國立臺灣大學生物資源暨農學院農業經濟學系

碩士論文

Department of Agricultural Economics
College of Bio-resources and Agriculture
National Taiwan University
Master's Thesis

瓜地馬拉原住民與非原住民之間教育落差的演變： Oaxaca-Blinder 分解模型比較分析

Evolution of the Educational Gap Between Indigenous and Non-Indigenous Populations in Guatemala: A Comparative Analysis Using the Oaxaca-Blinder Decomposition

李楷凡

Kevin Gamaliel Higueros Cruz

指導教授: 張宏浩 博士

Advisor: Hung-Hao Chang, Ph. D.

中華民國 114 年 7 月
July, 2025

doi:10.6342/NTU202504017

國立臺灣大學碩士學位論文

口試委員會審定書

MASTER'S THESIS ACCEPTANCE CERTIFICATE
NATIONAL TAIWAN UNIVERSITY

瓜地馬拉原住民與非原住民之間教育落差的演變：Oaxaca-Blinder 分解模型比較分析

Evolution of the Educational Gap Between Indigenous and Non-Indigenous Populations in Guatemala: A Comparative Analysis Using the Oaxaca-Blinder Decomposition

本論文係李楷凡（R12627035 在國立臺灣大學生農學院農業經濟學系暨研究所完成之碩士學位論文，於民國 2025 年 7 月 24 日承下列考試委員審查通過及口試及格，特此證明。

The undersigned, appointed by the Department of Agricultural Economic, College of Bioresources and Agriculture on July 24th, 2025 have examined a Master's Thesis entitled above presented by Kevin Gamaliel Higueros Cruz (student ID R12627035) candidate and hereby certify that it is worthy of acceptance.

口試委員 Oral examination committee:

(指導教授 Advisor)

doi:10.6342/NTU202504017

摘要

本研究深入探討了瓜地馬拉原住民與非原住民之間教育差距在過去二十年（2006-2023 年）間的演變。研究採用 Oaxaca-Blinder 分解模型，將學歷年限的差異分解為由可觀察特徵（稟賦）解釋的部分和無法解釋的部分。本研究使用了兩種模型：基準模型（考慮了年齡、性別和地理位置等基本人口統計與社會經濟因素）和擴展模型（在上述因素之外納入了父母教育程度）。

分析揭示，在兩個時期和兩種模式下，非原住民人口的學歷年限均較高，顯示出顯著但持續存在的差距。一個關鍵發現是，當擴展模型納入父母教育程度時，差距中可解釋部分的比例顯著增加，這突顯了人力資本代際傳遞在教育不平等持續存在中的深遠作用。例如，2006 年，總體差距中可解釋部分從基準模型的 44.2% 顯著擴大至擴展模型的 81.6%。儘管總體差距隨時間略有縮小，但無法解釋的部分仍然顯著，尤其是在 2023 年（基準模型為 69.0%，擴展模型為 33.2%）。這表明，諸如歧視、教育系統的文化相關性或教育的差異回報等不可觀察因素，在造成這些差異方面日益關鍵。本研究進一步按地理位置（農村/城市）和性別對結果進行了細分，持續證明父母教育在所有亞群體中均具有更高的解釋力，同時也揭示了持續存在的無法解釋部分的獨特構成模式。最終，這些結果強調了採取綜合政策干預措施的必要性，這些措施不僅要擴大教育機會，還要解決結構性不平等並推動文化相關的教育，以真正消除瓜地馬拉的教育差距。

關鍵詞：教育差距，原住民，非原住民，瓜地馬拉，父母教育，不平等

doi:10.6342/NTU202504017

Abstract:

This study conducts a deep examination of the evolution of the educational gaps between indigenous and non-indigenous population in Guatemala over two decades (2006-2023). Using the Oaxaca-Blinder decomposition model, the research breaks down the disparities in years of schooling into explained and unexplained components, by observable characteristics (endowments) and an unexplained part. Two models were used, the baseline model, which considered demographic and socioeconomic fundamental factors (age, sex, and geographic location), and an extended model, that incorporated additional factors to the variables.

The analysis reveals an important but persistent gap in both periods and modes, where the non-indigenous population achieve more years of schooling. A critical finding in the substantial increase in the proportion of the gap, when the parental education is included in the extended model, this highlights the profound role of the intergenerational transmission of the human capital in the perpetration of the education inequality. An example, in 2006, the explained component in the overall gap significantly expanded from 44.2% in the Baseline to 81.6% in the Extended model. While the total gap shows a slight reduction over time, the unexplained component remains significant, particularly in 2023 (69.0% in the Baseline, 33.2% in the Extended). This suggest that unobservable factors, such as discrimination, the cultural relevance of the educational system or the differential returns of education, are everyday more crucial in the generation of these disparities. The research further disaggregates finding by geographic location (rural/urban) and by gender, consistently demonstrating the increased explanatory power of parental education across all subgroups, while also revealing distinctive patterns in the composition of the persistence unexplained portion. Ultimately, these results emphasize

doi:10.6342/NTU202504017

the need of integral politics intervention, that not only expand the access to education, but that also address structural inequalities and promote culturally relevant education to truly eliminate educational gaps in Guatemala.

Keywords: Educational Gap, Indigenous, non-indigenous, Guatemala, Parental Education, Inequality.

doi:10.6342/NTU202504017

Contents

doi:10.6342/NTU202504017

doi:10.6342/NTU202504017

List of Tables

doi:10.6342/NTU202504017

List of Figures

doi:10.6342/NTU202504017

Chapter 1: Introduction

1.1 Objective of the Study

In Guatemala, an indigenous child has on average, almost 3 years less of education than a non-indigenous child (ENCOVI, 2023). This gap is not isolated, but reflects the structural inequalities that are part of Guatemalan history. Associated to factors, such as poverty, discrimination, geographic location, and ethnicity, despite Guatemala being characterized as one of the most stable economies in Central America, with a population exceeding 17 million and a Gross Domestic Product (GDP) that surpassed $104 billion in 2023. The nation has demonstrated consistent economic growth, which is reflected during 2014 and 2023 with an average rate of 3.2% exceeding the regional average. However, the economic expansion is not translated into result to reduce poverty or equitable improvement of living conditions, for its entire population (World Bank, 2025). With a significant portion of its population living in poverty conditions and an economy with a large informal sector, the development opportunities are not accessible for everyone. Inside this complex panorama, education emerges as one of the most powerful tools for social mobility and to reduce poverty, it is one of the fundamentals pillars for human development and to build a fairer society. However, in a multicultural society, the access and quality of education reveal a historical inequality between its ethnic groups.

Due to the high cultural and linguistic diversity in Guatemala, the State has taken various measures to preserve and promote respect for Indigenous groups. Although Spanish is the official language, Mayan languages are recognized as part of the national culture (Art. 143, Constitution of the Republic of Guatemala, 1985). The National Languages Law also

doi:10.6342/NTU202504017

states that the State recognizes, respects, and promotes the languages of the Mayan, Garifuna, and Xinca peoples.

Despite efforts to protect linguistic and cultural diversity, the Indigenous population has historically been subjected to racism, exclusion, and the impacts of the civil war (during the armed conflict from 1962 to 1997). With the signing of the Peace Accords, several institutions were established, such as the Commission Against Discrimination and Racism (CODISRA – Governmental Agreement No. 390-2002), the Ombudsman for Indigenous Women (Governmental Agreement No. 525-99), and the Law of the Academy of Mayan Languages of Guatemala (Decree No. 65-90). However, progress in improving the treatment of the Indigenous population remains limited, and significant gaps persist.

After the armed conflict and being one of the countries that invests the least in education, with only 3% of the total GDP—of which 99% in the primary sector is allocated to teachers—Guatemala still faces educational gaps that disproportionately affect the indigenous population. The Ministry of Education attempted to implement several programs to support and focus on the indigenous population, such as PRONADE (National Program for Educational Self-Management), which initially was a pilot program that showed deficiencies in implementation and limited resources (Alec Ian Gershberg B. M., 2009). PRONEBI (National Program of Intercultural Bilingual Education) was launched in 1965 to promote bilingual education, along with USAID; however, this pilot program lacked sufficient human capital (Harry Anthony Patrinos, 2009).

doi:10.6342/NTU202504017

However, comparative studies across Latin America show that the educational gap between the indigenous and non-indigenous population is still persistent and can largely be explained by family and contextual characteristics. Hernandez-Zavala et al. (2006) found that between 41% and 75% of the difference in school performance can be explained by endowments such as family and school resources, but the educational level of parents remains a predominant factor. On the other hand, Gershberg et al. (2009) emphasizes that educational decentralization in Guatemala does not guarantee proportional and equitable improvements, as local participation is important and still absent in many rural indigenous areas.

In this context our research is focusing into one of the most critical gaps that affects the future of the younger Guatemalans: the differences in average years of education between the indigenous and non-indigenous (ladino) population. Despite efforts, and general progress into educational coverage, the existence of this gaps highlights deep structural inequalities. Understanding the extent, this study focusing in the possibility of analyzing the educational gap in an analytical and disaggregated way, by examining its evolution between our two targets, addressing and empirical necessity by using a comparative approach analyzing two year (2006-2023). Through the Oaxaca-Blinder decomposition method, we are not only able to quantify the magnitude of the gap, but also to identify the key factors behind it. The used approach gives a clear understanding of the structural inequalities that persist in the country and provide valuable evidence for the future academic and policy debates regarding inclusive education.

In a context like Guatemala, where educational access shows that is limited to certain group, and the indigenous people have faced historically marginalization.

doi:10.6342/NTU202504017

The choice of the two periods compared is based on the information available through the ENCOVI surveys and taking into consideration the political and historical relevance of those periods in Guatemala.

The year 2006 represents the starting point of the study, as it is a decade after the Peace Accords (signed in 1996), where several important agreements were made and the Indigenous population was more specifically included in an effort to break inequalities and to see how effective these new policies and programs (PRONADE and PRONEBI) were in improving access to bilingual education and decentralizing the system.

The year 2023, almost two decades after our first reference point, not only reflects a period after the COVID-19 pandemic—when education worldwide was affected—but in Guatemala it had a greater impact on the most vulnerable areas and groups, such as rural zones and the Indigenous population. These two years allow us to analyze whether there have been changes in how education is delivered in Guatemala, and also whether there are effective and efficient policies that are contributing to more equitable education.

In this research three research questions were addresses: (1) What is the magnitude and composition of the gaps in years of education between the indigenous and non-indigenous population in Guatemala in 2006? (2) How has this gap evolved in magnitude and composition by 2023? And (3) What observable (endowments) and unobservable (coefficients and interaction) factors contribute to these gaps in each period, and how has their influence changed over the time? These questions allow for an static and dynamic analysis of the dimensions of educational inequality in Guatemala, providing a comparative and quantitative evidence that can inform the design of mor inclusive policies.

doi:10.6342/NTU202504017

The Oaxaca-Blinder decomposition method is a statistical technique used to analyze average differences and outcomes between two groups. This method separates the gaps into explained and unexplained components. The explained part (endowments) is reflected in the observable characteristics depending on the study—in this case, it could be age, sex, place of residence, etc. The unexplained component captures the differences in returns (coefficients), as well as the effects of unmeasured factors, such as structural discrimination, inequality, etc. This method has been previously used in studies on education, labor economics, and also in social research to analyze educational gaps, wage inequalities, and access to services.

Applying the Oaxaca-Blinder decomposition model is appropriate because it allows us to measure the magnitude of the educational gap between the two groups (indigenous and non-indigenous) in Guatemala, and also to compare the contribution of individual factors over time. Since the objective is to understand how educational inequalities are structured and evolve, this model is a powerful tool that helps distinguish between gaps derived from different characteristics (such as age, rural residence, parental education, sex, etc.) and gaps related to how these characteristics are valued in each group. This makes a perfect match with the research questions this study aims to address.

