J., Stegl, M., & Eng, P. (2013). The biological standard of living and body height in colonial and post-colonial Indonesia, 1770–2000. *Journal of Bioeconomics*, *15*(2), 103–122. Retrieved June 19, 2026, from https://ideas.repec.org//a/kap/jbioec/v15y2013i2p103-122.html

Becker, B. (2023, September 21). *Colonial Dates Dataset (COLDAT)*. Harvard Dataverse. https://doi.org/10.7910/DVN/T9SDEW

Bertocchi, G., & Canova, F. (2002). Did colonization matter for growth?: An empirical exploration into the historical causes of Africa's underdevelopment. *European Economic Review*, *46*(10), 1851–1871. Retrieved June 4, 2026, from https://ideas.repec.org//a/eee/eecrev/v46y2002i10p1851-1871.html

Besley, T., & Persson, T. (2011). *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton University Press. https://doi.org/10.2307/j.ctt1d2dmcd

Bleakley, H. (2007). Disease and Development: Evidence from Hookworm Eradication in the American South*. *The Quarterly Journal of Economics*, *122*(1), 73–117. https://doi.org/10.1162/qjec.121.1.73

Bundervoet, T., Verwimp, P., & Akresh, R. (2009). Health and Civil War in Rural Burundi. *Journal of Human Resources*, *44*(2). Retrieved April 26, 2026, from https://EconPapers.repec.org/RePEc:uwp:jhriss:v:44:y:2009:i2:p536-563

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Cameron, A. C., & Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, *50*(2), 317–372. https://doi.org/10.3368/jhr.50.2.317

Cappelen, C., & Sorens, J. (2018). Pre-colonial centralisation, traditional indirect rule, and state capacity in Africa. *Commonwealth & Comparative Politics*, *56*(2), 195–215. https://doi.org/10.1080/14662043.2017.1404666

Chancel, L., Piketty, T., Saez, E., & Zucman, G. (2022). *World Inequality Report 2022*. Harvard University Press. https://doi.org/10.2307/j.ctv3006zpt

Choi, S.-J., & Schwekendiek, D. (2009). The biological standard of living in colonial Korea, 1910-1945. *Economics and Human Biology*, *7*(2), 259–264. https://doi.org/10.1016/j.ehb.2009.04.007

Ciccia, D. (2024, November 1). *A short note on event-study synthetic difference-in-differences estimators*. arXiv: 2407.09565 [econ]. https://doi.org/10.48550/arXiv.2407.09565

Clarke, D., Pailañir, D., Carleton Athey, S., & Imbens, G. W. (2023). Synthetic Difference-in-Differences Estimation. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4346540

Cogneau, D., Dupraz, Y., & Mesplé-Somps, S. (2018). African states and development in historical perspective: Colonial public finances in British and French West. https://horizon.documentation.ird.fr/exl-doc/pleins_textes/divers18-07/010072984.pdf

Collins, M. (2017). 1. Nation, state and agency: Evolving historiographies of African decolonization. In *Britain, France and the Decolonization of Africa: Future Imperfect?* (pp. 17–42). University College London. Retrieved June 20, 2026, from https://uclpress.co.uk/book/britain-france-and-the-decolonization-of-africa/

Conley, T. G., & Taber, C. R. (2011). Inference with "Difference in Differences" with a Small Number of Policy Changes. *Review of Economics and Statistics*, *93*(1), 113–125. https://doi.org/10.1162/REST_a_00049

Correlates of War Project (CoW). (2025). *State System Membership* (Version 2024). https://correlatesofwar.org/data-sets/state-system-membership/

Cutler, D., & Miller, G. (2005). The role of public health improvements in health advances: The twentieth-century United States. *Demography*, *42*(1), 1–22. https://doi.org/10.1353/dem.2005.0002

Deaton, A. (2003). Health, Inequality, and Economic Development. *Journal of Economic Literature*, *41*(1), 113–158. https://doi.org/10.1257/002205103321544710

Deaton, A. (2007). Height, health, and development. *Proceedings of the National Academy of Sciences*, *104*(33), 13232–13237. https://doi.org/10.1073/pnas.0611500104

Dharmapala, D., & Suesse, M. (2025, July 1). *Decolonization, Legitimacy and Fiscal Capacity: Event Study Evidence from Africa*. 5392936. https://doi.org/10.2139/ssrn.5392936

Easterly, W., & Levine, R. (2016). The European origins of economic development. *Journal of Economic Growth*, *21*(3), 225–257. https://doi.org/10.1007/s10887-016-9130-y

Engerman, S. L., & Sokoloff, K. L. (2002, October). *Factor Endowments, Inequality, and Paths of Development Among New World Economics*. 9259. https://doi.org/10.3386/w9259

Fadakinte, M. M. (2020). The Struggle for Power in Post-Colonial Africa: Politics without Hegemony and the State. *Advances in Social Sciences Research Journal*, *7*(2), 133–146. https://doi.org/10.14738/assrj.72.7765

Ferman, B., & Pinto, C. (2019). Inference in Differences-in-Differences with Few Treated Groups and Heteroskedasticity. *The Review of Economics and Statistics*, *101*(3), 452–467. https://doi.org/10.1162/rest_a_00759

Hatton, T. J., & Bray, B. E. (2010). Long run trends in the heights of European men, 19th–20th centuries. *Economics & Human Biology*, *8*(3), 405–413. https://doi.org/10.1016/j.ehb.2010.03.001

Head, K., Mayer, T., & Ries, J. (2010). The erosion of colonial trade linkages after independence. *Journal of International Economics*, *81*(1), 1–14. https://doi.org/10.1016/j.jinteco.2010.01.002

Hendrix, C. S. (2010). Measuring state capacity: Theoretical and empirical implications for the study of civil conflict. *Journal of Peace Research*, *47*(3), 273–285. https://doi.org/10.1177/0022343310361838

Herbst, J. (2000). *States and Power in Africa: Comparative Lessons in Authority and Control - Second Edition* (REV - Revised, 2, Second Edition). Princeton University Press. Retrieved May 5, 2026, from https://www.jstor.org/stable/j.ctt9qh05m

Heyberger, L. (2022). A great colonial divergence in biological standards of living? The secular trend of stature among indigenous people and settlers in Algeria, 1880-1940. *Popolazione e storia*, *22*(2), 33–57. https://doi.org/10.4424/1497

Iyer, L. (2010). Direct versus Indirect Colonial Rule in India: Long-Term Consequences. *The Review of Economics and Statistics*, *92*(4), 693–713. https://doi.org/10.1162/REST_a_00023

Jones, C. I., & Klenow, P. J. (2016). Beyond GDP? Welfare across Countries and Time. *American Economic Review*, *106*(9), 2426–2457. https://doi.org/10.1257/aer.20110236

Komlos, J. (1998). Shrinking in a Growing Economy? The Mystery of Physical Stature during the Industrial Revolution. *The Journal of Economic History*, *58*(3), 779–802. https://doi.org/10.1017/S0022050700021161

Lange, M., Mahoney, J., & Vom Hau, M. (2006). Colonialism and Development: A Comparative Analysis of Spanish and British Colonies. *American Journal of Sociology*, *111*(5), 1412–1462. https://doi.org/10.1086/499510

Lee, A., & Paine, J. (2019). What Were the Consequences of Decolonization?: RESEARCH NOTE. *International Studies Quarterly*, *63*(2), 406–416. Retrieved June 4, 2026, from https://www.jstor.org/stable/48619262

Massey, D. S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., & Taylor, J. E. (1993). Theories of International Migration: A Review and Appraisal. *Population and Development Review*, *19*(3), 431–466. https://doi.org/10.2307/2938462

Meredith, M. (2013). *The State of Africa: A History of the Continent Since Independence*. Simon & Schuster Ltd. https://www.intuganda.org/wp-content/uploads/kanzu_intug/class24/The-State-of-Africa-A-History-of-the-Continent-Since-Independence-Martin-Meredith.pdf

Milanovic, B. (2016). *Global Inequality: A New Approach for the Age of Globalization*. Harvard University Press. https://doi.org/10.2307/j.ctvjghwk4

Minoiu, C., & Shemyakina, O. N. (2014). Armed conflict, household victimization, and child health in {Côte d'Ivoire}. *Journal of Development Economics*, *108*, 237–255. https://doi.org/10.1016/j.jdeveco.2014.03.003

Moradi, A. (2010). Nutritional status and economic development in sub-Saharan Africa, 1950-1980. *Economics and Human Biology*, *8*(1), 16–29. https://doi.org/10.1016/j.ehb.2009.12.002

Moradi, A., Austin, G., & Baten, J. (2013). Heights and Development in a Cash-Crop Colony: Living Standards in Ghana, 1870-1980. *African Economic History Working Paper*, (7/2013). Retrieved June 19, 2026, from https://ideas.repec.org//p/hhs/afekhi/2013_007.html

NCD Risk Factor Collaboration (NCD-RisC). (2016). A century of trends in adult human height. *eLife*, *5*, e13410. https://doi.org/10.7554/eLife.13410

North, D. C. (1990, October 26). *Institutions, Institutional Change and Economic Performance*. Cambridge University Press.

Nunn, N. (2009, April). *The Importance of History for Economic Development*. 14899. https://doi.org/10.3386/w14899

OECD. (2014, October 2). *How Was Life?: Global Well-being since 1820* (J. L. Van Zanden, J. Baten, M. Mira d'Ercole, A. Rijpma, & M. P. Timmer, Eds.). https://doi.org/10.1787/9789264214262-en

Ogle, V. (2020). 'Funk Money': The End of Empires, The Expansion of Tax Havens, and Decolonization as an Economic and Financial Event*. *Past & Present*, *249*(1), 213–249. https://doi.org/10.1093/pastj/gtaa001

Olsson, O. (2009). On the democratic legacy of colonialism. *Journal of Comparative Economics*, *37*(4), 534–551. https://doi.org/10.1016/j.jce.2009.08.004

Parkin, G. R. (1909). The railway development of Canada. *Scottish Geographical Magazine*, *25*(5), 225–250. https://doi.org/10.1080/00369220908733973

Piketty, T., & Goldhammer, A. (2014). *Capital in the Twenty-First Century*. Harvard University Press. Retrieved June 19, 2026, from https://www.jstor.org/stable/j.ctt6wpqbc

Pomeranz, K. (2000). *The Great Divergence: China, Europe, and the Making of the Modern World Economy*. Princeton University Press. https://doi.org/10.2307/j.ctt7sv80

Porreca, Z. (2022). Synthetic difference-in-differences estimation with staggered treatment timing. *Economics Letters*, *220*, 110874. https://doi.org/10.1016/j.econlet.2022.110874

Pritchett, L. (1997). Divergence, Big Time. *Journal of Economic Perspectives*, *11*(3), 3–17. https://doi.org/10.1257/jep.11.3.3

Reno, W. (2006). Congo: From State Collapse to 'Absolutism', to State Failure. *Third World Quarterly*, *27*(1), 43–56. Retrieved April 26, 2026, from https://www.jstor.org/stable/4017658

Rodrik, D. (2011, February 21). *The Globalization Paradox: Democracy and the Future of the World Economy*. W. W. Norton & Company.