This study can be considered multifaceted research. Viewing from a theoretical perspective, the research is focusing on how to understand the persistent education inequalities transmitted in a context of intragenerational ethnic diversity, enriching debates about human capital and social stratification. Empirically, this study fills a gap by providing comparative and longitudinal analyses based on recent national surveys, that allows to identify the trends and evaluate the impact of the previous policies. From a

doi:10.6342/NTU202504017

public policy perspective, the findings have the potential to inform the strategies more focalized and culturally relevant. By understanding the sources of inequality, more effective interventions can be designed to promote greater educational equity and contribute to breaking the cycle of poverty in the most vulnerable communities in Guatemala.

The structure is divided into five chapters. Following by the introduction, Chapter 2 presents the theoretical framework and the literature review, giving us a conceptual study, presenting the context of education in Guatemala and its relationship with poverty and ethnicity. Chapter 3 describes the details of the research methodology, including the description of the data and the application of the Oaxaca-Blinder decomposition model. Chapter 4 will present the results of the quantitative analyses, in this case Oaxaca-Blinder decomposition for 2006 and 2023. And finally, Chapter 5 offers a comprehensive discussion of the results, draws the study's main conclusions, and formulates concrete policy recommendations to address educational gaps in Guatemala.

doi:10.6342/NTU202504017

Chapter 2: Literature review

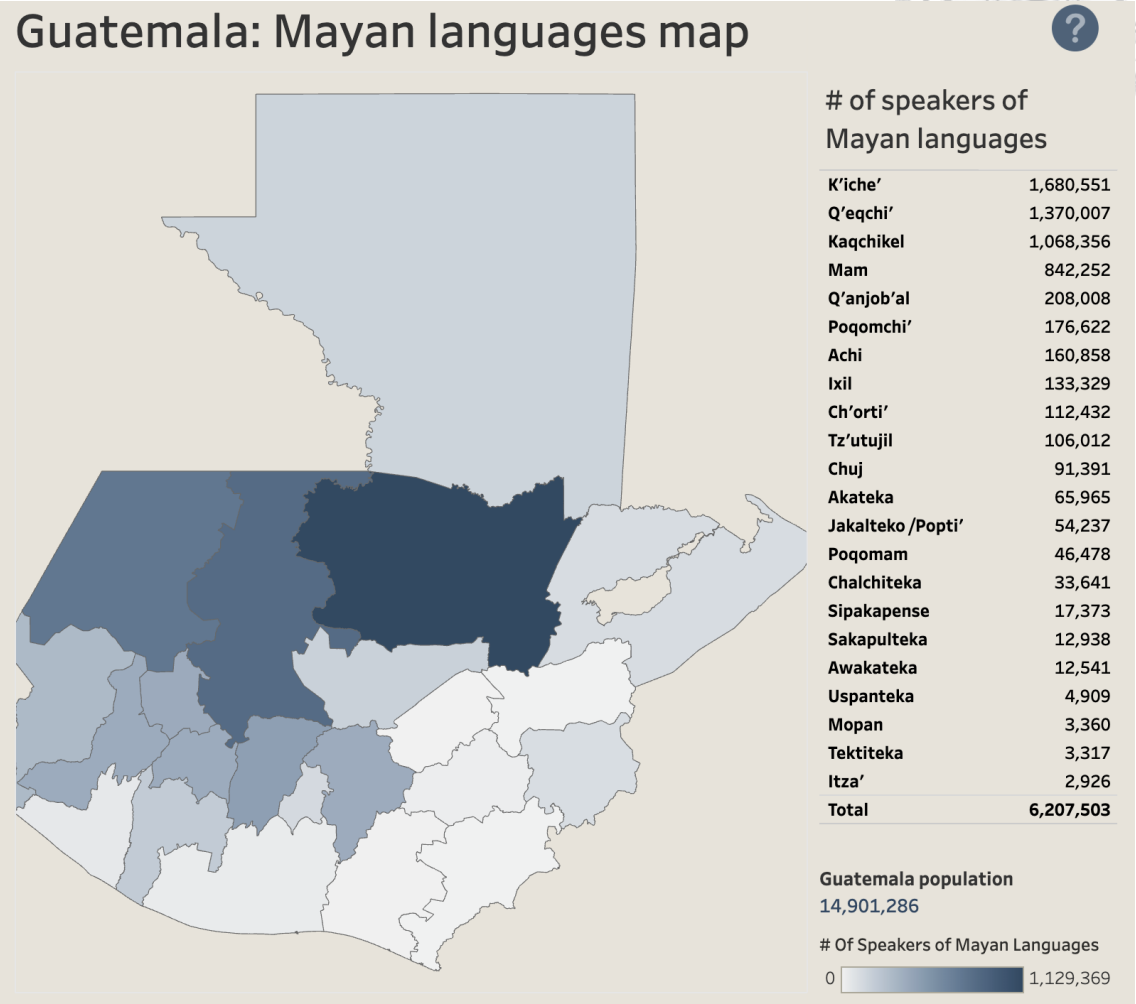
2.1 Socioeconomic and Demographic Background of Guatemala

Guatemala is a country in Central America, with an extraordinary cultural and geographic diversity, administratively divided in 22 departments, and 340 municipalities that are grouped in 8 regions. These elements have made of Guatemala a country with a complex scene in a socioeconomic and inequality perspective. With a population around 15 million inhabitants, according to the last census conducted it in 2018 (INE, 2018), It is considered one of the most populated countries in the region, hosting a rich multiethnic, multilingual and multicultural composition (PNDU, 2018).

To understand the history of Guatemala, it is important to mention that it was governed by indigenous/Mayan people for around 2000 years. These societies developed sophisticated systems of knowledge, writing, astronomy, and social and political organizations. They possessed their own methods to transmit cultural and education based on their languages, before the Spanish conquest in 1523. Therefore, part of the Guatemala's contemporary culture and customs is attributed to the legacy of the Mayan civilization. This diversity is manifested in the coexistence of 22 Mayan linguistic groups, together with Xinka, Garifuna, and the Ladino population, creating a social and cultural context (MINEX, 2020). The distribution of Mayans linguistics groups is showed in figure 1, We can observe the number of people who speak each of the registered languages in Guatemala. K'iche' is the most spoken language among the Indigenous population, with 1,680,555 speakers. The second most spoken Mayan language is Q'eqchi', with 1,370,007 speakers, followed by Kaqchikel in third place with 1,068,356. On the other hand, some languages have very small populations, such as Itza', which has only 2,926 speakers.

doi:10.6342/NTU202504017

Figure 1: Guatemala Division and Mayan Language Map



Note. Adapted from "Guatemala: Mayan Language map" by Translators without borders, using data from the 2018 census from INE

Nevertheless, the Spanish Conquest marked a devastating point in history for Guatemala. It was not merely a military confrontation and atrocities of war, but rather a violent imposition of a new colonial order that radically transformed social, economic, and cultural structures with Christianity, brought missionaries, that forcibly separated indigenous people from their ancient customs and traditions (Montejo, 2022). Also, Indigenous peoples were dispossessed of their lands, subjected to forced labor, under the rule of the Spanish Crown. Crucially, the Conquest established the Spanish language as the official and dominant language, relegating the various Mayan languages to a lower

doi:10.6342/NTU202504017

status, associated with oral tradition and the private sphere, and excluding them from the formal spheres of education, administration, and justice (Wiessner, 2020).

During the colonial period, accessed to formal education was extremely limited and stratified, reserved primarily for the principals Spanish and creole elites, with few opportunities for the mestizos and almost none for the vast majority of the indigenous population. The education was designed to consolidate the colonial control, promoting the language and Spanish culture, excluding the knowledge and local identity (Rivera, 1999). This systematic exclusion is one of the causes why the human capital is not reflected in the indigenous communities in Guatemala. Most of the indigenous communities only learned about religious doctrine and dogmas, which is one of the reasons that most of the missionaries learned the Mayan languages (Martínez, 1998)

After the independence in the 19th century, the colonial structures of ethnic and social inequality were not dismantled, but in many cases were reproduced or even accentuated under new republican forms. The Agrarian policies continued in favor of the landed elite, perpetuating the dispossession and marginalization of Indigenous communities. The investment in infrastructure, public services, and education, has been historically scarce in predominantly indigenous areas, widening the gap in the provision of quality services. Furthermore, discrimination and racism still persist, being significant barriers, that sometimes seem invisible, and limit the opportunities for social and educational mobility for indigenous peoples. Another conflict that affects Guatemala, was the internal armed conflict, which ravaged the country for decades in the 20th century, further exacerbated these inequalities, disproportionately, once again affecting indigenous communities and

doi:10.6342/NTU202504017

undermining their already limited access to essential services including education. (Neef, 2024)

Besides their cultural richness and being considered the country with the largest population and GPD in central America, with an average rate of 3.5% of growth (World Bank, 2025). Guatemala has faced persistent challenges in human development and equity. Historically speaking, the country has exhibited poverty and inequality indicators that are among the highest in Latin America, with significant gaps that are accentuated by geographic factors, like the division between its rural and urban areas, and, its ethnicity affiliation (CEPAL, 2022). One of the bases of the Guatemala's economy has been considering the agricultural sector, although with a growing participation of manufacturing and services industries, as well as a significant dependence on family remittances received from abroad. (World Bank, 2015). However, the distribution of the benefits of this growth has not been equitable, leaving a large segment of the population, particularly indigenous and rural communities in vulnerable conditions (Pop, 2013).

This context of profound socioeconomic disparities and marked cultural heterogeneity is the backdrop against which the study of living conditions and, specifically the relationship between education and poverty, with an ethnic focus is developed. Understanding, these particularities is essential to interpretating the data and findings related to the distribution of poverty and access to education in Guatemala.

doi:10.6342/NTU202504017

2.1.1 Guatemalan Educational System

The educational system in Guatemala is governed by the Political Constitution of the Republic and the National Educational Law (Decreto Legislativo 12-91). The educational system is structured into seven levels: pre-primary education, primary education (six years), secondary education which is divided into two cycles—the first one known as basic (three years), and diversified education (two or three years depending on the program)—and finally higher education, having only one public university known as Universidad de San Carlos de Guatemala (MINEDUC, 2024).

Despite legal frameworks and defined structure, the schooling system has historically faced significant challenges in terms of access, coverage, quality, and relevance, especially for the most vulnerable populations. Enrollment rates show a tendency to decrease dramatically as one advances through the educational levels, with a notable dropout rate in the transition from primary to high school education, and from high school to diversified education (Lopez, 2024). This educational progression gap-referring to the difference between initial enrollment and final completion rates- is even worse in rural areas and among the indigenous population, where factors like the distance to educational centers, the need to get involved in the labor market at a young age, and the lack of economic resources become insurmountable barriers for many.

2.1.2 Distribution of Poverty in Guatemala

Poverty in Guatemala is a complex reality and a socioeconomic indicator that has been present throughout the country's history, characterized by notable structural inequalities. In Guatemala, according to the Multidimensional Poverty Index (MIDES, 2018), poverty is commonly measured through a monetary approach using the poverty line, which

doi:10.6342/NTU202504017

establishes the cost of a basic food and non-food basket to satisfy the essential needs of Guatemalan households, 29% of Guatemalans live in multidimensional poverty, while an additional 21% are considered vulnerable. The average intensity of deprivation is 46%, and the national MPI is 0.134, revealing how non-income dimensions of poverty disproportionately affect indigenous and rural populations.

A distinction is made between general poverty and extreme poverty, the latter being the incapacity to cover the essential cost of a food basket. The ENCOVI (2023) shows that 56.0% of the population is living in poverty, an estimated 16.2% in extreme poverty, and 39.8% not in poverty, although 44% is considered non-poverty. At the beginning of the 2000s, Guatemala occupied one of the highest positions in poverty rates in Latin America, reflecting persistent socioeconomic challenges despite periods of economic growth.

The distribution of poverty in Guatemala shows a clearly differentiated path that sharpens the discussion on equity. Geographically speaking, poverty is predominant in rural areas according to INE (2024), with a rate of 66.3%, where communities living outside urban areas concentrate a higher number of poverty and extreme poverty cases compared with principal cities. This phenomenon is associated with factors such as the lack of infrastructure, restricted access to quality basic services (health and education), dependence on subsistence agriculture, and the scarcity of formal employment opportunities.

However, the most critical and persistent inequality in Guatemala is observed in the distribution of poverty by ethnicity. Data from the 2000s already consistently revealed that the indigenous population suffered from significantly higher poverty rates than the

doi:10.6342/NTU202504017

ladino population (INE, 2015). This gap was not marginal; poverty and extreme poverty rates among Indigenous peoples could be double or even triple those observed among the non-Indigenous population. This disparity is not explained solely by the rural location of most Indigenous communities, but is exacerbated by a series of historical and structural factors, including discrimination, unequal access to productive resources, limited political participation, and the language barrier to accessing services and opportunities (IWGIA, 2020).

The unbalanced distribution of poverty, concentrated in rural areas where the majority of the population is indigenous, has created disadvantages in access to education, quality services, and employment. The poverty condition, whether general or extreme, limits families' capabilities to invest in education and causes most children to work, perpetuating an intergenerational cycle. And it is here where the analysis of the educational gaps is important to understand this reality.

2.2 Theoretical and Conceptual Framework

2.2.1 Human Capital and Education

Human capital is a fundamental concept in economy and development, that is **recognized** as the stocks of skills, competences, knowledge, and attributes that a person **possesses** and are relevant to generate economic and social value (Diebolt, 2016). In the Oxford English Dictionary, human capital is defined as "the skills, knowledge, and experience possessed by a workforce, considered a valuable resource or asset." It encompasses the notion that there are investments in people (e.g., education, training, health) and that these investments increase an individual's productivity (Goldin, 2016).

doi:10.6342/NTU202504017

The idea that human capital is an investment began to gain strength towards the end of the 1950s, especially with the contributions of economists such as Jacob Mincer (1958). The education is the principal driver of human capital foundation and accumulation. By attending to school and advancing through the various educational levels, individuals acquire not only technical and academic knowledge, but also cognitive (critical thinking, problem-solving) and non-cognitive (discipline, perseverance, social skills) skills that are highly valued in the labor market and in civic life (Woessmann, 2008). Therefore, education is seen as a powerful tool to improve the quality life and employability of the people (Psacharopoulos, 2018)

This is showing that education can be an actor or opportunities, allowing people from disadvantages socioeconomic backgrounds to acquire the skills needed to access to better jobs, that will be reflected in higher wages and ultimately, break the intergenerational cycles of poverty (Solon, 1999). As Model of Lucas (1988) says that human capital is related with two factors: 1. Is the quality of education y and 2. Is the investment that people dedicate to study, and that the human capital through education is fundamental to social and economic development.

2.2.2 Poverty and Social Exclusion

Poverty is not just the lack of money; is a complex phenomenon with multiples facets that deprives people of opportunities and essential capabilities for a dignified life (Sen, 2001). Understand poverty required analyze the diversity theories focus that try to capture the true magnitude and their implications. Although, poverty has historically been

doi:10.6342/NTU202504017

measured using a monetary approach, but academic and public policy discussions have evolved toward a broader, multidimensional perspective.

The monetary approach to poverty measures poverty in terms of income or consumption and concentrates on a person's or household's capacity to get necessities (Ravallion, 1998). To this end, poverty lines are established, which represent the minimum income threshold necessary to cover food and non-food needs considered essential. If a person or family's income falls below this line, they are considered to be living in poverty. Within this approach, a distinction is often made between general poverty (inability to cover the entire basic food basket) and extreme poverty (inability to cover even the basic food basket) (Feres, 2001). This method is practical and easy to quantify, what makes it useful to faster comparations and tracking trends over time. However, one of their main limitations is that can't capture the crucial deprivations that affect quality of life, access to education, health and a decent housing.

In the face of the limitations of the monetary approach, a multidimensional approach to poverty has gained traction. This perspective holds that poverty is not simply a matter of low income, but rather an accumulation of deprivations across various dimensions of human well-being (Sabina Alkire, 2010). Poverty is considered when people experiment simultaneous deficiencies in areas as: health (lack of basic services and nutrition) Education (insufficient years of schooling, non-attendance) living conditions (as inadequate housing, lack of access to drinking water, sanitation, and clean energy) and also employment (Sabina Alkire J. E., 2015). This approach is very useful and helpful in a way to better understand poverty and the dimension of poverty of the individuals, to better design public policies.

doi:10.6342/NTU202504017

The social exclusion refers to the conditions where individuals are systematically marginalized or some of the basic resources, opportunities and laws are not available for them, but are available for the rest of the population (Gonzales, 2002) Some of the exclusion can be manifested in different aspects as were mentioned before, some of it could be access to labor market, quality public services, and more. The exclusion could exist in different ways and levels, some of the examples of social exclusions are when a certain individual the access to education is denied. The exclusion is manifested to different causes including economics, cultural, socials and politics (UNIR, 2024)

2.2.3 Intersection between: Education, Poverty and Ethnicity

Adopting an intersectional viewpoint is crucial to comprehending disparities in educational results and access in a nation like Guatemala. This idea, which was first put forth by Kimberlé Crenshaw [1989] in the framework of Black feminism, acknowledges that social identities—such as gender, ethnicity, socioeconomic status, and geography—do not function independently but rather interact to produce distinct experiences of privilege and oppression. Collins (2019) In others words, being an indigenous woman and poor is not simple the sum of being woman plus indigenous plus poor; it's a qualitatively experience that generate specific cumulative disadvantages.

The intersection between poverty, education and ethnicity in Guatemala is manifested through a set of barriers that limited the fully potential and development of the indigenous peoples. These barriers are not random: they are the result of structural mechanism and historically rooted forms of discrimination that have placed the indigenous population in a situation of systemic vulnerability.

doi:10.6342/NTU202504017

Structural Barriers and Discrimination:

I. Historical Discrimination: Indigenous people in Guatemala have historically been marginalized and excluded from the political and economic power. This has result in an inequality in the distribution of the resources and opportunities including the investment in infrastructures and basic services for their communities (Grandin, 2004). Leaving to indigenous people outside of the power atmosphere the public policies are created in a way that force to exclude indigenous people.

II. Discrimination in the Education System: Discrimination can also be a factor withing educational facilities, with stereotypes by teacher or classmates, or in the lack of recognition and appreciation of the indigenous culture and languages (UNESCO, 2023) Discrimination affects many factors as the academic performance, self-esteem and motivation to continue studying.

III. Territory and Geographic Location: Most of the indigenous population lives in disperse rural areas, where the access to different services is not easy (INE,2018). Regarding the quality of education, still exist marked disparities, which are notable in the rural area and indigenous communities where is often a lack of basis resources, such as teaching material, libraries, access to technology and, remarkably, teachers with adequate skills and working conditions (Figueroa, 2023). This precarity is translated into inequality result of learning, where students who live in these areas frequently obtain lower scores on standardized assessments (Lopez, 2024).

doi:10.6342/NTU202504017

According to the Inter-American Commission on Human Rights (2015) specific challenges facing the indigenous population and rural areas are profound and multifaceted consequences of the structural racism and discrimination. According to Jacir de Lovo (2022) the access gaps are aggravated by geographic dispersion, the lack of adequate school infrastructure in remote communities, and the distances that children must travel from their houses to the schools on foot because of the inadequate roads is around approximately 1 hour. Additionally, socioeconomic factor such as the need for children and adolescents between 11 and 13 years old to enter the labor market to support their families are overwhelming barriers that force to many students to drop out of school (ENCOVI, 2023)

Language Barriers and Lack of Cultural Relevance: Guatemala is a multilingual country, where the Mayan, Xinka, and Garifuna languages are the native languages of a significant proportion of the indigenous population. Cultural and linguistic relevance is another critical pillar; despite the recognition of Guatemala's cultural diversity, the effective implementation of bilingual and intercultural education that respects the native languages and worldviews of indigenous peoples remains a considerable challenge, which can lead to disengagement and demotivation in the learning process (Herrera Larios, Cifuentes Estrada, & Otzoy Chipix, 2023). When the teaching-learning process is conducted only in Spanish, without the effective implementation of bilingual education (IBE), a fundamental barrier to student understanding and participation is created (ICEFI, 2011). This language gap not only affects academic performance in specific subjects but can also lead to frustration and demotivation.

doi:10.6342/NTU202504017

In addition, the lack of cultural relevance into the curricula is a challenge. If the educational contents don't reflect the history, worldview, values and knowledge of the indigenous communities, the school could perceive as an alien to their reality and culture, limiting its ability to generate meaningful learning and connection with their roots, and this disconnect contributes and is one of the causes of the school drop-out and an accumulation of human capital and aspirations of Indigenous communities (Jacir de Lovo, 2022).

The intersectionality of ethnicity with the poverty and geographic location creates a complex set of disadvantages that limit the access of education and that indigenous population can succeed in Guatemala. Recognize the connection of this factors is important to understand the educational gaps that are going to be presented in the next chapters and the implication with the development of human capital.

doi:10.6342/NTU202504017

2.3 Review of the Empirical Literature

Several studies have been previously conducted at global, regional, and national levels, documenting how educational gaps are closely linked to socioeconomic conditions and racial affiliation, with indigenous populations being the most affected. This research contributes to existing research by offering a long-term comparative analysis (2006-2023) of how the educational gap has evolved between the indigenous and non-indigenous populations of Guatemala, using the Oaxaca-Blinder decomposition model. This model not only seeks to identify persistent inequalities, but also to understand the factors that explain them and how they have evolved or changed over time, providing useful empirical evidence for the design of inclusive and targeted public policies.

2.3.1. Studies about Education and Poverty in Latin-American

The relationship between education and how it reduces poverty has been a central subject of the research regarding development throughout the years. The global empirical literature has shown that the investment in education, in individuals' levels that later is reflected into national levels, is an important factor to improve mobility, quality life, and socioeconomic development. In terms of global return research has shown that it is a 9% of private returns, and in terms of social returns, it is seen as a strategy for development (Psacharopoulos, Patrinos, 2018). But education itself must be accompanied by good education, which is focused on quality rather than quantity, since simply increasing schooling does not increase the skills of individuals; it must be accompanied by useful teachings that train individuals in an accurate manner to achieve good development (Eric Hanushek, 2008).

In Latin America, this relationship is particularly relevant, give the persistence of high levels of inequality and poverty which have huge gaps of achievement and educational

doi:10.6342/NTU202504017

opportunities between socioeconomic groups. Some studies say that is important not only analyze the improvement in education, but also the persistent inequalities (Cruces et al., 2012).

Numerous studies in the region have analyzed how differences in human capital, primarily education, explain income gaps and socioeconomic disparities, in 2010 according to Barro (2010) the gap was around 4 years, that shows the persistent of inequality in education, and also that the gaps in education are attached to the persistent in poverty (Barro, 2010) this research often use robust quantity methodologies, including regression models, and decomposition technics to isolate the effect of education in others socioeconomic factors. An example, had been documented how the increase in schooling years in Latin-American countries had contributed to a reduction in poverty, even though this effect depends in other factors as education quality and the context of the labor market (Lopez-Calva, 2013)

The regional literature has also begun to place greater emphasis on the quality of education, beyond the years of schooling. It recognizes that school attendance alone does not guarantee the acquisition of skills relevant to the labor market of life and that deficiencies in educational quality can limit education's potential to foster social mobility and reduce poverty, even when coverage is high (Marina Bassi, 2012). This is important to understand: despite there being progress in access to education, the development gaps still persist.

The studies about education show and highlight the capability that education has and the transformational power it has for societies against poverty and social mobility, but they

doi:10.6342/NTU202504017

also remark about the complexity of this phenomenon, indicating that the effectiveness of education depends on its quality and how the structural inequalities and barriers are addressed.

2.3.2 Studies about Guatemala and Indigenous Groups

To understand the complex dynamics of poverty and education in Guatemala, it is essential to examine the research that has addressed the problematics from an ethnic perspective, with a special emphasis on indigenous peoples.

There are several investigations that have conducted studies on Guatemala and indigenous groups, providing solid evidence of the problems experienced in rural areas and among indigenous groups. Many of them lead to the conclusion that indigenous groups and rural populations experience significantly higher levels of poverty and lower educational achievements.

One of the pioneering works on bilingual education and conceptual perspectives is Klaus Zimmerman's (1997) work, "Modes of Interculturality in Bilingual Education: Reflections on the Case of Guatemala," which offers a crucial perspective for the evolution of intercultural bilingual education.