Roser, M., Appel, C., & Ritchie, H. (2021, May 1). *Human Height*. Our World in Data. Retrieved June 21, 2026, from https://ourworldindata.org/human-height

Steckel, R. H. (1995). Stature and the Standard of Living. *Journal of Economic Literature*, *33*(4), 1903–1940. Retrieved June 20, 2026, from https://ideas.repec.org//a/aea/jeclit/v33y1995i4p1903-1940.html

Steckel, R. H. (2009). Heights and human welfare: Recent developments and new directions. *Explorations in Economic History*, *46*(1), 1–23. https://doi.org/10.1016/j.eeh.2008.12.001

Stiglitz, P. J. E., Sen, P. A., & Fitoussi, P. J.-P. (2009). Report by the Commission on the Measurement of Economic Performance and Social Progress. https://ec.europa.eu/eurostat/documents/8131721/8131772/Stiglitz-Sen-Fitoussi-Commission-report.pdf

Strang, D. (1990). From Dependency to Sovereignty: An Event History Analysis of Decolonization 1870-1987. *American Sociological Review*, *55*(6), 846–860. https://doi.org/10.2307/2095750

Subrahmanyam, G. (2006). Ruling continuities: Colonial rule, social forces and path dependence in British India and Africa. *Commonwealth & Comparative Politics*, *44*(1), 84–117. https://doi.org/10.1080/14662040600624478

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199. https://doi.org/10.1016/j.jeconom.2020.09.006

Thelin, J. R. (2004, June 11). *A history of {American} higher education*. JHU Press.

Thomas, D., & Strauss, J. (1992). Prices, infrastructure, household characteristics and child height. *Journal of Development Economics*, *39*(2), 301–331. https://doi.org/10.1016/0304-3878(92)90042-8

Tödt, D. (2021). *The Lumumba Generation: African Bourgeoisie and Colonial Distinction in the Belgian Congo*. De Gruyter. https://doi.org/10.1515/9783110709308
Accepted: 2021-12-09T10:08:08Z.

Vander Eycken, H., & Vander Vorst, F. (1967). Le blocage de la croissance en République démocratique du Congo. *Revue Tiers Monde*, *8*(30), 411–434. https://doi.org/10.3406/tiers.1967.2361

Veenman, M., Stefan, A. M., & Haaf, J. M. (2024). Bayesian hierarchical modeling: An introduction and reassessment. *Behavior Research Methods*, *56*(5), 4600–4631. https://doi.org/10.3758/s13428-023-02204-3

Wallerstein, I. (2011, June 10). *The Modern World-System I: Capitalist Agriculture and the Origins of the European World-Economy in the Sixteenth Century*. University of California Press.

Wimmer, A., & Min, B. (2006). From Empire to Nation-State: Explaining Wars in the Modern World, 1816-2001. *American Sociological Review*, *71*(6), 867–897. http://www.jstor.org/stable/25472435

Wolfe, P. (2006). Settler colonialism and the elimination of the native. *Journal of Genocide Research*, *8*(4), 387–409. https://doi.org/10.1080/14623520601056240

Xu, Y. (2017). Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models. *Political Analysis*, *25*(1), 57–76. https://doi.org/10.1017/pan.2016.2

Zijdeman, R. (2015). *Life Expectancy at Birth (Total)*. Retrieved June 21, 2026, from https://clio-infra.eu/Indicators/LifeExpectancyatBirthTotal.html

Ziltener, P., & Künzler, D. (2017). Research Note: Measuring the Impacts of Colonialism: A New Data Set for the Countries of Africa and Asia. *Journal of World-Systems Research*, *23*(1), 156–190. https://doi.org/10.5195/jwsr.2017.683

A Data Properties and Baseline Calibration

A.1 Master Sample

Table A.1: Overview: Sample Composition, Independence Timing, and Transition Mechanisms

Country	**Region**	**Indep. (Mean)**	**Indep. (Max)**	T_0 **(Rounded)**	**Transition**
		COLONIZED (N = 16)			
Extractive Colonies (N = 14)					
Burkina Faso	Sub-Saharan Africa	1960	1960	1960	Unplanned
Ghana	Sub-Saharan Africa	1957	1957	1960	Planned
Ivory Coast	Sub-Saharan Africa	1960	1960	1960	Planned
Mali	Sub-Saharan Africa	1960	1960	1960	Unplanned
Niger	Sub-Saharan Africa	1960	1960	1960	Unplanned
Nigeria	Sub-Saharan Africa	1960	1960	1960	Planned
Senegal	Sub-Saharan Africa	1960	1960	1960	Unplanned
Tanzania	Sub-Saharan Africa	1961	1963	1960	Planned
Cambodia	Southeast Asia	1952	1953	1950	Unplanned
India	South Asia	1947	1947	1950	Planned
Indonesia	Southeast Asia	1952	1962	1950	Unplanned
Myanmar	Southeast Asia	1948	1948	1950	Unplanned
Laos	Southeast Asia	1952	1954	1950	Unplanned
Vietnam	Southeast Asia	1951	1954	1950	Unplanned
Settler Colonies (N = 2)					
Canada	North America	1877	1919	1870*	–
Australia	Oceania	1904	1919	1900	–
		NEVER COLONIZED (N = 19)			
Austria	Central Europe	–	–	–	–
China	East Asia	–	–	–	–
Czech Republic	Central Europe	–	–	–	–
Denmark	Northern Europe	–	–	–	–
Ethiopia	Sub-Saharan Africa	–	–	–	–
Finland	Northern Europe	–	–	–	–
France	Western Europe	–	–	–	–
Germany	Central Europe	–	–	–	–

Continued on next page

Table A.1 – *Continued from previous page*

Country	**Region**	**Indep. (Mean)**	**Indep. (Max)**	T_0 **(Rounded)**	**Transition**
Italy	Southern Europe	–	–	–	–
Japan	East Asia	–	–	–	–
Netherlands	Western Europe	–	–	–	–
Norway	Northern Europe	–	–	–	–
Portugal	Southern Europe	–	–	–	–
Russia	Eastern Europe	–	–	–	–
Spain	Southern Europe	–	–	–	–
Sweden	Northern Europe	–	–	–	–
Thailand	Southeast Asia	–	–	–	–
Turkey	Western Asia	–	–	–	–
United Kingdom	Western Europe	–	–	–	–

Notes: This table provides a comprehensive overview of the estimation sample. The sample is stratified into *Colonized* and *Never Colonized* units, with the former further divided into extractive and settler institutional regimes. *Indep. (Mean)* indicates the simple unweighted average used in the COLDAT database (Becker, 2023) to aggregate underlying historical sources (Correlates of War Project (CoW), 2025; Lange et al., 2006; Olsson, 2009; Wimmer & Min, 2006), while *(Max)* denotes the latest recorded sovereignty year across those sources. T_0 signifies the rounded start of treatment employed in this paper. The asterisk (*) for Canada indicates that the treatment decade was operationalized using the majority-consensus rule (1867; see Section **??**) rather than the COLDAT mean to avoid a historically vacant "phantom date." *Transition* classifies the mode of power transfer into planned and unplanned trajectories.

A.2 Descriptive Statistics

Table A.2: Summary Statistics: Treated Cohorts and Global Control Pools

Sample	**Units** (N)	**Periods** (T)	**Total Obs.** ($N \times T$)	**Mean (cm)**	**Std. Dev.**	**Min**	**Max**
Panel A: Extractive Cohort & Global Control Pool							
Extractive Cohort	14	8	112	166.4	4.7	158.0	174.7
Control Pool (Extr.)	16	8	128	171.9	5.6	159.6	183.2
Panel B: Settler Cohort & Global Control Pool							
Settler Cohort	2	8	16	171.4	1.0	169.9	173.0
Control Pool (Settler)	21	8	168	165.3	4.4	156.7	176.3
Panel C: Regional Sub-Samples (Extractive Cohort Only)							
South/Southeast Asia	6	8	48	161.7	2.0	158.0	165.0
Sub-Saharan Africa	8	8	64	169.5	2.6	164.6	174.7
Panel D: Transition Sub-Samples (African Cohort Only)							
Planned Transition	4	8	32	168.0	1.9	164.6	171.3
Unplanned Transition	4	8	32	171.9	1.3	168.5	174.7

Notes: This table presents pooled summary statistics for the height outcome variable. Panels A and B outline the primary treatment cohorts (extractive and settler) and their respective global donor pools. Panels C and D decompose the extractive cohort by geographical region, and the African sub-cohort by transition mode, respectively. Descriptive statistics for the control pools are omitted in the sub-sample panels, as the SDiD algorithm mathematically constructs counterfactuals from the full global donor pools regardless of the specific treatment sub-category.

A.3 Supplementary SDiD Weight Distributions

Tables A.3 and A.4 present the assigned synthetic weights for the settler cohort and the transition types (planned vs. unplanned) within the African sub-sample.

For the settler cohort, the unit weights are distributed across all 21 available donor countries, indicating that the ridge penalty successfully mitigated the risk of overfitting by preventing overreliance on a single unit. The two highest-weighted donors are Laos (0.090) and Indonesia (0.074). Given that the settler cohort includes Australia, the algorithm's reliance on geographically proximate Indo-Pacific nations suggests that it accurately captured underlying latent regional factors. The time weights are distributed between 1850 and 1890, with the heaviest emphasis placed on 1860 (0.418) and 1890 (0.319). This indicates that these specific decades provided the strongest structural predictive power for anchoring the post-independence trajectory.

Regarding the African sub-sample, the algorithm selected an almost identical set of donor countries to construct the counterfactuals for both transition types, albeit with slightly varying weight distributions. Because both sub-samples consist entirely of nations located in Sub-Saharan Africa, this strong overlap of donor countries is theoretically consistent and further confirms that the estimator effectively optimized on shared latent geographic and structural factors. Furthermore, the temporal optimization yielded identically distributed time weights for both groups. The algorithm anchored the counterfactuals on the immediate pre-treatment decade of 1950 (0.666) and the earlier baseline of 1910 (0.316), suggesting that these two specific periods provided the most robust historical baseline for predicting post-independence anthropometric outcomes across the region.

Table A.3: Settler Cohort: SDiD Unit and Time Weights

Panel A: Unit Weights (ω_i^{sdid})

R.	**Country**	**Weight**	**R.**	**Country**	**Weight**
1	Laos	0.090	12	Cambodia	0.038
2	Indonesia	0.074	13	Russia	0.038
3	China	0.068	14	Portugal	0.037
4	Spain	0.062	15	Denmark	0.034
5	Malaysia	0.053	16	Netherlands	0.029
6	France	0.053	17	United Kingdom	0.027
7	Italy	0.051	18	Myanmar	0.024
8	Vietnam	0.046	19	Turkey	0.024
9	Norway	0.042	20	Sweden	0.023
10	Austria	0.039	21	Germany	0.023
11	Czech Republic	0.039			

Panel B: Time Weights (λ_t^{sdid})

R.	**Year**	**Weight**	**R.**	**Year**	**Weight**
1	1860	0.418	3	1880	0.182
2	1890	0.319	4	1850	0.082

Notes: This table presents the SDiD weights assigned to the settler cohort ($N = 2$), ranked by their magnitude (R. = Rank). Panel A details the averaged distribution of the unit weights (ω_i^{sdid}) across the two units. Panel B displays the averaged distribution of the time weights (λ_t^{sdid}).