Zimmerman (1997) emphasizes that intercultural bilingual education is the most advanced and necessary response to centuries of assimilationist education in Latin America. The reason is because, according to Zimmerman, IBE aims for the self-determined development of Amerindian populations, as it promotes the integration of their own culture and language into the educational process. Interculturality aims at and seeks the coexistence and mutual enrichment of cultures and is not merely seen as a policy derived from the negative consequences of assimilation. In the case of Guatemala, the

doi:10.6342/NTU202504017

situation of indigenous peoples is complicated by the integration of Mayan languages into the educational system (Zimmerman, 1997).

Zimmerman recognizes that historically, bilingual education systems have been limited to simple translation or transitional bilingualism, without addressing the intrinsic structure and cultural richness of indigenous languages. He argues that the difficulty lies not in an inherent "lack of structure" of the Mayan languages, since these, like any language, are complete and coherent linguistic systems, but in the absence of formal recognition and pedagogical standardization, taking into account their academic value and social equivalence to those of the Spanish language. His conclusions are that true interculturality demands going beyond the mere implementation of the vernacular language in the classroom; it requires a paradigm shift that validates and enhances the worldviews and forms of transmission of indigenous cultures. This requires not only an educational methodology that integrates Mayan languages and knowledge, setting aside any discrimination (Zimmerman, 1997).

In a recent analysis, Jacir (2022) details research highlights how existing gaps in Guatemala hinder equitable access to quality education. She identifies three main types of gaps:

1. Access and coverage: She mentioned that despite the progress in educational coverage, a large part of the population remain underserved, especially in rural areas. An example is that in 2018 there was still an educational lag in 4 out of 10 people aged 7 years and older. Dropout in Guatemala is one of the highest in Latin America and the Caribbean. Also, she highlights the problem of distance from schools, which is a significant barrier.

doi:10.6342/NTU202504017

2. Quality and relevance, this includes gaps in infrastructure, maintenance, and teaching materials, as well as the quality and preparation of teachers, and curricular relevance.

3. Socioeconomic and cultural: Poverty, which is closely related to the difficulties that arise in accessing education. Discrimination against ethnic minorities and groups that have historically been affected, and finally, it mentions gender roles and stereotypes, emphasizing that girls face greater obstacles, citing marriage, childhood, and teenage pregnancy as examples.

Some of the conclusions that Jacir (2022) reaches are that, despite the advances in educational coverage rates, there are still gaps in quality, access, and efficiency persist in the most vulnerable populations. These gaps underscore an urgent inclusive and quality transformation is necessary, taking into account equity in learning. She also concludes that indigenous peoples, who have been historically marginalized must be taken into account and must be included in decision-making, since they have the highest illiteracy rate among rural indigenous women, and that bilingual education must be seen as a necessity for the entire educational system and not as a subgroup.

In a more practical and critical look at the implementation of intercultural bilingual education (EBI) in Guatemala, Mo Isem (2019) traces the trajectory of EBI from its initial recognition in the 1980s with the return of democracy and the recognition of Indigenous rights to its establishment as an official modality within the Guatemalan education system at the beginning of the 21st century.

doi:10.6342/NTU202504017

Mo Isem (2019) highlights that in the initial phases of implementation and transition, not only were Mayan languages made visible, but also there was an increase in Indigenous enrollment and a reduction in school dropout rates in these communities. He also argues that EBI has not only been a tool for assimilation (or transition from Spanish) but also strengthens the ethnic and linguistic identities of Indigenous peoples.

Among the gaps that exist are the shortage of teachers trained in Mayan languages, the lack of relevant materials, and the weak state commitment to Indigenous education. Furthermore, the value of indigenous languages as learning tools and fundamental components for the comprehensive development of indigenous children is emphasized, taking into account that the integration of EBI should not only be linguistic but also encompass the worldview and epistemology of indigenous peoples, recognizing them as the axis of knowledge. Another gap is the monolinguistic challenge in Spanish, since many indigenous children entering the school system are now monolingual in Spanish, which becomes a fundamental challenge for the effectiveness of EBI programs.

doi:10.6342/NTU202504017

2.3.3 Oaxaca-Blinder Decomposition Studies

The Oaxaca-Blinder decomposition model has been widely used in the social sciences to analyze observed disparities between groups, especially in education and income. One of its key features is that it can separate the observed gap or difference between two groups into a portion explained by observable characteristics (age, education, geographic area) and an unexplained portion, which may be differences in returns to the characteristics or factors not included in the model.

The Oaxaca-Blinder model has been used in various studies on wage gaps, analyzing income differences between men and women or between ethnic groups, to identify discrimination and segregation that is partly due to disparities in education or experience. For example, for studies regarding income deprivation, there are studies that have used Oaxaca-Blinder decompositions and provide models estimating the probability of a household falling into poverty (Ponce 2006, for Ecuador; Telles and Lim 1998; and Montaño 2004, for Brazil). Using the latter approach, Hall and Patrinos (2005) find that being indigenous increases the probability of being poor in Bolivia (13 percent), Ecuador (16 percent), Guatemala (14 percent), Mexico (30 percent), and Peru (11 percent). In terms of people with disabilities, poverty seems to exacerbate disabilities (Dudzik et al. 2002; Hernández-Jaramillo and Hernández-Umaña 2005) (Gandelman, 2011).

In the area of segregation and discrimination, we can mention studies such as Marconi (2004), which analyzes the wage gaps between the private and public sectors in Brazil. Even after adjusting for productive and demographic factors, the disparity remains apparent. The existence of gender and skin-color salary disparities in the public sector is also documented in this paper, with the latter being greater than the former. Arias et al.

doi:10.6342/NTU202504017

(2010) carried out a similar study, where whites and Afro-descendants in Brazil were included in their study to analyze the wage gaps between these two groups, using quantile regressions and Oaxaca-Blinder decomposition. The results indicate that wage differences that cannot be attributed to differences in human capital increase with job responsibilities, regardless of skill level. It is also observed that average wage gap measures provide limited information on the situation of disadvantaged groups, which indicates that it is necessary to use techniques that emphasize the distributions of the variables of interest (Gandelman, 2011).

In research related to education, McEvan (2008) used the Oaxaca-Blinder decomposition method to analyze the differences in average performance of indigenous and non-indigenous students in eighth grade in Chile. His results show that in 3 years the average differences decreased by 0.2 standard deviations in Spanish and 0.1 in mathematics. The decomposition method showed that the decreases were not due to convergences in the observed family characteristics, but rather that the majority of decreases occurred within the school with students of similar socioeconomic levels, and to a lesser extent between schools, in the measure of school quality (McEwan, 2008).

The ability to unravel the sources of disparities makes the Oaxaca-Blinder decomposition an invaluable tool for this research. It allows us to go beyond simply describing the gaps in years of education between ethnic groups in Guatemala to quantify how much of that gap is due to differences in their characteristics (such as parental education or age) and how much remains "unexplained," opening the door to a deeper discussion about structural and contextual factors.

doi:10.6342/NTU202504017

The literature review shows us the importance of education as a tool to reduce poverty and close gaps in inequalities that have persisted over the years. However, we can also see that despite the efforts and policies that have been attempted to implement that we saw throughout the literature, the efforts are still not sufficient, inequalities are still persistent, and the efforts that have been made so far still do not fully meet the needs of certain groups that have been marginalized for years. This research focuses on providing long-term empirical evidence by incorporating differences from the years 2006 and 2023, allowing us to identify the factors behind this inequality and generate information that can be useful for the design of more conscious policies taking into account the needs of the entire population.

doi:10.6342/NTU202504017

Chapter 3. Data

3.1 Data Source and Sample

The data used in this research comes from the National Survey of Living Conditions (ENCOVI) 2006 and 2023, collected by the National Statistical Institute (INE) of Guatemala. The national representative national surveys which include detailed information on the sociodemographic, educational, economic, and housing characteristics of Guatemalan households. The ENCOVI, as its name suggests, is a nationally representative household survey designed to provide a broad set of detailed indicators on issues like housing, health, education, unemployment, poverty, and inequality is its main objective. Because of this, it is a useful resource for studies on growth and well-being.

The temporality is due to the fact that these two years respond to the interest in analyzing the evolution of educational gaps between ethnic groups over time, in different sociopolitical contexts. ENCOVI 2006 offers a post-armed conflict panorama, with structural changes following the signing of the peace accords in 1996- including policy shift in education coverage and the promotion of intercultural bilingual education (as discussed in the literature review). while ENCOVI 2023 presents a more recent context, with ongoing challenges and following the impact of the COVID-19 pandemic.

The data were implemented using a complex sampling design, which includes stratification and cluster selection (primary sampling units). This design guarantees the representativeness of the sample at the national level and for key domains such as urban/rural areas and major geographic regions. To ensure that the results obtained from the analysis are representative and valid for the entire Guatemalan population, the use of

doi:10.6342/NTU202504017

expansion factors or sample weights, (pweight = factor), is incorporated. The use of sampling weights is crucial for statistical analysis, as they adjust the contribution of each observation in the sample to reflect a true proportion of the total population (Rutstein, 2006). Their correct use is essential to obtain unbiased estimates and adequate variances in any statistical inference.

We included only individual aged 15 and older, because this group allows us to observe educational trajectories based on concluded or nearly concluded periods of education. We excluded foreigners because they are not part of the Guatemalan population, and those who don't belong to any of the categories of non-indigenous or indigenous. The ethnic group was defined according to self-identification, based on the specific criteria of the survey. In addition, all missing values were excluded in any of the main variables, in order to consider a consistent sub-sample.

After applying those filters, the final samples are 23,522 for 2006 and 16,228 for 2023, considered a wide sample that allows a robust comparative analysis between both periods, in general terms and by analytical subgroup (ethnicity, gender, area, etc.). This chapter present a clear methodological foundation that supports the robustness and relevance of the findings discussed in the following section.

doi:10.6342/NTU202504017

3.2 Variables

This section describes each variable used in the Oaxaca-blinder decomposition model.

3.2.1 Dependent variable:

- **Years_education (years of education)** In Guatemala, education is compulsory for nine years—six years of primary and three years of lower secondary education (MINEDUC, 2023). This legal threshold helps contextualize the average years of schooling observed across population groups in this study.

  Are the years of formal education achieved per person. It's a standard measure of the human capital and the educational level achieved.

  We measure as a continuous variable that indicates the total numbers of approved years in the educational system in Guatemala (as an example, 6 years for primary education, 9 years for secondary education, 12 years for high school, etc.) For those who haven't receive education the value is 0. The calculation was made according to the highest education level obtained, according to the categories declared in the survey. This constitutes as a result variable for this analysis.

3.2.2 Independent Variables/explicative

The following independent variables were included in the regression model to explain the variation in the educational years and were part of the Oaxaca-Blinder decomposition model as observable characteristics

- **Age**

  Age of the person in years.

  This is measured as a continuous variable and represent the age of the person at the moment of the survey. This variable was used to control the difference related with the accumulation of education in the life cycle of the individual, and the

doi:10.6342/NTU202504017

effects of generational cohorts on education. A non-linear or negative relationship is expected at older ages, reflecting access to education in different historical periods.

- **Female**

  This is a dichotomous variable to identify the gender of each person.

  As a binary variable we defined as following 1=woman and 0=man, as is incorporated to evaluated the differential impact of the gender to access and accumulate education, considering possible historical and cultural gaps that affect in a particular way to women.

- **Rural**

  This is a dichotomous variable to identify the residence area where the person lives. Was defined as a binary variable, where 1=rural area and 0=urban area.

  Was included to capture the structural differences to access to education and the quality of the educational institutions between rural and urban environments that can also affects the educational levels.

- **Father's Years of education (Age_educ_dad)**

  As the name says, this continuous variable expressed in years represents the total years of formal education of the father of each individual.

  Reflects the human capital of the father, that is also a key indicator of family socioeconomic status and the educational environment at home. It's expected that a highest father's education can have a positive association with son's education.

- **Mother's Years of education (Age_educ_mom_)**

  This variable is similar to the father's education, and as the name says, this continuous variable expressed in years represents the total years of formal education of the mother of each individual.

doi:10.6342/NTU202504017

Reflects the human capital of the mother, and is considered that maternal education is a robust predictor of children's educational achievement, often with an even greater impact due to the mother's role in early education and socialization, and as in the father's education a positive association is expected in this one.

- **Moderate Poverty (_IPOBREZA_2)**

  Dichotomous variable that indicates if the individual is in a situation of moderate poverty. The classification is not self-reported, but rather determined by the official monetary poverty thresholds established by the INE, based on whether the individual's household income allows for the purchase of a basic food and non-food basket. This binary variable takes the value 1 = moderate poverty and 0 = not poor (reference category corresponds to _IPOBREZA_1). This variable captures the potential impact of economic hardship on human capital accumulation, as financial limitations may affect access to or permanence in the educational system.

- **Extreme Poverty** (_IPOBREZA_3)

  Dichotomous variable that indicates if the individual is in a situation of extreme poverty. This is a binary variable where 1=Extreme Poverty and 0=no poor. The category that is the implicit reference for this analysis is the one for "No Poor" (that belongs to _IPOBREZA_1). Similar to the variable of moderate poverty and also is threshold-based according to the official INE definitions, but this one capture the impact of the severe conditions to access to education. It's expected that both poverty variables show a negative association with the years of education, but the result will be interpreted according to the model. variable This variable allows to evaluate the impact of the economic condition of the family regarding

doi:10.6342/NTU202504017

the human capital. The economic limitations can restrict the access or permanence into the educational system.

- **FACTOR**

  This variable represents the sample weight assigned to each individual in the survey. It's a continuous variable that reflects the number of persons on the population that represent each observation in the sample. It's used as a countability weight (pweight) in all the models to make sure that the result is representative of the population of Oaxaca and to adjust the design in the complex survey.

- **Ethnic**

  Dichotomous variable, it is based on self-identification, following the structure used in the ENCOVI surveys conducted by the National Institure of Statistics (INE).

  Respondents were asked: **"¿Se considera perteneciente a algún pueblo o grupo indígena?"** (Do you consider yourself a member of any Indigenous people or group?) If the respondent answered **"Sí"**, they were then asked to specify the group (e.g., Maya, Xinka, or Garífuna). Those who answered "No" were classified as non-Indigenous

  For this analysis, a binary variable was constructed where:

  1 = Indigenous (including all those who self-identified as Maya, Xinka, or Garífuna)

  0 = Non-Indigenous

This operationalization is consistent with INE's classification system and aligns with national standards for disaggregating data by ethnic origin. The variable allows us to analyze differences in educational outcomes between Indigenous and non-Indigenous

doi:10.6342/NTU202504017

populations, reflecting ethnic self-identification as recorded in the original survey instruments.

3.3Descriptive Statistics

This section presents descriptive statistics for the variables used in the analysis, for the 2006 and 2023 ENCOVI surveys. These means and their linearized standard errors, along with confidence intervals, provide a first approximation of the characteristics of the study population and the temporal trends of the variables of interest. It is important to highlight that all estimates have been weighted using expansion factors (FACTOR) to ensure the representativeness of the results at the national level, considering the complex sample design of both surveys.

Table 1:Summary Statistics (ENCOVI 2006-2023)

Variable	Mean 2006	95% CI 2006	Mean 2023	95% Ci 2023
Years of Education	5.30	[5.20 - 5.40]	6.12	[6.019 - 6.21]
Age	42.04	[41.71 - 42.37]	45.16	[44.83 - 45.48]
Father's Education	2.16	[2.08 - 2.23]	2.40	[2.31 - 2.48]
Mother's Education	1.56	[1.49 - 1.62]	1.79	[1.71 - 1.85]
Female	0.56	[0.55 - 0.57]	1.58	[1.57 - 1.59]
Rural	0.49	[0.48 - 0.50]	0.47	[0.45 - 0.47]
Ethnicity	0.38	[0.37 - 0.39]	0.41	[0.40 - 0.42]
Poverty	0.44	[0.43 - 0.45]	0.50	[0.48 - 0.50]

Source: Author's own calculations based on data from the National Survey of Living Conditions (ENCOVI) 2006 and 2023. Analyses conducted using Stata. Sample weights provided by the survey were applied.

The table 1 presents the descriptive statistics of the main demographic and socioeconomic characteristics of the sample for the years 2006 and 2023. This data was obtained from the National Survey on Living Conditions (ENCOVI) for each period. This detailed comparison helps to identify significant changes in the profile of the Guatemalan population over nearly two decades, which is crucial information for understanding the observed disparities.

doi:10.6342/NTU202504017

A key finding from Table 1 is the notable increase in the average years of education for the general population. It rose from approximately 5.30 years in 2006 to 6.12 years in 2023. This increase, while modest, suggests general progress in the country's educational level. This could be due to greater coverage and retention rates at the primary and basic education levels. However, this advancement happened alongside a worrying change in poverty conditions. The proportion of the Guatemalan population living in poverty significantly increased from 43.64% in 2006 to 49.89% in 2023. This shows a general decline in socioeconomic well-being, indicating that almost half of the population was living in poverty in the most recent year.

Regarding demographic characteristics, the average age of respondents slightly increased from 42.04 years in 2006 to 45.16 years in 2023, which reflects a slightly older sample. Parental education levels also showed a modest rise; the average education for fathers increased from 2.16 to 2.40 years, and for mothers, it went from 1.56 to 1.79 years during the same period. While these increases are subtle, they could point to a gradual improvement in the human capital of parental generations, which is relevant for the intergenerational transmission of education.

The gender composition of the sample remained relatively stable, with the proportion of women showing a minimal change from 56% in 2006 to 58% in 2023. Concerning residence, the proportion of the rural population slightly decreased from 48.72% to 46.60% between 2006 and 2023, suggesting a gradual trend towards urbanization. Finally, the proportion of the population self-identifying as indigenous in the sample saw a slight increase, from 37.93% in 2006 to 41.48% in 2023. This change in demographic

doi:10.6342/NTU202504017

composition is particularly relevant for our decomposition analysis, given our focus on the educational gap between ethnic groups.

Overall, these descriptive data illustrate a complex picture of progress mixed with persistent challenges. Improvements in years of education contrast with the worsening poverty conditions. These average characteristics for both periods represent the fundamental 'endowments' for the subsequent Oaxaca-Blinder decomposition analysis. This allows us to understand how differences and changes in these characteristics contribute to the educational gap between indigenous and non-indigenous populations.

doi:10.6342/NTU202504017

Chapter 4 Data analysis methods:

This study uses a quantitative method to analyze educational gaps in Guatemala, focusing on differences geographic regions, gender and ethnic groups. This approach was chosen because of its ability to examine relationships between variables, identify patterns, test hypotheses, and generate results that may be extrapolated to a broader population though the use of statistical methods. The study is analytical and explanatory in character, aiming to identify the sources of these disparities in order to understand them, as well as to define the differences that currently exist.

The methodology was designed in the interest of estimate the differences in the average years of schooling between indigenous and non-indigenous people, controlling for variables such as age, sex, area of residence, language learned in childhood, poverty, and parental educational level. Combining descriptive analysis with econometric techniques allows us to break down the observed educational gap into explained components (associated with differences in observable characteristics) and unexplained components (which, may be linked to structural discrimination or other variables that were not directly measured).

To quantify and understand the reasons behind the gap in years of education between the ethnic groups (indigenous and non-indigenous) in Guatemala, the Oaxaca-Blinder decomposition method was used. This method is a standard tool in econometrics and the social sciences, it is designed to analyze group differences in an average outcome variable between two groups by separating the contributions of differences in observable characteristics from differences in the coefficients (or "returns" to those characteristics) (Oaxaca, 1973). This method was selected for its ability to provide a detailed view of the

doi:10.6342/NTU202504017

driver of educational inequality in a context of social heterogeneity. The Oaxaca-Blinder model is based on Ordinary Least Squares (OLS) regressions, estimated independently for each group. In this study, two separate regressions were estimated: one for the indigenous population and one for the non-indigenous population, following the general formulation:

$$Y_g = X_g \beta_g + \epsilon_g$$

Where:

- Y_g: represents the years of education for group g (indigenous or non-indigenous).
- X_g: is a vector of observable characteristics (age, sex, area of residence, years of education of father and mother, and poverty status) for group g.
- β_g: is a vector of estimated regression coefficients for group g, representing the "returns" or marginal impact of each characteristic on years of education.
- ϵ_g: is the error term.

Once the regressions are estimated for each group, the total gap in average years of education between the non-indigenous group (N) and the indigenous group (I) ($\Delta Y = \bar{Y}_N - \bar{Y}_I$) is decomposed into three principal components. The standard formulation of the decomposition is as follows:

$$\Delta Y = (\bar{Y}_N - \bar{Y}_I) = \underbrace{(\bar{X}_N - \bar{X}_I)\hat{\beta}_N}_{\text{Endowment Effect}} + \underbrace{\bar{X}_I(\hat{\beta}_N - \hat{\beta}_I)}_{\text{Coefficient Effect}} + \underbrace{(\bar{X}_N - \bar{X}_I)(\hat{\beta}_I - \hat{\beta}_N)}_{\text{Interaction Effect}}$$

For practical interpretation purposes in the analysis of educational inequality, the decomposition is commonly presented by grouping the 'Coefficient Effect' and

doi:10.6342/NTU202504017

'Interaction Effect' into a single 'Unexplained Component'. This approach allows for a clearer distinction between differences attributed to observable characteristics and those reflecting structural factors or unobservable effects. Thus, the decomposition is analyzed in two main components:

1. **Explained Component (characteristics and functions):**

$$[(\bar{X}_N - \bar{X}_I)\hat{\beta}_I]$$

This component measures the portion of the gap in years of education, that is the result of the differences in observables characteristics between the indigenous and non-indigenous population. For example, if the non-indigenous group, has on average, parents with more years of education o with a lower proportion of members in poor conditions, this component quantifies how much of the total gap can be explained regarding their “endowments” of characteristics. It is interpreted as the gap that would exist if both groups have the same return to their characteristics, but the endowments of the reference group (indigenous group)

2. **Unexplained component (Due to coefficient/Discrimination/Unobservables)**

$$[\bar{X}_I(\hat{\beta}_N - \hat{\beta}_I) + (\bar{X}_N - \bar{X}_I)(\hat{\beta}_I - \hat{\beta}_N)]$$

This component captures the portion of the gap that cannot be attributed to observable characteristics. It reflects the differences in the 'returns' (coefficients) that each group receives for their characteristics, even if these characteristics were the same across groups. This component can indicate factors such as structural discrimination, institutional barriers, differential quality of education received, or a lack of cultural and linguistic relevance, elements that are not directly measured

doi:10.6342/NTU202504017

as explanatory variables in the model. While 'discrimination' is a common interpretation, it is important to recognize that this component can also include the effect of relevant unobservable variables.

All statistical analyses, including the estimation of OLS regressions and the Oaxaca-Blinder decomposition, were performed using the statistical software Stata (version 16) using the oaxaca command, the one allows to made detailed decompositions and specific for linear models. It is important to highlight that, to ensure the representativeness of the results at the national level and to correct for the complex sample design of the 2006-2023 ENCOVI, expansion factors or sampling weights (pweight = FACTOR) were applied in all estimates. This guarantees that each observation contributes to the analysis in proportion to its actual representation in the Guatemalan population.

doi:10.6342/NTU202504017

4.1 Model Specification:

Two versions of the model were run per year:

- **Baseline Model**: Includes sociodemographic variables: age, sex, area (rural/urban), language learned in childhood, ethnicity, and poverty.
- **Expanded Model**: Adds parental education (father's and mother's years of schooling) to capture the intergenerational effects of human capital.

It is important to note that the sample size differs between the baseline and expanded models due to missing data in the parental education variables. As a result, the expanded model uses a slightly smaller sub-sample, which may affect direct comparability between the two specifications, but increases explanatory power by integrating intergenerational human capital factors.