Table A.4: Planned vs. Unplanned Transitions: SDiD Unit and Time Weights

Panel A: Unit Weights (ω_i^{sdid})

Planned Transition (N = 4)			**Unplanned Transition (N = 4)**		
R.	**Country**	**Weight**	**R.**	**Country**	**Weight**
1	Russia	0.130	1	Portugal	0.149
2	Portugal	0.127	2	Russia	0.132
3	Japan	0.098	3	Japan	0.112
4	Norway	0.083	4	Norway	0.106
5	Spain	0.080	5	Spain	0.089
6	Italy	0.077	6	Ethiopia	0.084
7	China	0.076	7	Italy	0.080
8	Thailand	0.068	8	China	0.060
9	Ethiopia	0.058	9	Germany	0.050
10	France	0.056	10	Thailand	0.046
11	United Kingdom	0.052			

Panel B: Time Weights (λ_t^{sdid})

Planned Transition			**Unplanned Transition**		
R.	**Year**	**Weight**	**R.**	**Year**	**Weight**
1	1950	0.666	1	1950	0.666
2	1910	0.316	2	1910	0.316

Notes: This table presents the SDiD weights assigned to the African sub-cohorts based on their transition type, ranked by their magnitude (R. = Rank). Panel A details the weighted averaged distribution of the unit weights (ω_i^{sdid}) within each sub-sample. Panel B displays the weighted averaged distribution of the time weights (λ_t^{sdid}).

B Robustness Checks

B.1 Treatment Timing Sensitivity Check

Section 3.3.2 established that utilizing the alternative "maximum" independence date shifts the assigned treatment timing (T_0) for three specific countries in the sample: Indonesia (1950 to 1960), Australia (1900 to 1920), and Canada (1870 to 1920). Figure B.1 presents a sensitivity analysis comparing the baseline models against alternative specifications utilizing these maximum dates.

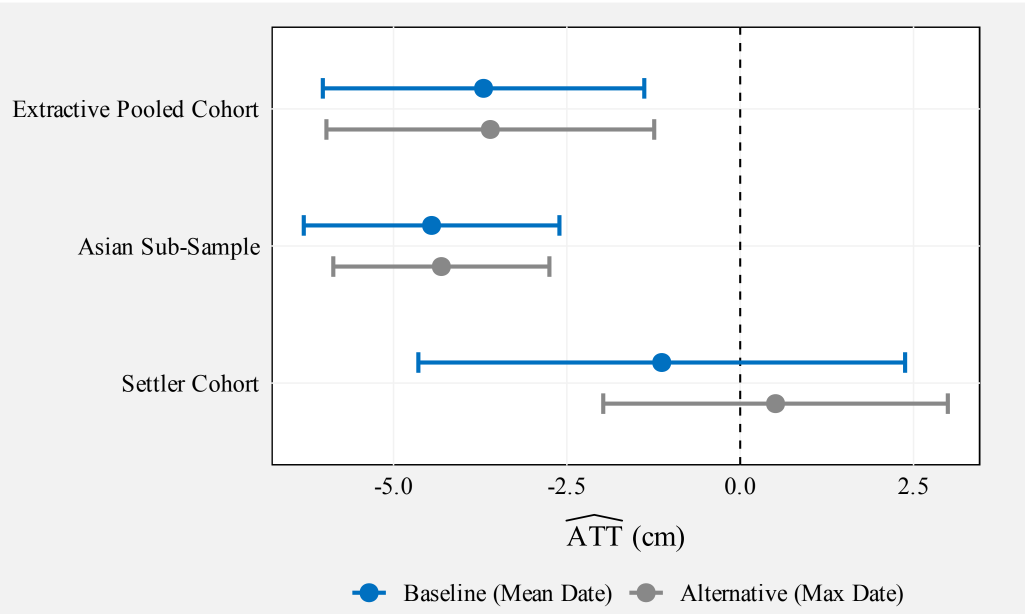


Figure B.1: Sensitivity to Treatment Timing: Mean vs. Max Date

Notes: This figure compares the $\widehat{ATT}$ when using the baseline specification (mean independence date) versus an alternative specification (maximum independence date). The blue dots represent the baseline, while the grey dots represent the alternative. Estimates are shown for the main pooled cohort, the Asian sub-sample, and the settler cohort. Models restricted exclusively to the African sub-sample were not re-estimated, as none of the constituent countries experienced decadal date deviations. Error bars represent 95% confidence intervals. The vertical dashed line indicates a treatment effect of zero.

The plot demonstrates that the primary findings remain highly robust to changes in treatment timing. For both the main pooled extractive cohort and the Asian sub-sample, the point estimates are virtually unchanged and statistically significant, with their 95% confidence intervals consistently falling well below the zero line.

Similarly, the findings for the settler cohort demonstrate comparable stability; the confidence intervals in both specifications cross the zero line, confirming a persistent null effect. However, the alternative specification shifts the point estimate for the settler cohort from a negative to a positive value. The underlying mechanics driving this specific directional shift are explored in detail in the subsequent unit-by-unit analysis.

B.2 Unit-by-Unit Analysis

As discussed in Chapter 5, a primary limitation of the pooled models is the constrained sample size imposed by data availability and the strict balanced panel requirements. To enhance the robustness of the baseline estimates, a disaggregated, unit-by-unit analysis was conducted. By relaxing the panel restrictions, this approach allowed for the inclusion of a broader set of countries, utilizing all nations that fulfilled the baseline data requirements outlined in Section 4.4.2. Consequently, the number of treated units in the extractive cohort expanded from 14 to 23. Similarly, the settler cohort increased from two to three units, allowing for the inclusion of Argentina (generally regarded as a quasi-settler colony, see Easterly & Levine, 2016). In addition to increasing the sample size, disaggregating the effects also provides a direct test of whether the results of the pooled model were driven by extreme outliers or if they represent a consistent, structural pattern across the colonies.

The results of this unit-by-unit analysis are visualized in the coefficient plots for both the mean-date (Figure B.2) and max-date (Figure B.3) independence specifications. The estimates demonstrate strong directional consistency across the extractive cohort. Regardless of the date utilized, all 23 extractive point estimates are negative, with their respective confidence intervals heavily skewed away from zero. While not every individual country achieves statistical significance—a common artifact of reduced statistical power at the single-unit level—this widespread downward trajectory confirms that the baseline extractive findings hold across a larger sample size and were not driven by outliers.

Conversely, the individual estimates for the settler colonies reveal important nuances regarding treatment timing. Under the mean-date specification (Figure B.2), Australia and Argentina yield the theoretically expected null effects, but Canada emerges as a heavily negative anomaly. However, when applying the max-date specification (Figure B.3), this anomaly disappears entirely, shifting Canada's estimate upward and aligning all three settler colonies with the theoretical expectation of no severe anthropometric penalty.

This result indicates three critical insights. First, it shows that the pooled null effect observed for the settler cohort in the main analysis (which relied on the mean date) was actually dragged downward by Canada. This clarifies not only the negative $\widehat{ATT}$ of the baseline model, but also the directional shift highlighted in the preceding timing sensitivity check: the pooled settler estimate flipped to a positive value because Canada's downward pull was eliminated under the max-date specification. Second, it underscores that for Canada—where the underlying COLDAT sources diverge significantly on the timing of sovereignty—the exact definitional choice of independence becomes highly consequential. Third, it explicitly illustrates why the settler findings must be interpreted as suggestive: in a highly restricted sample setting (like in this case of $N_{tr} = 2$ or $N_{tr} = 3$ when including Argentina), the Law of Large Numbers does not apply. Without the capacity for individual variances to converge toward a true expected value, it becomes difficult to definitively distinguish which countries represent isolated outliers and which capture a broader structural pattern. Consequently, looking strictly at the unit-by-unit mean-date specification, it cannot be definitively concluded whether the negative Canadian trajectory or the stable Australian and Argentine trajectories more accurately reflect the true anthropometric consequence of settler decolonization. However, because the pooled settler model yields an overall null effect and the max-date specification fully aligns with the theoretical predictions, the majority of evidence points toward Canada's mean-date trajectory being an isolated outlier in this case.

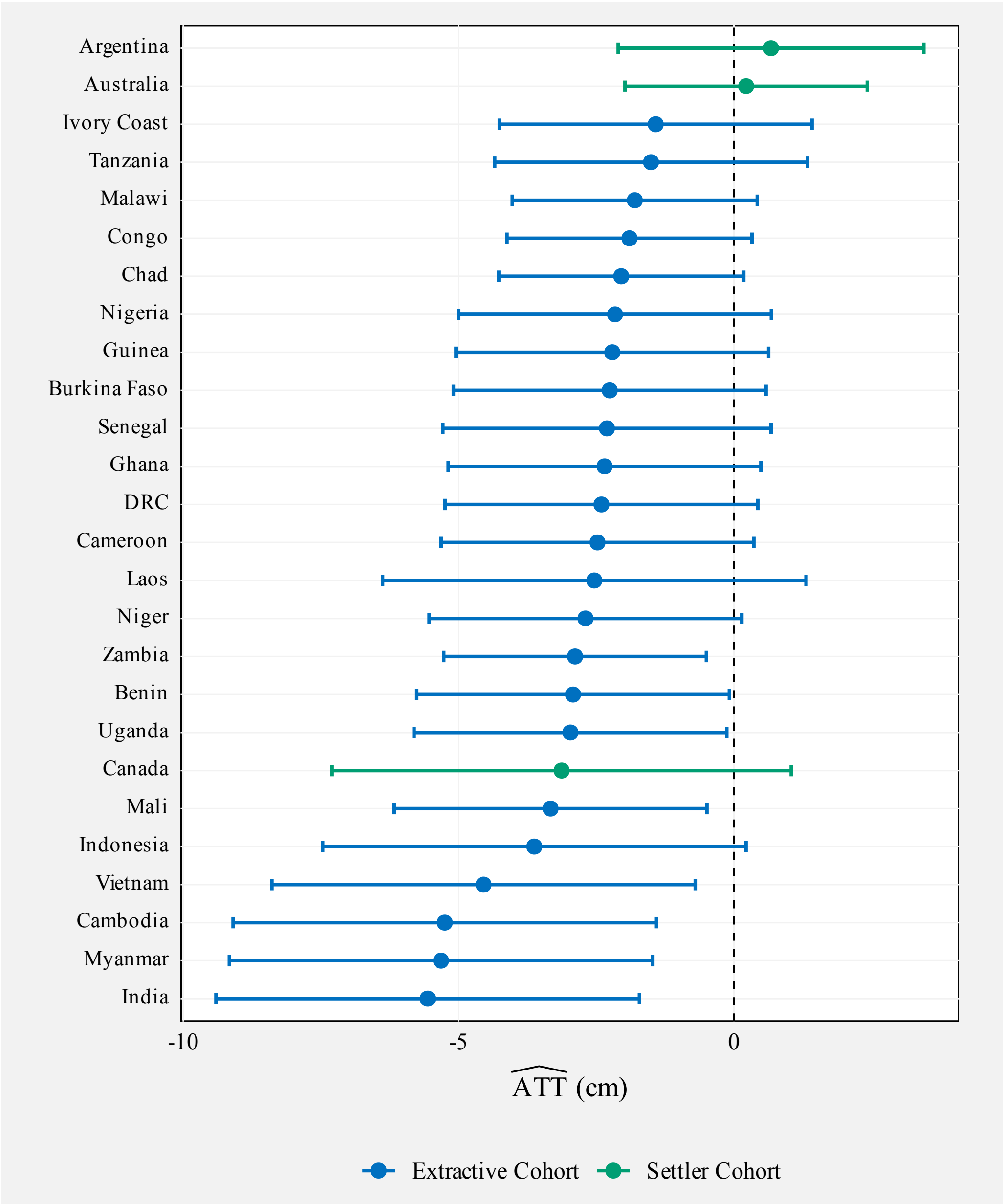


Figure B.2: Country-Specific $\widehat{ATT}$ Estimates: Mean Date

Notes: This figure presents disaggregated, country-specific SDiD point estimates ($\widehat{ATT}$) ordered by the magnitude of the treatment effect. The specification utilizes the baseline approach (mean independence dates) on an expanded sample ($N_{tr} = 23$ for the extractive cohort, blue markers; $N_{tr} = 3$, green markers). Error bars represent 95% confidence intervals. The vertical dashed line indicates a treatment effect of zero.

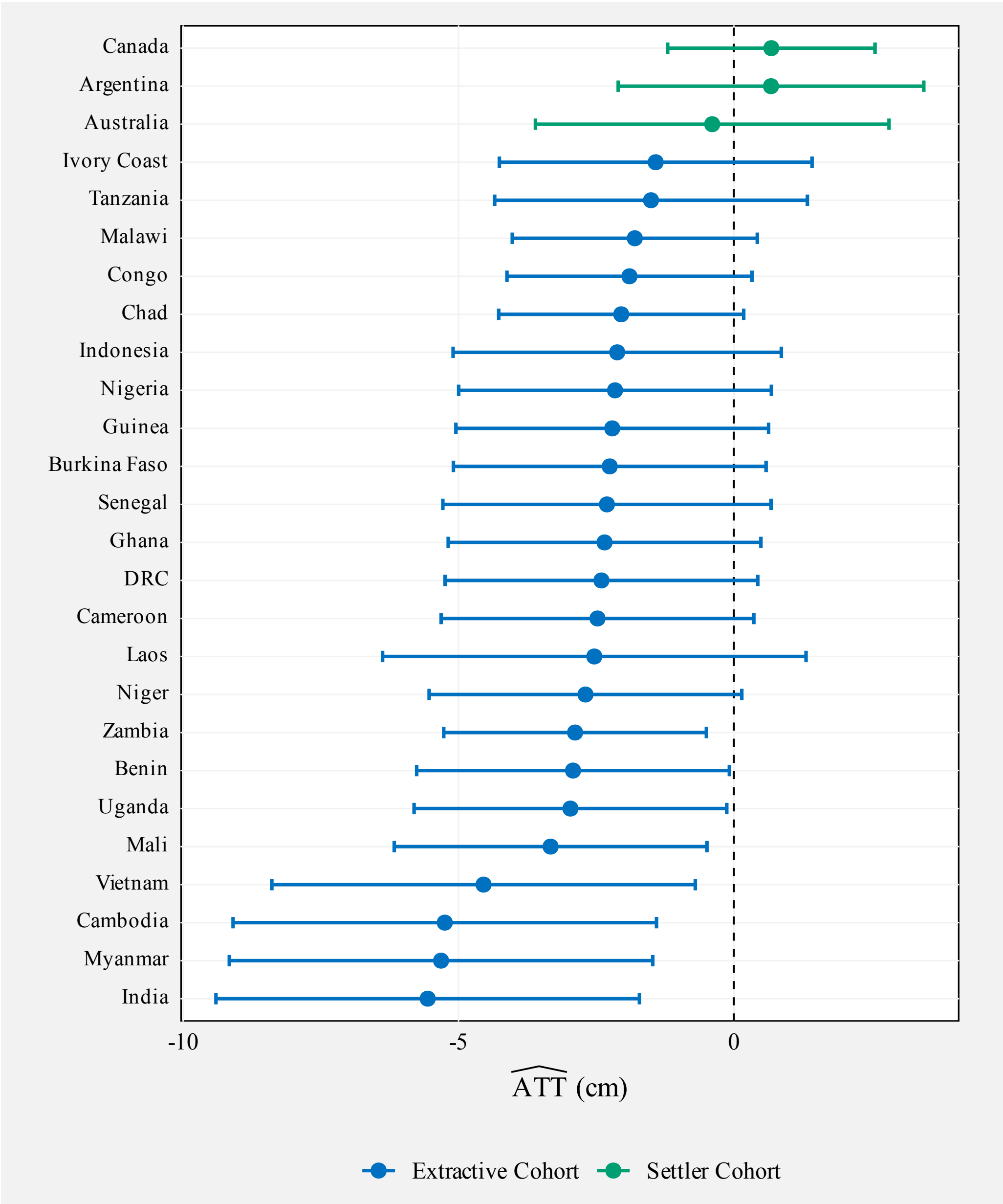


Figure B.3: Country-Specific $\widehat{ATT}$ Estimates: Max Date

Notes: This figure presents disaggregated, country-specific SDiD point estimates ($\widehat{ATT}$) under the alternative coding specification (maximum independence dates). The expanded sample remains identical to Figure B.2 ($N_{tr} = 23$ for the extractive cohort, blue markers; $N_{tr} = 3$ for the settler cohort, green markers). Error bars represent 95% confidence intervals. The vertical dashed line indicates a treatment effect of zero.

B.3 Leave-One-Out Sensitivity Check

While the unit-by-unit analysis already provides evidence that the results were not driven by outliers, it does not demonstrate how the exclusion of a specific country actually influences the pooled model. To explicitly test this, a systematic "leave-one-out" robustness check was conducted. This procedure iteratively excludes one country at a time from the sample and re-estimates the baseline model, allowing for a direct test of whether the aggregate findings are overly reliant on a specific outlier. Because the settler cohort comprises only two units, it was excluded from this exercise, as dropping one unit would merely revert the pooled estimate to an individual, unit-level effect.

The results for the overall extractive cohort are presented in Figure B.4. The estimates demonstrate stability: iteratively excluding each of the 14 extractive colonies yields pooled $\widehat{ATT}$s that remain virtually unchanged from the primary baseline estimate (-3.7 cm). Furthermore, all 95% confidence intervals remain strictly negative and well separated from the zero line.

This robust stability extends to the comparative sub-sample analyses. Figure B.5 visualizes the leave-one-out estimates for the African and Asian cohorts. The macro-regional dynamic established in the main results remains consistent: regardless of which specific nation is excluded from the estimation, the Asian sub-sample consistently exhibits a more severe anthropometric drop than the African sub-sample.

Finally, Figure B.6 confirms the resilience of the institutional transition mechanisms within the African cohort. The systematic exclusion of individual units preserves the core divergence outlined in Chapter 5. Unplanned transitions consistently produce a larger negative penalty than planned transitions. Ultimately, these analyses confirm that the directional findings and comparative magnitudes of the baseline pooled models are not statistical artifacts generated by single extreme cases.

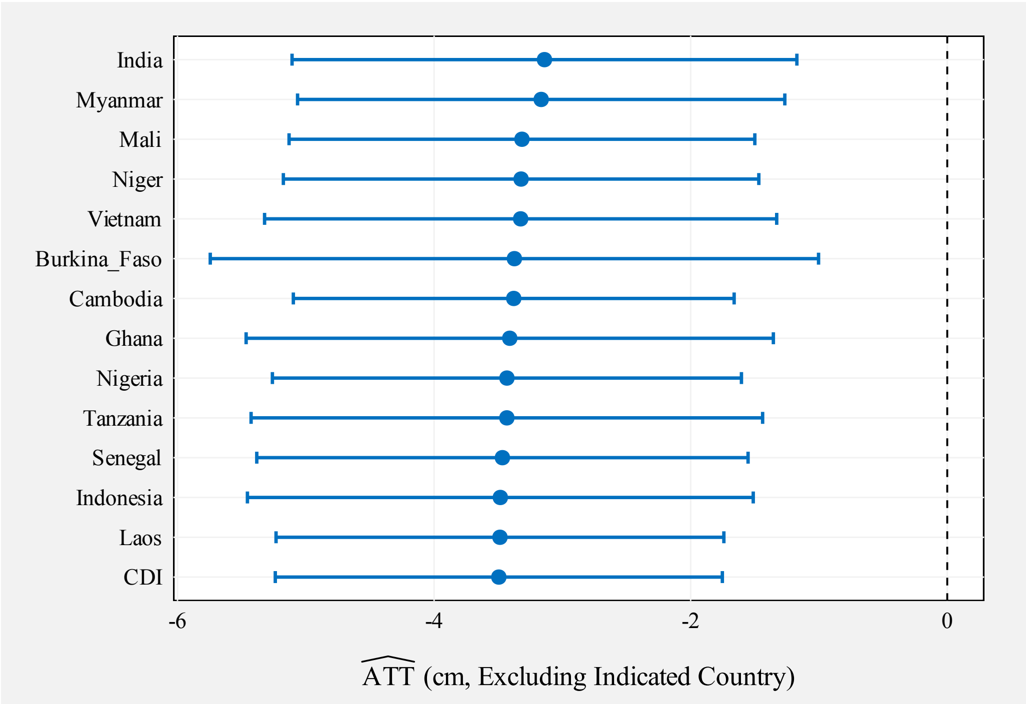


Figure B.4: Leave-One-Out Estimates: Extractive Cohort

Notes: This figure displays the leave-one-out point estimates ($\widehat{ATT}$) alongside their respective 95% confidence intervals for the extractive cohort. Each row represents the pooled $\widehat{ATT}$ when the indicated country is sequentially excluded from the estimation sample. The vertical dashed line represents a treatment effect of zero.

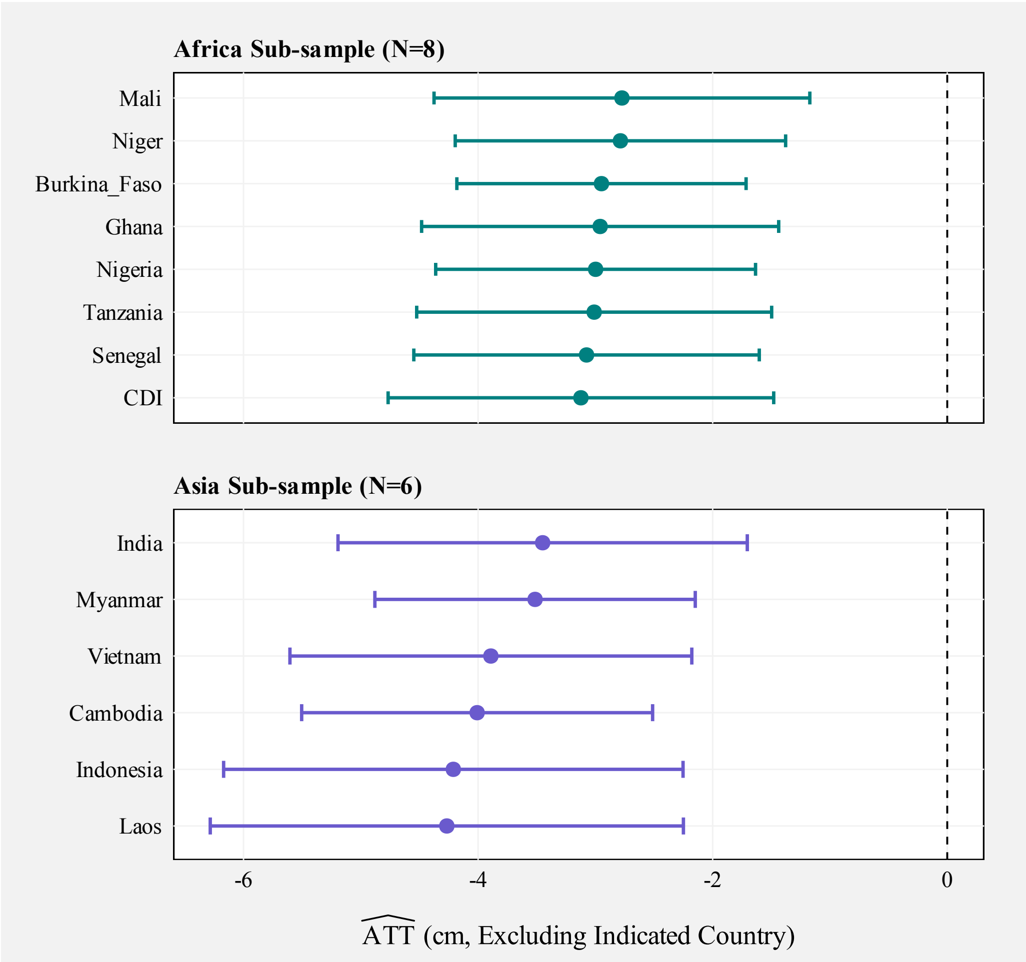


Figure B.5: Leave-One-Out Estimates: Africa vs. Asia

Notes: This figure displays the leave-one-out point estimates ($\widehat{ATT}$) disaggregated by regional sub-samples (Africa and Asia) within the extractive cohort. Each row represents the $\widehat{ATT}$ when the indicated country is excluded from the estimation sample. Error bars represent 95% confidence intervals. The vertical dashed line indicates a treatment effect of zero.