Methodology Considerations:

- The results are reported with robust standard errors and 95% confidence intervals.
- Binary dummies were created for categorical predictors such as ethnicity and poverty.
- Emphasis is placed on interpreting the explained and unexplained components and connecting them to structural processes like discrimination, exclusion, and inequality.

doi:10.6342/NTU202504017

Chapter 5. Results of the Analysis

The results of the Oaxaca-Blinder decomposition are presented to quantify the gaps in years of education for the indigenous and non-indigenous population in Guatemala during 2006 and 2023. The findings are broken down to show the magnitude of the gap, the distribution of observable (endowments) and unobservable (coefficients and interaction) factors. The general results are explored first, and then will go deep in the specific contribution of each variable, and finally, the dynamics across different subgroups will be analyzed.

The following section examines how the national average educational gap, has evolved over nearly two decades.

5.1 Overall Educational Gap Decomposition Baseline and Extended Models

The results of the general models of the Oaxaca-Blinder decomposition (baseline and extended) in 2006 and 2023 show significative trends of educational inequality. Table 2 summarizes the total gap and its principal components, meanwhile Figure 2 shows in a visual way the proportion of each component.

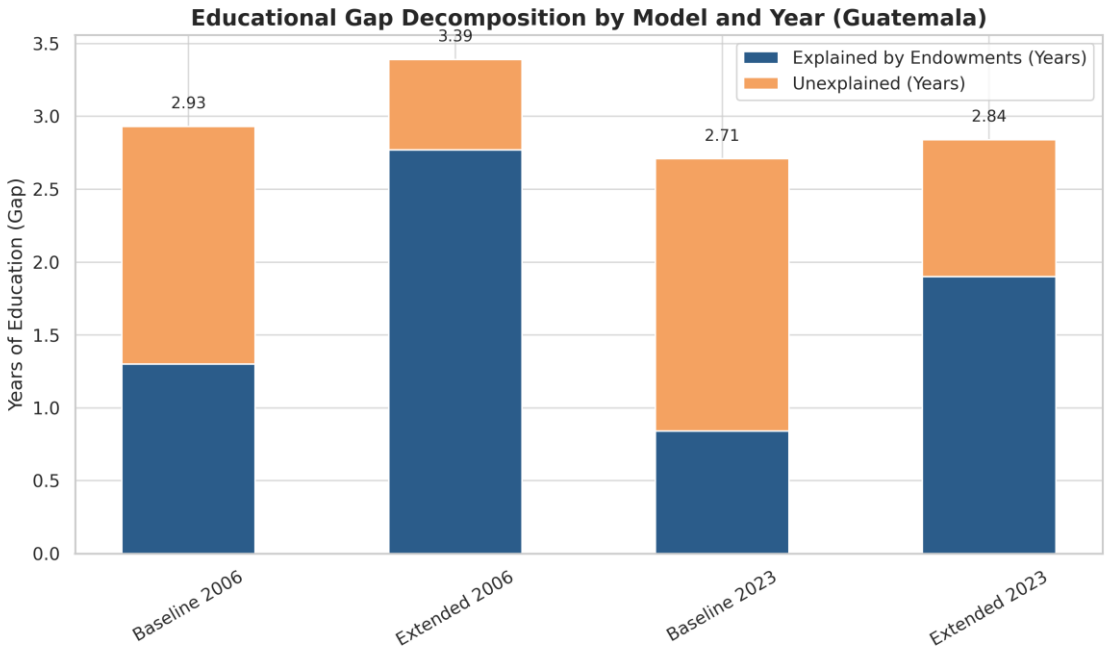
doi:10.6342/NTU202504017

Table 2: Overall Oaxaca-Blinder Decomposition (2006 & 2023)

Model and year	Total Gap (Years)	Explained by Endowments (Years)	% Explained by Endowments	Unexplained (Years)	% Unexplained
Baseline 2006	2.93	1.3	44.4	1.63	55.6
Extended 2006	3.39	2.77	81.7	0.62	18.3
Baseline 2023	2.71	0.84	31.0	1.87	69.0
Extended 2023	2.84	1.9	66.9	0.94	33.1

Source: Author's own calculations based on data from the National Survey of Living Conditions (ENCOVI) 2006 and 2023. Analyses conducted using Stata. Sample weights provided by the survey were applied.
Note: Percentages are calculated based on the Total Gap. The Explained Component (Endowments) was calculated using the Oaxaca-Blinder decomposition, specifically the term $(\bar{X}_N - \bar{X}_I)\hat{\beta}_I$, which represents the portion of the gap explained by differences in observable characteristics (age, sex, area, poverty, language, and parental education). The Unexplained Component is the sum of the Coefficients and Interaction effects. Due to missing data on parental education, the sample size used in the extended model is smaller than in the baseline model

Figure 2: Educational Gap Decomposition by Model and Year



Note. Author's own elaboration based on data from the National Survey of Living Conditions (ENCOVI) 2006. Sample weights were applied
(Note: This figure illustrates the gap in years of education between the non-indigenous and indigenous populations within each sex category.)

doi:10.6342/NTU202504017

The graph summarizes how the educational gap between the indigenous and non-indigenous populations in Guatemala has changed, considering the studied years (2006 and 2023) and using two analytical approaches: the baseline model (which does not include parental education) and the extended model (which includes parental education as an explanatory variable).

Decomposition Results:

In 2023, the total gap remained relatively stable: 2.71 years in the base model and 2.84 years in the extended model. However, what changed was the composition of the gap. In the base model, the unexplained component increased to 69%, suggesting that observable differences are no longer sufficient to explain the gap. With the inclusion of parental education in the extended model, the explained component reached 66.9%, a value still relevant but lower than the 81.56% observed in 2006.

This graph shows that parental education has historically been one of the most important factors in explaining educational inequality in Guatemala. Although it remains relevant in 2023, its relative weight has decreased compared to other unobserved factors. Furthermore, the persistence of the unexplained component, especially in the base models, reveals the possible existence of more complex structural barriers, such as discrimination, unequal quality of schools, or cultural factors that are not directly captured by the surveys. These results suggest that public policies should go beyond access and consider equity in learning environments.

In 2006, the general educational gap stood at 2.93 years for the baseline model and 3.39 years for the extended model. The extended model shows that when parental education is

doi:10.6342/NTU202504017

included, it captures a significantly larger portion of the gap (81.7% in 2006), highlighting the importance of the parents' characteristics.

5.1.1 Contribution of Variables to the Explained Gap (Extended Models)

In order to understand the observables factors that drive the educational gap, it is essential to examine the individual contribution of each variable within the explained component by endowments. Table 3 and 4, with figures 3 and 4, details this contribution for extended models of 2006 and 2034.

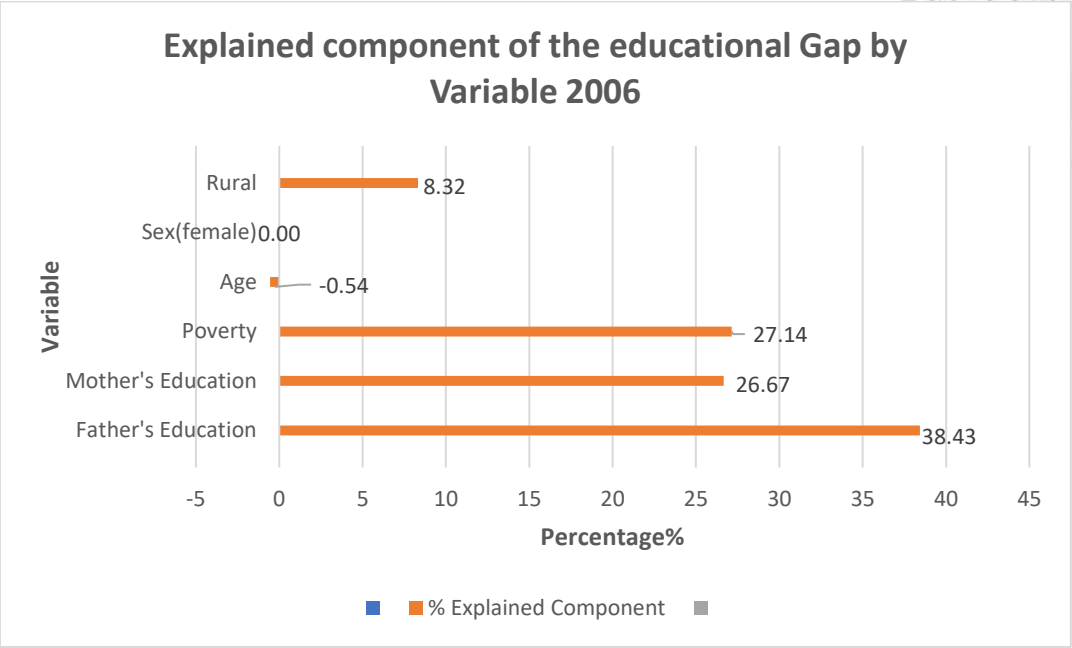
Table 3: Variable Contributions to Explained Gap: Extended Model (2006)

Variable	Endowments (Years)	% Explained Component
Father's Education	1.062	38.42
Mother's Education	0.737	26.66
Poverty	0.75	27.13
Age	-0.015	-0.54
Sex(female)	0.0001	0.00
Rural	0.23	8.32
Total of endowments explained	2.76	100

Source: Prepared by the authors based on data from ENCOVI 2006
Note: Percentages are calculated based on the Total Explained Endowments from the 2006 Extended Model.

doi:10.6342/NTU202504017

Figure 3: Explained Component of the Educational Gap by Variable



Note. Author's own elaboration based on data from the National Survey of Living Conditions (ENCOVI) 2006. Sample weights were applied

The results show that the variables with the greatest explanatory weight in the educational gap between the two groups studied (indigenous and non-indigenous) are parental education variables-especially father's education- with approximately 1.06 years of education, representing 38% of the total explained. This demonstrates the influence of intergenerational human capital, with an emphasis on the paternal side, on educational opportunities.

In second and third place are poverty and maternal education. Poverty represents 0.75 years of education, which translates to 27.14%, and maternal education, 0.74 years of education, or 26.67%. Adding the percentages of these three variables, we can see that they represent and explain more than 90% of the total explained of the gap, suggesting that structural disadvantages in families and the socioeconomic background of individuals are the main factors behind the observed inequalities. Variables such as sex, age, and area

doi:10.6342/NTU202504017

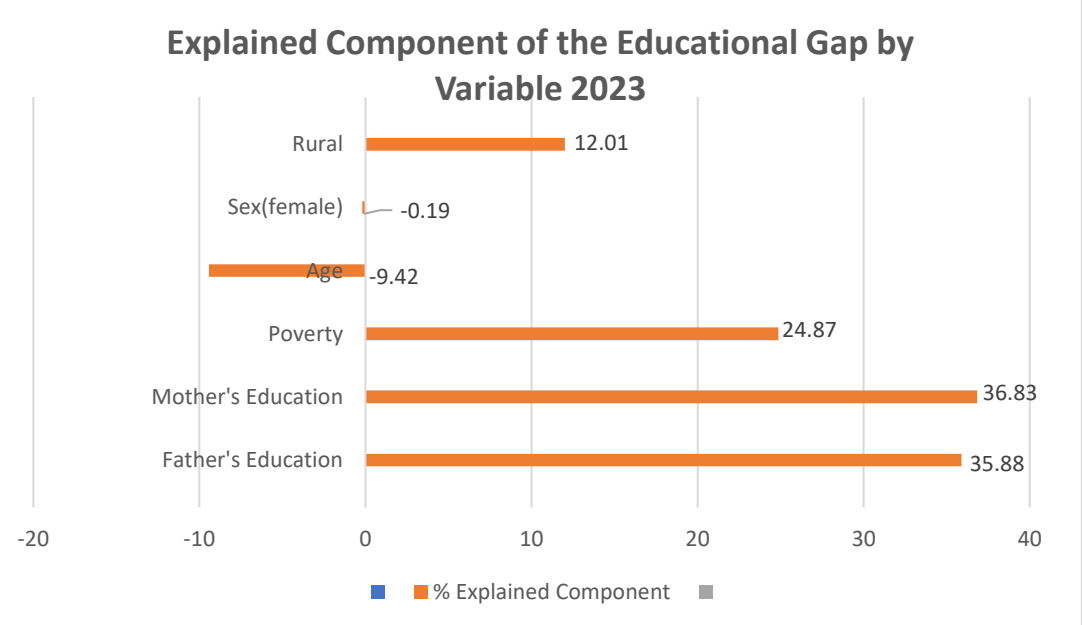
(urban and rural) have a minimal effect, as we can see they are close to zero. The age variable shows a negative contribution of -0.015, which can be interpreted as not favoring either group.