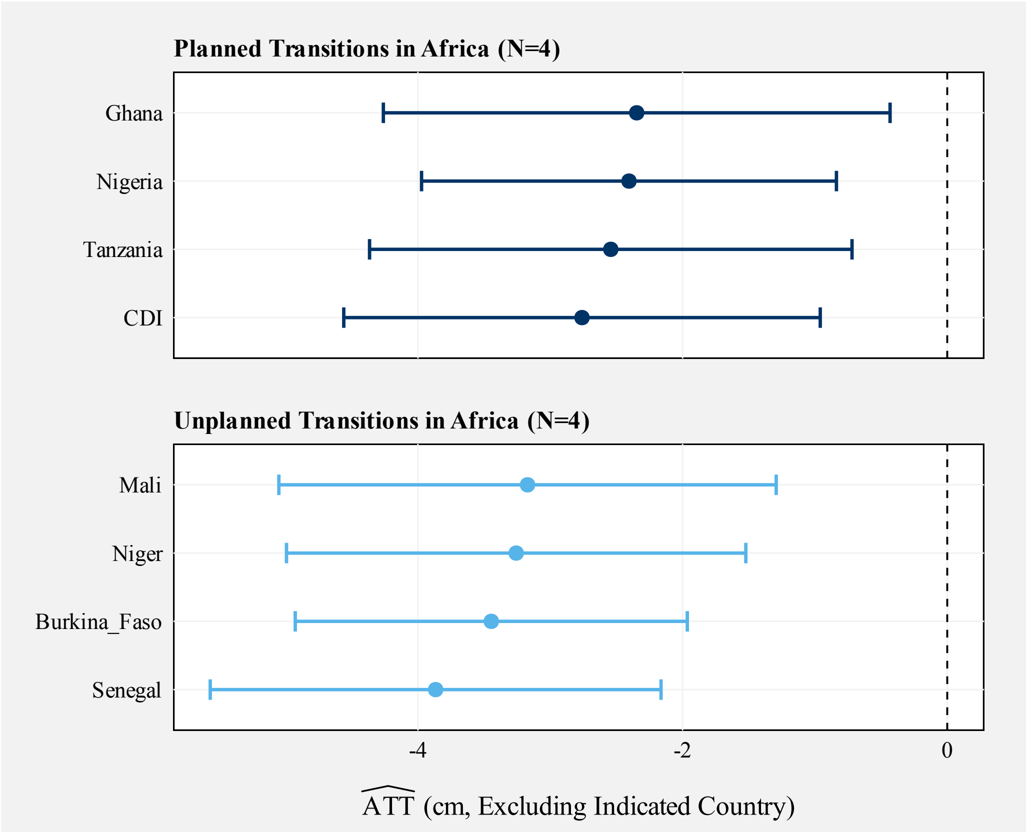


Figure B.6: Leave-One-Out Estimates: Planned vs. Unplanned Transitions

Notes: This figure displays the leave-one-out point estimates ($\widehat{ATT}$) split by the mechanism of independence (planned vs. unplanned transitions) for the African sub-sample. Each row represents the $\widehat{ATT}$ when the indicated country is excluded from the estimation sample. Error bars represent 95% confidence intervals. The vertical dashed line indicates a treatment effect of zero.

B.4 Sample Composition Sensitivity Check

As discussed in Section 4.4.1, constructing a strictly balanced panel for the extractive cohort required a trade-off between the temporal length of the pre-treatment observation window and the sample size of the treated cohort. Specifically, the baseline model prioritizes temporal depth by retaining the 1940 observation decade in order to provide the algorithm with a longer pre-treatment runway to construct a counterfactual.

Figure B.7 presents the event study plot for the alternative specification, which excludes the 1940 decade from the pre-treatment period. By relaxing this constraint, the extractive cohort expands from 14 to 19 treated units, allowing for the inclusion of Benin, Cameroon, the Democratic Republic of the Congo (DRC), Guinea, and Uganda. While only period $\ell = -4$ statistically violates the parallel trends assumption—with both the point estimate and its confidence interval resting entirely above the zero line—a broader, downward trajectory is visible across the entire pre-treatment window. This yields a worse pre-treatment fit than the baseline specification (see Section 5.1.2).

Figure B.8 demonstrates the impact of this alternative specification on the pooled estimates compared to the baseline model. Despite the degraded pre-treatment fit, the results demonstrate directional consistency across all cohort specifications. The re-estimated $\widehat{ATT}$s remain persistently negative and statistically significant, though the point estimates become slightly attenuated (less negative) in the expanded sample. This attenuation is likely a direct consequence of the downward pre-treatment trajectory, which suppresses the true size of the treatment effect.

Nevertheless, the macro-regional dynamics established in the baseline models remain intact: the Asian sub-sample continues to exhibit a more severe anthropometric penalty than the African sub-sample. Regarding the institutional transition mechanisms within Africa, the alternative specification produces a narrower gap between the two transition types. This compression indicates a degree of sensitivity to sample composition, which is an expected dynamic when disaggregating the data into small, distinct cohorts. However, despite this compositional sensitivity, the fundamental directional finding is not disrupted: unplanned transitions continue to yield a larger negative drop than planned transitions. Ultimately, the fact that the core results remain consistent even when tested against an alternative, sub-optimal sample composition further corroborates the resilience of the institutional interpretation.

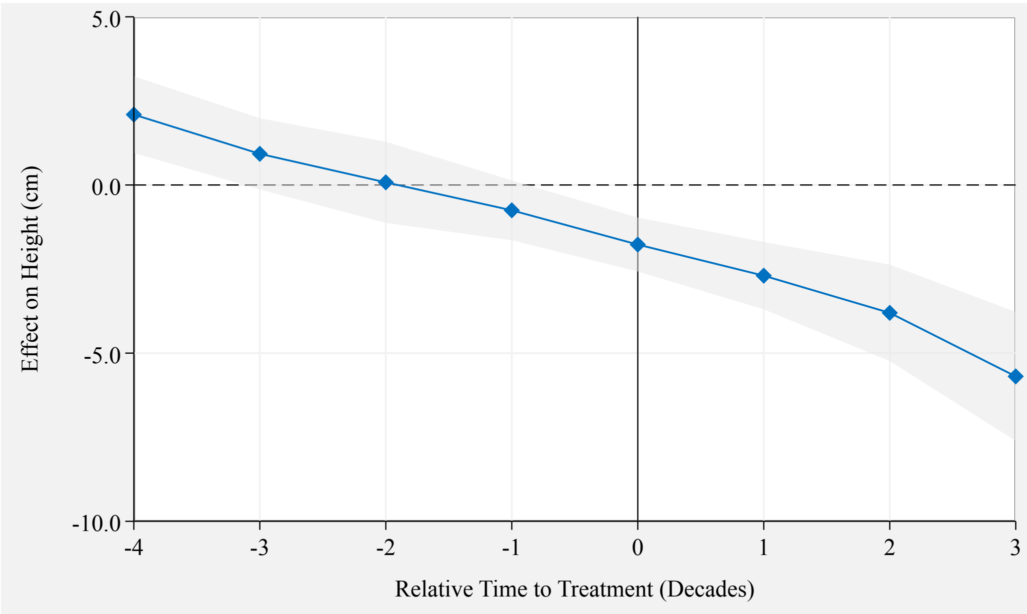


Figure B.7: SDiD Event-Study Estimates: 1940 Cohort Excluded

Notes: This figure presents the SDiD event-study estimates for the alternative specification (1940 excluded). The solid markers represent the point estimates ($\widehat{\tau}_{\ell}^{sdid}$) of the difference in mean terminal adult stature (in cm) between the treated cohort and its synthetic counterfactual, normalized relative to the temporal point of independence (T_0). The shaded regions represent the 95% confidence intervals. The vertical solid line indicates the exact decade of transition to independence ($\ell = 0$), and the horizontal dashed line marks zero effect.

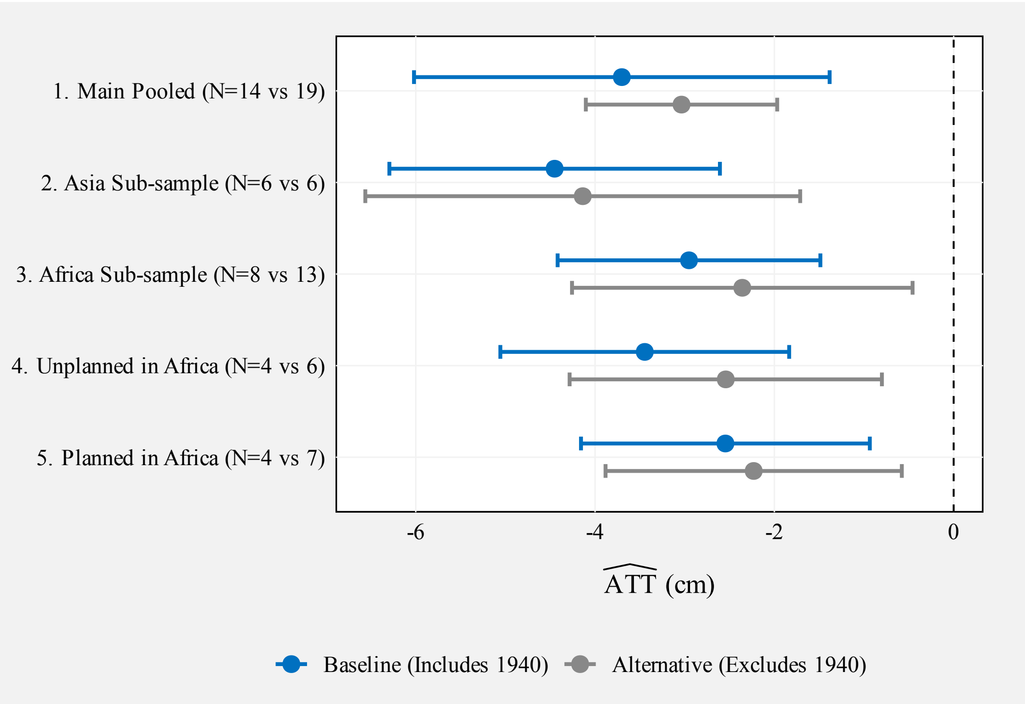


Figure B.8: Sensitivity to Sample Composition: 1940 Cohort Included vs. Excluded

Notes: This figure compares the $\widehat{ATT}$ across different sub-samples when including versus excluding the 1940 cohort. The blue dots denote the baseline specification, while the grey dots denote the alternative sample. Error bars represent 95% confidence intervals. The vertical dashed line indicates a treatment effect of zero.

B.5 In-Space Placebo Check

In a classical DiD setting, “in-time” placebo tests—which artificially assign a fake treatment date in the past or future—are conducted to provide support for the parallel trends assumption. However, this test carries a different interpretation within a SDiD framework. Because the SDiD algorithm mathematically optimizes unit and time weights specifically to construct a fitting counterfactual, an in-time placebo does not yield valuable information regarding underlying parallel trends. Instead, it merely assesses the treated-donor fit for that shifted time window. Consequently, in an SDiD setting, an in-time placebo can only be utilized to verify whether the estimator mechanically generates spurious effects at randomly chosen dates.

In the context of this study, executing a reliable in-time placebo for the main extractive sample proved mathematically infeasible. Shifting the fake treatment dates forward (e.g., after the primary 1950 and 1960 independence waves) would contaminate the test; because the available data terminates in 1980, this forward shift would overlap with the actual post-treatment period, capturing the persistent negative treatment effects observed in the main event study plots. Conversely, a backward shift was similarly unviable. Starting the treatment window prior to the date chosen for the main model (1910) results in a substantial drop in the available sample size. As a consequence, the estimator struggles to construct a precise counterfactual, leading to highly unstable pre-trends and rendering the test unreliable.

Because a conventional in-time placebo test could not provide meaningful results, this analysis relies on an “in-space” placebo. The settler cohort serves as this natural placebo since treatment is assigned to a group where theoretical literature predicts no negative institutional shock. The null effect for this cohort (see Section 5.1.2) successfully confirms that the SDiD estimator does not simply manufacture mechanical, negative structural breaks at assigned treatment dates.

Consequently, this validates that the observed anthropometric penalties reflect actual historical impacts rather than spurious algorithmic artifacts.

B.6 Error Structure and Inference

To perform inference, the SDiD framework provides three methods: the bootstrap, the jackknife, and placebo (permutation) tests (Arkhangelsky et al., 2021). However, the bootstrap and jackknife estimators rely on large-sample asymptotic assumptions to guarantee consistent inference. As previously discussed, these asymptotic properties do not hold in this analysis, as the largest model employs a treated sample size of only $N_{tr} = 14$.

Arkhangelsky et al. (2021) show that for small-N scenarios, the placebo method is required, as it can generate valid inference for sample sizes as small as $N_{tr} = 1$ since it does not rely on those large-sample assumptions. Instead, this permutation method operates "in-space" by utilizing the donor pool: while keeping the actual treatment date fixed, the algorithm repeatedly draws random combinations of control units to form fake treatment groups—equal in size to the actual treated cohort—and runs the SDiD estimator on them. This process generates a distribution of placebo effects representing the natural noise in the data. The actual coefficient of the treated group is then compared against this distribution. For example, at an $\alpha = 0.05$ significance level, an estimated treatment effect that falls outside the central 95% of the placebo distribution is deemed statistically significant.

Crucially, for this permutation approach to produce reliable results, two assumptions must hold. First, the donor pool must strictly outnumber the treated group (Arkhangelsky et al., 2021). This is a combinatorial necessity: without a donor pool that strictly exceeds the treated cohort, it is mathematically impossible for the algorithm to generate the multiple unique permutations of fake treatment groups required to build the distribution. As shown in Table A.2, this condition holds true across all specified models.

The second requirement is homoscedasticity across all units; the treated and donor pools must share a structurally similar variance profile (Arkhangelsky et al., 2021). Because the donor distribution is utilized as the baseline to test the treated group's coefficient, a severe divergence in variance between the two groups invalidates the donor pool as a reliable mechanism for inference. However, there is currently no classical statistical test available to formally verify homogeneity within an SDiD setting (Arkhangelsky & Samkov, 2025; Ferman & Pinto, 2019).

To overcome this, Ferman and Pinto (2019) propose an alternative approach: mathematically isolating the idiosyncratic error term of the dependent variable. To accurately measure this variance, the underlying structural differences between units and aggregate time shocks must first be partialled out. By regressing the pre-treatment data on unit and time fixed effects, this confounding heterogeneity is stripped away. The standard deviations of the resulting isolated residuals are then calculated for both the treated and donor pools to allow for direct visual comparison. Figure B.9 executes this exact diagnostic.

The resulting box plots demonstrate that in the main pooled extractive cohort, the medians of the pre-treatment residual standard deviations for both the treated and donor units are nearly identical. This homoscedasticity holds equally true for the African, Asian, and Planned subsamples. Consequently, the placebo inference for these models is credible.

However, heteroscedasticity presents a challenge for the Settler cohort and the Unplanned subsample. In the Settler cohort, the median volatility of the treated units is substantially smaller than that of the donor pool. Because the algorithm constructs the placebo distribution from this highly volatile donor pool, the resulting distribution is artificially wide. This introduces a conservative bias, significantly increasing the probability of a Type II error (failing to reject the null hypothesis) (Ferman & Pinto, 2019). Conversely, the Unplanned sub-sample exhibits the opposite violation: the median volatility of the treated pool is larger than that of the donor pool. Here, the placebo distribution is artificially narrow, creating an anti-conservative bias that increases the probability of a Type I error (finding a statistically significant effect where none exists) (Ferman & Pinto, 2019).

Usually, to address problematic inference, two methodological possibilities arise: transforming the dependent variable (e.g., applying a natural logarithm) or introducing post-estimation corrections. The first approach yielded no effect in the given setting; because the variance of human height is naturally bounded, the logarithmic transformation did not meaningfully alter the structural variance mismatch between the cohorts.

In the second case, classical TWFE models would typically employ clustered or robust standard errors to perform such variance corrections. Nonetheless, because these adjustments rely on large-sample asymptotics that are mathematically invalid when the number of treated units is small (Cameron & Miller, 2015; Conley & Taber, 2011), such post-estimation corrections are not possible here. Furthermore, while alternative resampling methods have been developed for small-sample standard DiD models, they have not currently been adapted for the SDiD framework, rendering them infeasible (Ferman & Pinto, 2019).

Consequently, it cannot be excluded that the statistical results of the Settler and the Unplanned cohorts are mechanical artifacts of biased placebo distributions. This underscores why the findings for both groups must be interpreted strictly as suggestive mechanism evidence rather than definitive proof.

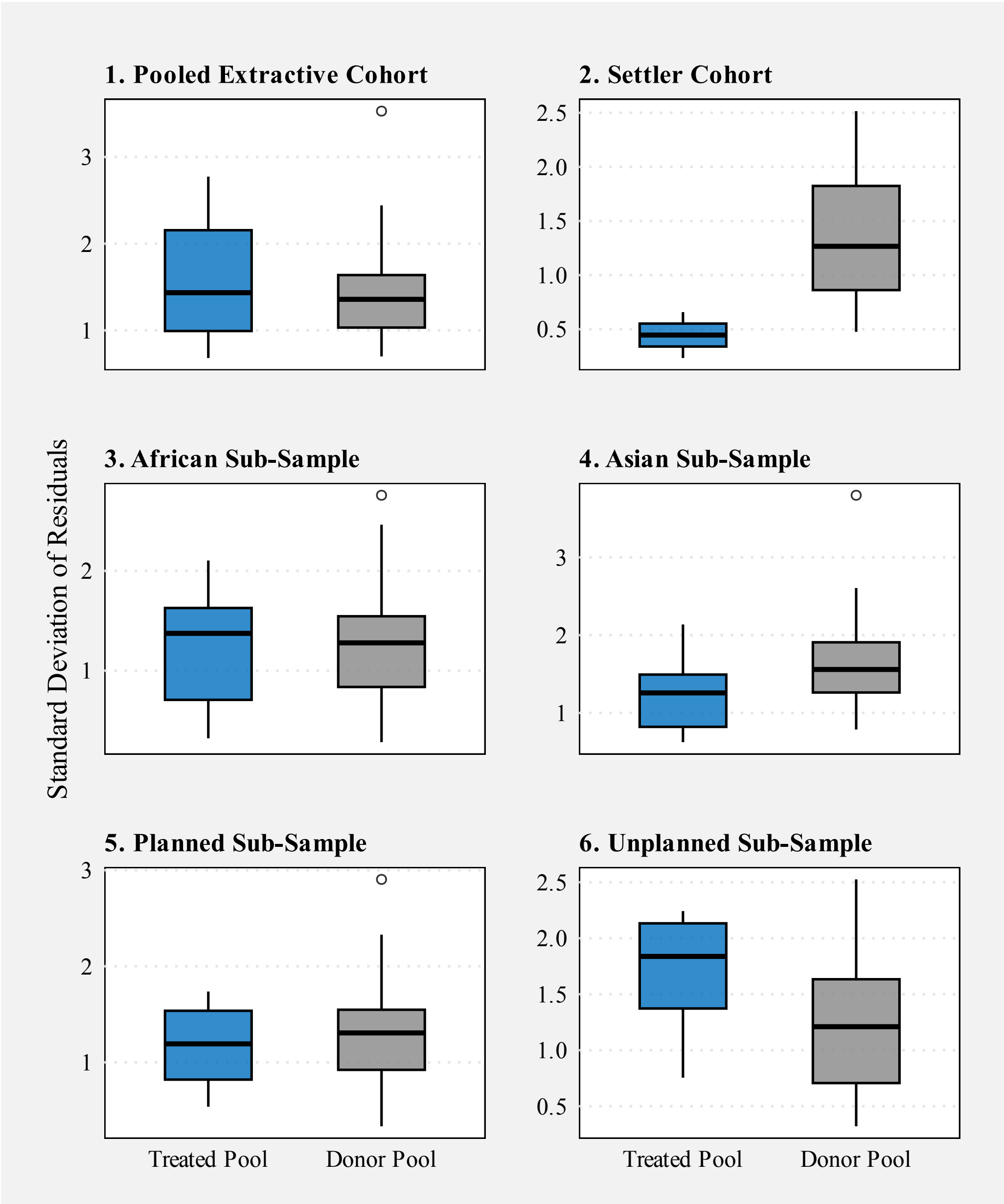


Figure B.9: Homoscedasticity Check of Residuals

Notes: This figure assesses the homoscedasticity assumption by comparing the standard deviation of residuals between the treated pool (blue) and donor pool (grey) across all samples and sub-samples. The boxplots display the median and interquartile ranges, while circles represent outlier observations. Equal spread across groups indicates stable variance within the models.

C Identification Assumptions Checks

As established in Section 4.2.2, for the ATT in an SDiD setting to represent a truly unbiased causal effect, the standard identification assumptions of the classical DiD approach must be fulfilled (Arkhangelsky & Samkov, 2025): the no-anticipation assumption, the Stable Unit Treatment Value Assumption (SUTVA), and the absence of time-varying confounders at the time of treatment. The subsequent sections systematically evaluate whether these assumptions hold within the context of this thesis.