Table 4: Variable Contributions to Explained Gap: Extended Model (2023)

Variable	Endowments (Years)	% Explained Component
Father's Education	0.68	35.88
Mother's Education	0.70	36.83
Poverty	0.47	24.87
Age	-0.18	-9.42
Sex(female)	0.00	-0.19
Rural	0.23	12.01
Total of endowments explained	1.89	100.00

Source: Author's own calculations based on data from the National Survey of Living Conditions (ENCOVI) 2006 and 2023. Analyses conducted using Stata. Sample weights provided by the survey were applied Note: Percentages are calculated based on the Total Explained Endowments from the 2023 Extended Model.

Figure 4: Explained Component of the Educational Gap by Variable 2023



Note. Author's own elaboration based on data from the National Survey of Living Conditions (ENCOVI) 2023. Sample weights were applied

doi:10.6342/NTU202504017

In 2023, father's (35.88%), and mother's education (36.83), are still predominant factors explaining the educational gap, maintaining a similar proportion of the total explained contribution (more than 70% combined) However, their absolute contribution in years shows that father's education significantly decreased, while mother's education had a slightly increased compared to 2006. This indicates a possible adjustment in differences in parental endowments between groups, or in the way that are translated into educational achievements.

Between 2006 and 2023, the main explanatory factors of the educational gap remained centered in parental education, while fathers' education was the strongest factor in 2006 (38.43%), by 2023 its contribution had decreased (35.88%), while mother's education increased (from 26.67% to 36.83%). This suggests a more balanced influence of both parents over the time. Others changes observed, were poverty that shows a slightly decreased, while the negative effect of age become more pronounced in 2023 (-9.42%), which could indicate deeper structural shifts in how age correlates with educational outcomes.

doi:10.6342/NTU202504017

5.2 Educational Gaps by Subgroups: Specific Analysis (2006 and 2023)

Beyond the overall educational gap, this section delves into a more detailed analysis of the differences in years of education by disaggregating the Oaxaca-Blinder decomposition across key dimensions. This approach allows us to observe how the ethnic educational gap (between Indigenous and non-Indigenous populations) varies across two distinct time periods (2006 and 2023), by geographic environment (rural and urban areas), by gender (men and women), and under both Baseline and Expanded model specifications. This comprehensive disaggregation is crucial for understanding how these contextual factors and model choices shape educational opportunities and inequalities for the Indigenous population.

5.2.1 Differences Gaps between Rural and Urban Areas

The distribution of the educational gap between areas (rural and urban) reveals important contrasts in the Guatemalan context. Table 5 presents the decomposition results for both environments in 2006 and 2023, and Figure 5 provides a visual comparison.

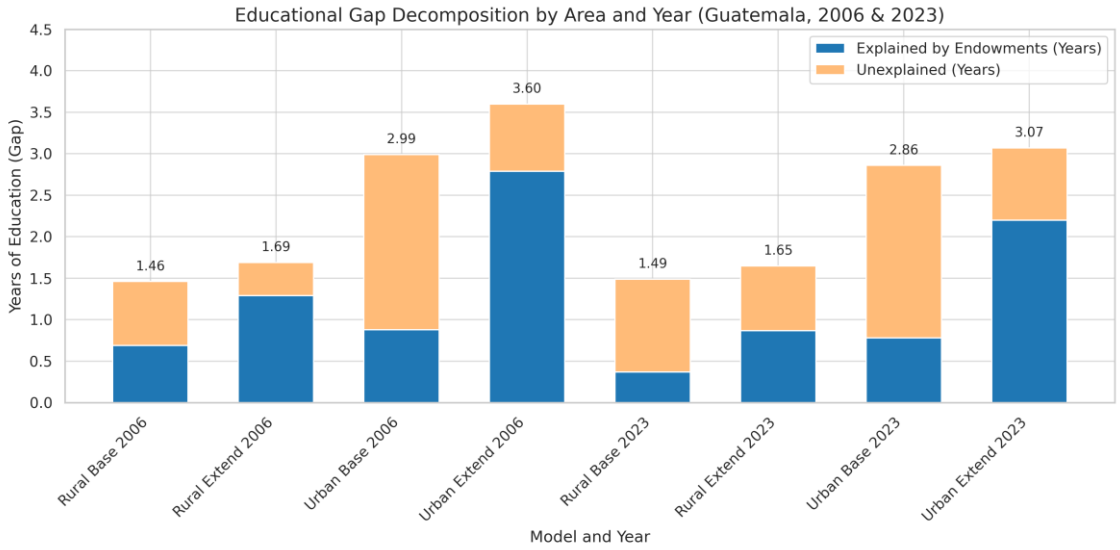
doi:10.6342/NTU202504017

Table 5: Oaxaca-Blinder Decomposition by Area & Model (2006 & 2023)

Year	Area	Model	Total Gap (Years)	Explained by Endowments (Years)	% Explained by Endowments	Unexplained (Years)	% Unexplained
2006							
	Rural	Base	1.46	0.69	47.3	0.77	52.7
	Rural	Extend	1.69	1.29	76.3	0.4	23.7
	Urban	Base	2.99	0.88	29.4	2.11	70.6
	Urban	Extend	3.6	2.79	77.5	0.81	22.5
2023							
	Rural	Base	1.49	0.37	24.8	1.12	75.2
	Rural	Extend	1.65	0.87	52.7	0.78	47.3
	Urban	Base	2.86	0.78	27.3	2.08	72.7
	Urban	Extend	3.07	2.2	71.7	0.87	28.3

Source: Author's own calculations based on data from the National Survey of Living Conditions (ENCOVI) 2006 and 2023. Analyses conducted using Stata. Sample weights provided by the survey were applied
Note: Percentages are calculated based on the Total Gap. The Unexplained Component is the sum of the Coefficients and Interaction components. An explained percentage greater than 100% or a negative percentage for the unexplained component indicates that a group's endowments (in this case, urban Indigenous people) are even better, but other unfavorable gaps persist.

Figure 5: Educational Gap Decomposition by Area and Year



Note. Author's own elaboration based on data from the National Survey of Living Conditions (ENCOVI) 2006. Sample weights were applied

doi:10.6342/NTU202504017

Figure 5 illustrates the educational gap between the indigenous and non-indigenous (Ladino) populations in Guatemala, disaggregated by areas of residence (rural vs. urban), and year (2006 and 2023). Each bar shows the total gap in years of education, divided into two components:

1. The blue represents the explained component (attributable to observable differences such as age, sex, area, and in the extended model, with the parental education)
2. Orange bar represents the unexplained component that could be associated with discrimination, educational quality, or unobserved factors.

Year 2006:

In rural areas, the gap was smaller (1.46 years), and almost half (47.3%) could be explained by endowments. With the inclusion of parental education in the extended model, the explained component increased significantly (76.3%), reducing the weight of the unexplained component to only 23.7%. This indicates that the educational disadvantages of rural Indigenous peoples were largely explained by their lower access to structural conditions such as family education.

In urban areas, the gap more than doubled (2.99 years). In the base model, only 29.4% was explained by observable endowments, and the remainder (70.6%) could not be easily explained. However, when considering parental education, the explained component rose to 77.5%, revealing the strong role of family background in urban contexts.

doi:10.6342/NTU202504017

Year 2023:

In rural areas, although the total gap barely changed (1.49 years), the explained component fell significantly in the base model (24.8%), while the unexplained component rose to 75.2%. This suggests that in rural areas, visible structural factors explain less of the gap than in the past, which could be due to changes in population composition or persistent inequalities that are difficult to capture. However, the extended model improves the explanation to 52.7%, although it still leaves 47.3% unexplained.

In urban areas, the total gap remained high (2.86 years in the base model), and the unexplained component remained dominant (72.7%). In the extended model, the inclusion of parental education increased the explained component to 71.7%, confirming that this factor remains decisive in urban settings.

The results show that parental education is a key element for understanding educational inequalities between ethnic groups, especially in urban areas. In rural areas, while its addition improves the explanation, significant unexplained components persist, suggesting the need for more in-depth and culturally relevant policies. Furthermore, the increase in the unexplained component in 2023, especially in rural areas, poses a challenge for equity strategies, as it reveals that there are still invisible barriers that have not been resolved despite educational advances.

In 2006, the baseline model shows a gap of 1.46 years, with a 47.3% explained by basic endowments. However, when the parental education was incorporated into the extended the model, the educational gaps in rural areas increased 1.70 years, and the percentage explained by endowments soared to 76.3%, highlighting the strong intergenerational

doi:10.6342/NTU202504017

influence of the human capital in this environment. In 2023, the rural gap stayed in similar ranges (1.50 years in baseline and 1.66 years in the extended model), but the explained power of endowments decreases in both models (24.8% in baseline and 52.7% in extended), suggestion that a large part of the inequality is based into unobservable factors.

In contrast, in urban areas, the total gap was considered higher than in the rural areas in both years. In 2006, the urban gap was 2.99 years in the baseline model (29.4% explained) and of 3.60years in the extended model (77.5% explained). In 2023, the urban gap remained higher (2.86 year in the baseline, and 3.07 in extended model), even though the extended model still explains a significative part (71.4%), the unexplained component persists strongly. The key difference is that, when parental education is included, the gap in the Extended model is consistently larger than in the Baseline, indicating that differences in parental human capital are a significant driver of educational inequality across cities.

5.2.2 Gender Gap (Men and Women)

The analysis of the educational gap base on the gender variable offers a value perspective, about how the dynamics of gender interact with the ethnic identities in the access to the education. Table 6 and figure 6 present a result of the decomposition for men and women in 2006 and 2023.

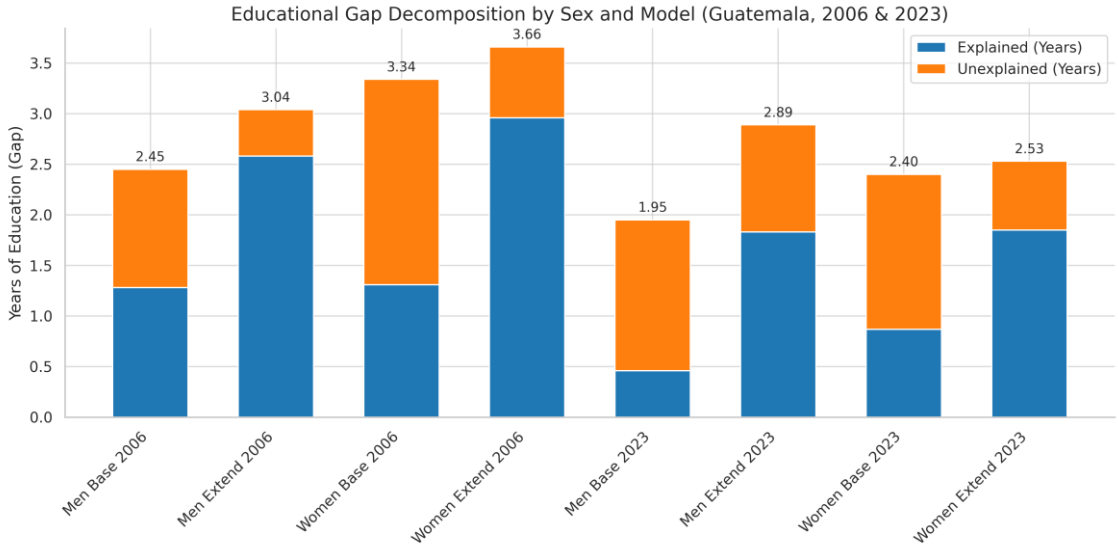
doi:10.6342/NTU202504017

Table 6: Oaxaca-Blinder Decomposition by Gender & Model (2006 & 2023)

Year	Gender	Model	Total Gap (Years)	Explained by Endowments (Years)	% Explained by Endowments	Unexplained (Years)	% Unexplained
2006							
	Men	Base	2.45	1.28	52.2	1.17	47.8
	Men	Extend	3.04	2.58	84.9	0.46	15.1
	Women	Base	3.34	1.31	39.2	2.03	60.8
	Women	Extend	3.66	2.96	80.9	0.7	19.1
2023							
	Men	Base	1.95	0.46	23.6	1.49	76.4
	Men	Extend	2.89	1.83	63.3	1.06	36.7
	Women	Base	2.4	0.87	36.3	1.53	63.8
	Women	Extend	2.53	1.85	73.1	0.68	26.9

Source: Prepared by the authors based on ENCOVI data from 2006 and 2023.
Note: Percentages are calculated based on the Total Gap. The Unexplained Component is the sum of the Coefficients and Interaction components.