C.1 Treatment Anticipation

The no-anticipation assumption requires that the treatment exerts no effect on the outcomes prior to its official implementation ($\ell = 0$). While this paper conceptualizes the ultimate transfer of power as a shock-like institutional disruption, the eventuality of regime change was not completely unforeseeable (Cogneau et al., 2018). Because colonial actors were aware of mounting anti-colonial pressures, they had the opportunity to alter their behavior prior to the actual event of independence. Therefore, the potential for anticipatory effects must be evaluated.

Cogneau et al. (2018) demonstrate one potential mechanism for an anticipation effect. According to their findings, real public expenditure per capita quadrupled in several African colonies between 1940 and 1955. They attribute this increase to "appeasement" policies, noting that colonizers pumped subsidies into the territories in an attempt to quell decolonization movements. Theoretically, this pre-treatment spending surge could introduce amplification bias if it artificially inflated the pre-independence biological baseline. However, as established in Chapter 3, terminal adult height is a highly inelastic indicator. Because shifting cohort trajectories requires profound and sustained improvements in the nutritional and epidemiological environment, it is highly questionable whether such a transient increase in public spending exerted a strong enough biological effect. This is corroborated by Figure 5.1, which confirms a stagnation of aggregated mean height during this exact pre-treatment period, indicating that late-colonial appeasement spending failed to translate into structural biological gains. Thus, amplification bias is improbable.

Conversely, a second theoretical mechanism involves the opposite effect: in some cases, as independence became inevitable, colonial administrators and foreign enterprises engaged in anticipatory capital flight and institutional withdrawal in the immediate period preceding the legal transfer of power (Ogle, 2020). Nevertheless, even if such anticipatory behavior did occur, the inherently inelastic nature of human stature again provides a natural defense against this empirical threat. First, a gradual, preemptive withdrawal of capital and administrative staff would not immediately generate a structural rupture severe enough to stunt physical growth. Second, even if the peak of institutional rupture occurred slightly before the exact, historically defined date of independence, the fact that terminal stature captures the cumulative biological environment of the entire childhood and adolescence ensures that a short-term anticipatory shock is absorbed into a much longer developmental window. This provides robustness against minor temporal measurement errors. Thus, a scenario in which the pre-treatment anthropometric trajectory was systematically depressed (attenuation bias) prior to $\ell = 0$ also remains improbable.

Consequently, severe violations of the no-anticipation assumption can be excluded.

C.2 Confounding at Time of Treatment

While the SDiD estimator successfully accounts for unobserved, time-varying confounders through its latent factor structure, it cannot neutralize confounders that coincidentally align with the exact year of independence ($\ell = 0$). This limitation arises because the algorithm's weighting procedure is strictly anchored to the pre-treatment trajectory (Arkhangelsky et al., 2021). If the estimator were allowed to optimize weights using post-treatment observations, it would inappropriately absorb the actual treatment effect while attempting to construct a valid counterfactual. Consequently, any exogenous shock occurring simultaneously with the institutional transition becomes mathematically indistinguishable from the treatment. However, the threat posed by such concurrent confounders can be dismissed due to two structural defenses inherent in the research design.

First, at the measurement level, the inelasticity of stature once again provides protection against this identification problem. A short-term environmental or macroeconomic shock—such as a severe drought or a global commodity price spike—occurring at the moment of independence would leave a negligible effect on average terminal stature. Only a persistent, structural deterioration in childhood living conditions—such as the collapse of public health infrastructure produced by sustained institutional decay—could generate the magnitude and growing temporal persistence of the anthropometric declines documented in this analysis.

Second, even if a historical shock satisfied the conditions laid out above, the sequential stratification of the empirical design acts as a further safeguard. For a concurrent shock to fully explain these results, it would need to simultaneously affect geographically dispersed treated units across two different continents, perfectly align with their staggered moments of independence, and produce directionally consistent anthropometric penalties in almost every unit-by-unit analysis (see Appendix B.2)—all while leaving the settler-colony cohort unaffected.

The only phenomenon theoretically capable of fulfilling all of these conditions would be the outbreak of severe armed conflict. As established in Chapter 2, many formerly colonized nations became embroiled in conflicts around the time of their independence, and armed conflict induces shocks severe enough to permanently stunt anthropometric outcomes. However, as also discussed in the theoretical framework, such violence was frequently a direct consequence of the struggle for independence and the subsequent institutional vacuum. Therefore, these conflicts must be understood as a mechanism of the institutional rupture itself, rather than an exogenous confounder.

Consequently, the combination of a structurally resistant biological outcome measure and a directionally consistent pattern across multiple empirical stratifications constitutes a robust defense against concurrent shock explanations.

C.3 Cross-Border Spillovers

Finally, SUTVA mandates that the treatment assignment of one unit does not affect the potential outcomes of other units in the sample. In the context of this thesis, this means there must be no cross-border spillover effects. The first potential source of such a spillover arises between colonized nations, which could have manifested in two distinct forms. First, armed conflicts or civil unrest arising during one nation's struggle for sovereignty could have spilled over into neighboring territories. Second, an early-decolonizing country could have acted as a political catalyst, inspiring anti-colonial movements in neighboring states and thereby generating early institutional friction. Strang (1990) demonstrates that such regional contagion did indeed occur.

However, these intra-cohort spillovers do not invalidate the empirical findings. First, as argued regarding the no-anticipation assumption, any instability would need to be severe enough to affect terminal adult stature. Second, even if regional instability from an early-decolonizing nation did negatively impact the anthropometric trajectory of a neighboring, soon-to-be-independent colony, this spillover effectively functions as an anticipation effect. By prematurely depressing the neighbor's pre-treatment baseline, this dynamic introduces attenuation bias, ensuring that the estimated anthropometric penalty represents a conservative lower bound of the true effect.

The second potential source of spillover exists between the treated colonies and the never-colonized control group. For the vast majority of the donor pool—comprising European nations—a violation here is highly implausible. As established previously, human stature is primarily determined by domestic nutritional intake and the local disease environment. No plausible channel exists through which a foreign political transition in Africa or Asia could persistently degrade these domestic conditions sufficiently to stunt terminal adult height in a distant, uncolonized control unit.

Environmental factors or long-term trends could drive this pre-independence shift, making the post-treatment interpretation as an upper bound essential. Nevertheless, one must account for never-colonized states geographically adjacent to the treated cohort, such as Thailand or Ethiopia. In these specific cases, it is theoretically possible that regional instability spilling over from neighboring independence struggles induced localized economic or epidemiological distress. Yet, even if such cross-border spillovers did negatively impact the anthropometric trajectories of these specific control units, it would not invalidate the findings. First, the SDiD algorithm constructs the synthetic counterfactual as a weighted average, meaning any localized shock in a single border nation is diluted by the unaffected remainder of the donor pool.

Second, and more importantly, if a control unit's stature were artificially depressed by a neighbor's decolonization, the post-treatment height of the synthetic control group would be biased downward. This would shrink the estimated gap between the treated and control cohorts, once again introducing attenuation bias.

To conclude, while the complex reality of decolonization means that spillovers cannot be entirely ruled out, their presence in either the treated or control group would only bias the results toward zero. Therefore, even if such contagion occurred, the estimated $\widehat{ATT}$ remains a conservative lower bound of the actual treatment effect.

D Econometric Specifications

D.1 Formalization Weight Optimization

As outlined in Section 4.2.1, this section formally defines how the unit ($\hat{\omega}^{sdid}$) and time weights ($\hat{\lambda}^{sdid}$) are selected.

D.1.1 Unit Weights

Arkhangelsky et al. (2021) and Clarke et al. (2023) show that the unit weights are determined by solving the following optimization problem:

$$(\hat{\omega}_0, \hat{\omega}^{sdid}) = \arg \min_{\omega_0 \in \mathbb{R}, \omega \in \Omega} \underbrace{\sum_{t=1}^{T_{pre}} \left(\omega_0 + \sum_{i=1}^{N_{co}} \omega_i Y_{it} - \frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} Y_{it} \right)^2}_{\text{Unit Weight Loss Function}} + \underbrace{\zeta^2 T_{pre} \|\omega\|_2^2}_{\text{Ridge Penalty}} \quad \text{(D.1)}$$

where:

$$\Omega = \left\{ \omega \in \mathbb{R}_+^N, \text{ with } \sum_{i=1}^{N_{co}} \omega_i = 1 \text{ and } \omega_i = \frac{1}{N_{tr}} \text{ for all } i = N_{co}+1, \dots, N \right\} \quad \text{(D.2)}$$

To formalize this optimization, the sample is partitioned into distinct groups and time horizons. The variable N_{co} represents the total number of untreated control units (never-colonized or not-yet-independent nations), while N_{tr} denotes the total number of treated units (colonized nations). The parameter T_{pre} represents the number of pre-treatment (pre-independence) periods, and Y_{it} denotes the observed outcome (decadal mean male height) for a specific unit i (country) in period t (decade).

Using these variables, the arg min operator seeks to minimize the loss function to find the optimal set of unit weights ($\hat{\omega}^{sdid}$) and the optimal unpenalized intercept ($\hat{\omega}_0$). It does so by minimizing the distance between the unweighted average of the treated units ($\frac{1}{N_{tr}} \sum_{i=N_{co}+1}^{N} Y_{it}$) and the conditionally weighted average of the control units ($\sum_{i=1}^{N_{co}} \omega_i Y_{it}$) across all pre-treatment periods. During this optimization process, $\hat{\omega}_0$ dynamically absorbs the absolute baseline differences between the treated cohort and the donor pool. This feature ensures that the loss function evaluates deviations from parallel trajectories rather than mismatches in absolute levels, thereby allowing for a constant additive shift between the treated nations and their synthetic counterfactual. By allowing this additional shift, $\hat{\omega}_0$ serves as the component that accounts for unit fixed effects during the optimization procedure—which, as discussed in Chapter 4.2, distinguishes SDiD from traditional Synthetic Control (SC) methods.

The outlined minimization process is strictly bound by the constraint set Ω, which defines three constraints on the unit weights. First, $\omega \in \mathbb{R}_+^N$ enforces non-negativity, ensuring the estimator does not extrapolate outside the support of the data by assigning negative weights to donor countries. Second, $\sum_{i=1}^{N_{co}} \omega_i = 1$ mandates that the weights assigned to the donor pool sum to exactly one. Combined, these two restrictions ensure that the resulting synthetic counterfactual is a convex combination of observed control units; i.e., the generated counterfactual cannot mathematically exceed the maximum or fall below the minimum bounds of human stature observed in the donor pool. Third, ($\omega_i = \frac{1}{N_{tr}}$) fixes the weights of the treated cohort as a uniform average. This restricts the optimization algorithm to exclusively calibrate the control units, requiring it to construct a synthetic counterfactual that mimics the simple arithmetic mean trajectory of the treated group. Because the optimization target is anchored to this unweighted treated baseline, the resulting causal parameter ($\hat{\tau}^{sdid}$) identifies the ATT, rather than the ATE.

Ultimately, the expression $\zeta^2 T_{pre} \|\omega\|_2^2$ represents the unified regularization term. Within this block, $\|\omega\|_2^2$ denotes the squared Euclidean norm, which operates as a ridge penalty. By penalizing highly concentrated weight distributions, this mechanism systematically forces the optimization algorithm to build its synthetic counterfactual from a broader combination of donor countries rather than overfitting to a single, noisy control unit.

Multiplying the penalty by the length of the pre-treatment horizon (T_{pre}) serves as a normalization anchor. Because the loss function naturally grows with the number of observed periods, the inclusion of T_{pre} ensures that the penalty scales in tandem with the baseline timeline, preserving temporal invariance.

Finally, the overall intensity of this penalty is governed by ζ^2. Formally, ζ is defined as $\left(N_{tr}T_{post}\right)^{1/4}\hat{\sigma}$, where:

$$\hat{\sigma}^2 = \frac{1}{N_{co}(T_{pre}-1)}\sum_{i=1}^{N_{co}}\sum_{t=1}^{T_{pre}-1}(\Delta_{it}-\bar{\Delta})^2 \tag{D.3}$$

and $\Delta_{it} = Y_{i,(t+1)} - Y_{it}$. Here, Δ_{it} calculates the first difference of the outcome variable for each donor country. This is done to eliminate time-invariant country fixed effects and isolate the within-unit variation. Subsequently, subtracting the global mean ($\bar{\Delta}$) of these first differences, removes common time effects, ensuring that the variance calculation is based on the residual idiosyncratic variation within the donor pool. Finally, to obtain the average variance ($\hat{\sigma}^2$), these squared differences are divided by the total number of pre-treatment control observations ($N_{co}(T_{pre}-1)$). Thus, ζ represents the standard deviation of the control units ($\hat{\sigma}$) after purging all unit and time fixed effects, scaled by the total number of treated observation cells ($N_{tr}T_{post}$) raised to the power of $\frac{1}{4}$. Consequently, ζ^2 makes the degree of regularization adaptive to the characteristics of the donor pool. Greater volatility among the control units increases the penalty and discourages highly concentrated weight distributions, whereas a more stable donor pool allows the algorithm to tolerate more concentrated weights without substantially increasing the risk of overfitting. Together, these optimization procedures ensure that the SDiD estimator constructs a synthetic counterfactual that closely reproduces the pre-independence trajectory of the treated colonies while limiting the risk of overfitting to idiosyncratic fluctuations in individual donor countries.

D.1.2 Time Weights

Symmetrically to the unit weights, the time weights ($\hat{\lambda}^{sdid}$) are determined by solving the following optimization problem (Arkhangelsky et al., 2021; Clarke et al., 2023):

$$(\hat{\lambda}_0, \hat{\lambda}^{sdid}) = \arg \min_{\lambda_0 \in \mathbb{R}, \lambda \in \Lambda} \underbrace{\sum_{i=1}^{N_{co}} \left(\lambda_0 + \sum_{t=1}^{T_{pre}} \lambda_t Y_{it} - \frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it} \right)^2}_{\text{Time Weight Loss Function}} + \underbrace{\zeta^2 N_{co} \|\lambda\|_2^2}_{\text{Ridge Penalty}} \quad \text{(D.4)}$$

where:

$$\Lambda = \left\{ \lambda \in \mathbb{R}_+^T, \text{ with } \sum_{t=1}^{T_{pre}} \lambda_t = 1 \text{ and } \lambda_t = \frac{1}{T_{post}} \text{ for all } t = T_{pre}+1, \ldots, T \right\}. \quad \text{(D.5)}$$

The notation follows that introduced in the previous subsection, with the additional parameter T_{post}, which denotes the number of post-treatment periods.

In this case, the loss function minimizes the distance between the unweighted post-treatment average ($\frac{1}{T_{post}} \sum_{t=T_{pre}+1}^{T} Y_{it}$) and the conditionally weighted pre-treatment average ($\sum_{t=1}^{T_{pre}} \lambda_t Y_{it}$) across all control countries. During this optimization, the unpenalized intercept $\hat{\lambda}_0$ absorbs common shifts affecting all of these units over time. Consequently, the loss function focuses on differences in temporal trajectories rather than absolute level differences between periods. In this sense, $\hat{\lambda}_0$ represents the mechanism through which SDiD incorporates time fixed effects during the optimization stage, constituting the second feature that distinguishes the estimator from traditional SC methods.

This minimization problem is subject to the constraint set Λ, which imposes restrictions analogous to those applied to the unit weights. First, the non-negativity ($\lambda \in \mathbb{R}_+^T$) and sum-to-one

($\sum \lambda_t = 1$) constraints ensure that the counterfactual baseline is constructed as a convex combination of historical pre-independence periods. This eliminates temporal extrapolation, limiting the risk of linking the post-treatment counterfactual to an artificial baseline value that lies outside the actual boundaries of the data. Second, fixing the post-treatment weights uniformly to $\lambda_t = \frac{1}{T_{post}}$ ensures that the algorithm exclusively calibrates the pre-treatment decades to target the unweighted average trajectory of the post-treatment era.

Finally, the regularization term $\zeta^2 N_{co} \|\lambda\|_2^2$ serves a parallel purpose to that of the units weights. The squared Euclidean norm $\|\lambda\|_2^2$ introduces a ridge penalty that encourages temporal weight dispersion. Multiplying this penalty by the number of control units (N_{co}) acts as a normalization device, preserving invariance to the size of the donor pool.

However, in contrast to the optimization of the unit weights, where ζ^2 actively scales up to penalize data volatility in order to prevent excessive reliance on individual donor countries, the temporal regularization parameter is intentionally fixed at a small value, $\zeta = 1 \times 10^{-6} \times \hat{\sigma}$, where $\hat{\sigma}$ is defined as in the previous subsection. This reflects the fact that, in time-series settings, observations immediately preceding the treatment are often times the strongest predictors of post-treatment outcomes due to their temporal proximity. Consequently, concentrated weights on the final pre-treatment periods represent an expected feature rather than evidence of overfitting. Strong regularization is therefore unnecessary. Instead, the small penalty serves primarily a computational purpose by ensuring a unique solution whenever multiple combinations of historical periods generate identical pre-treatment trajectories.

Ultimately, the time-weight optimization ensures that the post-treatment divergence is evaluated relative to the pre-treatment periods most predictive of the subsequent trajectory.

D.2 Formalization Latent Factor Model

As introduced in Section 4.3.1, the inclusion of time-varying covariates is generally redundant within the SDiD framework. Arkhangelsky et al. (2021) demonstrate that this result follows from the underlying Data-Generating Process (DGP) specified by the estimator.

In this setting, the true trajectory of the data is assumed to be driven by a latent factor model (also known as an interactive fixed-effects model).[1] At the individual observation level, this DGP is written as:

$$Y_{it} = \gamma_i \upsilon_t^\top + \tau W_{it} + \varepsilon_{it} \tag{D.6}$$

Y_{it} represents the observed outcome (mean height) for unit i (country) at time t (decade). The causal parameter of interest is τ, W_{it} is the binary treatment indicator, and ε_{it} is the idiosyncratic error term. The critical component of this equation is the inner product $\gamma_i \upsilon_t^\top$. Here, γ_i is an unobserved $1 \times R$ row vector containing unit-specific latent factors (e.g., a nation's underlying health infrastructure), and υ_t is a $1 \times R$ row vector containing time-specific latent factors (e.g., global epidemiological shocks), which is transposed into an $R \times 1$ column vector to yield a scalar outcome. Because these vectors multiply together, the model explicitly allows for an interactive relationship where a global time shock (υ_t) impacts each country at a varying rate depending on its local vulnerabilities (γ_i).

This contrasts with classical TWFE models, which assume a strictly additive structure ($Y_{it} = \alpha_i + \beta_t + \tau W_{it} + \varepsilon_{it}$). Additive models force global shocks to impact every unit identically. Therefore, in such settings, it is often necessary to explicitly include time-varying covariates to capture potential diverging effects. Since the interactive structure of SDiD is designed to

[1]For a detailed discussion of interactive fixed-effects models, see, e.g., Xu (2017) and Bai (2009).

accommodate such unit-varying dynamic trajectories, the inclusion of additional time-varying covariates is generally unnecessary.

While Equation D.6 explains the intuition behind the redundancy of covariates, it leaves open the question of how the estimator accounts for these interactive effects. In order to recover τ, the estimator must neutralize $\gamma_i v_t^\top$. However, an identification problem arises. In a TWFE framework, identification is mathematically straightforward. Because TWFE imposes an additive structure, it is only necessary to estimate one scalar parameter for each of the N units and one for each of the T time periods. In an $N \times T$ panel dataset, the $N \times T$ available observations vastly exceed the $N + T + 1$ unknown parameters (comprising the unit fixed effects, time fixed effects, and the single treatment parameter τ), making the system overidentified and easily solvable.

Nevertheless, since the SDiD framework abandons this additive assumption, a degrees-of-freedom problem is introduced. Because the model requires estimating an unobserved $1 \times R$ vector (γ_i) for all N units, alongside an unobserved $1 \times R$ vector (v_t) for all T time periods, the total number of unknown parameters expands to $(N \times R) + (T \times R) + 1$. If the dimension R is left unrestricted, the number of unknowns will rapidly exceed the $N \times T$ available data points. Consequently, the unconstrained interactive model is underidentified. In such a scenario, any estimation attempt would simply overfit the data, mistaking random historical anomalies for systematic trends. Because this dimensionality problem affects the entire dataset simultaneously, the mathematical constraint that resolves it cannot be applied at the individual unit level. Instead, it is necessary to stack all observations into a matrix equation:

$$Y = L + \tau W + E \quad \text{where} \quad L = \Gamma \Upsilon^\top \tag{D.7}$$

In this aggregated notation, the observed outcomes (Y), treatment assignments (W), and error terms (E) expand into $N \times T$ matrices. More importantly, the individual interactive components

($\gamma_i v_t^\top$) combine to form the systematic baseline matrix L. This matrix is defined as the product of the $N \times R$ matrix of unit factors (Γ) and the $R \times T$ matrix of global time factors ($\Upsilon^\top$). To solve the underidentification problem highlighted above, the following assumption is applied: $R \ll \min(N, T)$. This constraint asserts that the true dimension of unobserved factors (R) is strictly smaller than the number of units (N) or time periods (T), reflecting the reality of a highly interconnected world: the vast majority of cross-country variation is driven by a concentrated set of shared latent factors rather than a multitude of independent, idiosyncratic shocks.

Because $R \ll \min(N, T)$, the baseline matrix L possesses a low-rank structure. This mathematically limits the degrees of freedom and resolves the dimensionality problem of a full-rank matrix. The key insight of the SDiD algorithm lies in how it navigates this constrained system. It never explicitly estimates the underlying parameters (i.e., it never calculates the exact numerical values of Γ or $\Upsilon^\top$). Instead, it accounts for them implicitly through its double-weighting procedure. Because the unit weights ($\hat{\omega}^{sdid}$) and time weights ($\hat{\lambda}^{sdid}$) are calibrated to closely align the observed historical trajectories of the treated and synthetic control groups, any weight combination that achieves this alignment neutralizes the unobserved latent factors embedded within the L matrix. Since this matching procedure relies on the geometric alignment of complete trajectories across units and periods, it also provides the intuition for the balanced-panel requirement of SDiD. Missing observations break this alignment and prevent the estimator from constructing a coherent synthetic counterfactual. Consequently, the estimator controls for interactive confounders purely through structural re-weighting. In doing so, the SDiD framework transforms the challenge of dynamic confounding from a data-collection problem into a geometric one, identifying the treatment effect through the structure of the baseline outcomes rather than through the explicit modeling of every potential confounder.