Figure 6: Educational Gap Decomposition by Sex and Model



Note. Author's own elaboration based on data from the National Survey of Living Conditions (ENCOVI) 2006. Sample weights were applied

This graph presents a visual comparison of the educational gaps between the indigenous and non-indigenous populations, disaggregated by sex, year, and model type (baseline vs.

doi:10.6342/NTU202504017

extended). Each bar represents the total gap in years of schooling, broken down into two components: the explained component (attributed to differences in observable characteristics such as age, residence, and parental education) and the unexplained component (potentially related with discrimination, unobserved barriers, or other structural factors).

5.2.3 Decomposition Results:

Year 2006:

Women was the larger gap (3.34 years in the baseline model and 3.66 in the extended model) compared to men (2.45 and 3.04 years, respectively). However, when parental education was included, the explained component increased considerably for both cases, especially for women (80.9%), indicating that a larger part of the educational inequality can be attributed to unequal family conditions.

Despite this, the unexplained component remained significant, especially in the base model for women (60.8%), suggesting the presence of structural factors that disproportionately affect Indigenous women.

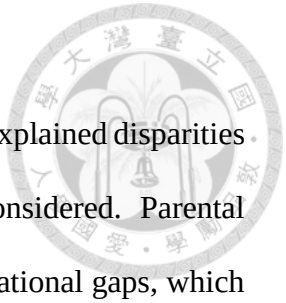
Year 2023:

The total educational gap decreased for both genders compared to 2006, reflecting progress in educational access.

However, in the baseline model, for men the unexplained component increased (76.4%), suggesting that, without considering family factors such as parental education, inequality appears less attributable to observable characteristics.

Once again, when parental education is included in the extended model, the explained component rises (up to 73.1% for women and 63.3% for men), confirming the continuing importance of intergenerational human capital on current educational inequalities.

doi:10.6342/NTU202504017



This result highlights that although gaps have narrowed over time, unexplained disparities remain significant, especially when parental education is not considered. Parental education is also observed to have a stronger effect on women's educational gaps, which could be associated with gender roles in the Guatemalan context.

In 2006, the educational gap for men was 2.45 years in the baseline model (52.2% explained), increasing to 3.04 years in the extended model, where 84.9% was explained, indicating a strong role of parental education. In the case of indigenous women in 2006, the gap was higher 3.34 years in the baseline model (39.2% explained) and 3.66 years in the extended model (80.9% explained. These results highlight, that, even women faced a higher gap, the parental education was an important fact to explain part of the differences for both genders.

In 2023, the total gap for men decreased to 1.95 yeas in the baseline model (23.6% explained), rising 2.89 years in the extended model (63.3% explained) In the case of women in 2023, the gap remained higher: 2.40 years in the baseline model (36.3% explained) and 2.53 years in the extended model (73.1% explained). It’s crucial to observe than, even though parental education continues to be an important explanatory factor in the extended model for both genders, the unexplained component had a significative increase in 2023, specially for women (26.9% in the extended), suggesting that the barriers persist, limiting educational opportunities beyond the observables characteristic can capture.

doi:10.6342/NTU202504017

5.3 Summary of the Main Findings of the Chapter:

This chapter presents the results of the Oaxaca-Blinder decomposition, offering a deep vision of the educational gap between the indigenous and non-indigenous population in Guatemala during the years of 2006 and 2023. The main findings from this analysis are as follows:

- **Persistence of the educational Gap:** The gap in years of education between the two groups has persists over the two decades, with similar values in 2006 and 2023. This highlights a structural challenge in the educational system in Guatemala.
- **Change in the composition of the gaps:** While the overall of the gap remained relative stable, a slightly change occur in its composition. In 2006, the biggest part of the gap was explained by the differences in its observable characteristics (endowments) between the two groups. However, by 2023, the unexplained component (attributable to differences in the "returns" to those characteristics or unmeasured factors such as quality of education, discrimination or systematic barriers) gained a considerable weight, suggesting that the causes of inequality have become more complex and subtle.
- **Dominant influence of the parental education:** Across the two periods, parental education emerged as the more influential factor to explain the educational gap. This highlights the importance of the intergenerational transmission of human capital and the environment at home in determining the opportunities of indigenous youth.
- **Geographic disparities:** The rural areas continue experiment a prominent educational gap than urban areas. In 2006, differences in endowments explained the most of the rural gap. However by 2023 a significant increase in the

doi:10.6342/NTU202504017

unexplained component was observed in these areas, pointing to growing challenges beyond basic socioeconomic characteristics.

- **Differences by gender:** The educational gap affects in a higher proportion to indigenous women, who in 2006 showed a slightly higher gap than indigenous men. By 2023, the unexplained gap component increases for both genders, been slightly more prominent for women, which could indicate the presence of specific gender barriers que impede equitable progress in the educational attainment.

doi:10.6342/NTU202504017

Chapter 6

6.1 Discussion of Results

Building upon the notable educational gap observed in the descriptive statistics (as presented in Table 1, for instance) between the indigenous and non-indigenous populations in Guatemala, the application of the Oaxaca-Blinder decomposition model allows for a deeper understanding of the factors contributing to this difference. This analysis reveals the underlying components of a gap that, unfortunately, proved to be persistent across both 2006 and 2023.

While the unexplained component of the Oaxaca-Blinder decomposition cannot be directly measured, its persistence—especially in 2023—suggests the influence of deeper structural barriers. Prior research indicates that these may include systemic discrimination, the limited presence of culturally relevant pedagogical approaches, and the unequal valuation of educational achievement in the labor market for Indigenous populations (Montejo, 2022; Psacharopoulos & Patrinos, 2018). Furthermore, the limited availability of trained bilingual teachers and the tendency to reduce intercultural education to mere translation have contributed to an ineffective inclusion of Indigenous communities in the educational system (Herrera Larios et al., 2023).

In her analysis of educational inequality in Guatemala, Jacir de Lovo (2022) emphasizes that the country continues to prioritize coverage over quality, failing to adapt teaching practices and curricula to the sociocultural realities of Indigenous students. Therefore, the unexplained component may reflect the accumulation of these systemic gaps—elements that are not easily captured by sociodemographic variables but that profoundly affect the educational trajectories of marginalized populations.

doi:10.6342/NTU202504017

In both periods, the results are clear: non-indigenous people have achieved, on average, more years of schooling than indigenous peoples. This disparity, while concerning, is not surprising if we consider the deep historical context and structural inequalities that, over time, have relegated indigenous peoples in access to essential basic services such as education, health, and employment opportunities (Jacir de Lovo, 2022; ECLAC, 2021).

Analyzing the education gap in depth with the Oaxaca-Blinder decomposition provide a more detailed look, breaking that difference down into what we can "explain" and what seems "unexplainable." In the baseline model, that includes fundamental socioeconomics and demographics characteristics such age, sex, and area of residence, we find that a significant portion of the disparity in years of schooling is due to factors we can observe and measure. This explained component in the baseline model, show that, unfortunately, indigenous students often start from less favorable backgrounds and conditions. For example, they are more likely to live in rural areas with less school infrastructure, have less or limited access to educational materials, and come from homes where their parents had fewer opportunities to study.

But what is most revealing is that a significant portion of this gap persists, particularly in the baseline model, even after accounting for all these observable differences. This segment, which we call "unexplained" or "residual," may be due to more complex and sometimes less visible factors. These includes factors such as the discrimination they still face, the low cultural relevance of the educational system that doesn't connect with their knowledge, or even differences in the "returns" (benefits) that education offers them in the labor market compared to other groups. These differences, which are not readily

doi:10.6342/NTU202504017

apparent, are particularly important because they suggest that, even if the starting conditions were similar, Indigenous students could still encounter additional obstacles in their educational journey (Zimmermann, 1997; Mó Isém, 2019).

Comparing the analyses of the baseline and extended model, where we include the parental educational variable, we can notice something crucial: the inclusion of this factor consistently and significantly increases the portion of the gap that can be explained with observable characteristics. This underscores the fundamental role played by human capital passed down from generation to generation in maintaining inequalities. In other words, children of parents with low educational attainment—a reality more common among the Indigenous population, according to our data—have fewer opportunities to advance in their studies, which unfortunately fuels a cycle of poverty and exclusion.

Fortunately, when looking at the total gap in years of education over time, we see a slight reduction in both models, even though values of the total gap were different. This could be an encouraging sign that there is progress in educational coverage or that some public policies targeting vulnerable populations are bearing fruit. However, and this is important, the "unexplained" component persist in both models and years, even when its relative magnitude change when we incorporate the parental education. This warns us that deep and entrenched inequalities still exist that cannot be resolved solely with formal access to school, but rather require structural actions far beyond that.

Ultimately, our results reinforce a vital idea for Guatemala: to truly close the educational gaps, it is not enough to ensure that all children can enroll into school. It is essential to transform the education system so that it is genuinely relevant to their culture, inclusive,

doi:10.6342/NTU202504017

and sensitive to the realities of life in Indigenous communities. At the same time, it is crucial to address the root causes of the structural factors that perpetuate poverty and social exclusion in the country.

doi:10.6342/NTU202504017

6.2 Conclusions

After a rigorous analysis of the educational gap between the indigenous and non-indigenous population in Guatemala, this study has revealed persistent patterns and significant changes in the dynamics of inequality over nearly two decades (2006-2023). The main conclusions drawn from our research are the following:

- **Gap Persist, despite slightly improvements:** Despite overall advancements in educational coverage across Guatemala (as evidenced by an increase in average years of education for the general population from approximately **5.30** years in 2006 to **6.12** years in 2023, as shown in Table 1), the average educational gap between the indigenous and non-indigenous populations remained remarkably persistent between 2006 and 2023. Specifically, the total gap in years of education, according to our decomposition models, was **2.93 years for the Baseline Model and 3.39 years for the Extended Model** in 2006, and **2.71 years for the Baseline Model and 2.84 years for the Extended Model** in 2023 (see Table 2). This minimal change in magnitude over nearly two decades underscores that, while general educational indicators may improve, deep-seated structural inequalities continue to disproportionately affect the indigenous population, preventing a significant reduction in this disparity.

- **Evolution of Essential Causes:** Crucial changes in the composition of the gap have been observed over time and between models. Meanwhile, the baseline model that includes fundamental sociodemographic characteristics shows that a part of the gap is already explained by differences in observable characteristics (such as geographic location and age), It is in the extended model that the

doi:10.6342/NTU202504017

inclusion of parental education reveals a substantially higher explanatory power. In 2006, most of the explained portion by differences in characteristics was observable, particularity when the extended model is considered, that capture the effect of the parental education. However, in 2023, the unexplained proportion of the gap increased significantly in both models, that is more notorious in the baseline model and persist in the extended one. This suggest that more complex and often invisibles factors, such as potential discriminatory practices, a lack of cultural relevance in education, or difference in how educational achievement are valued in labor market are becoming more relevant to the persistence of this disparity.

- **The parental educational legacy:** the parental education had showed to be a consistent observable factor in the educational gap. This conclusion reinforced by comparing the results of the baseline model with those of the extended model, where the inclusion of the parental education substantially increases the explained portion. These emphasizes how the human capital that is transmittable at home and family environment are fundamental pillars for educational advancement. The initial disadvantages of previous generations are unfortunately replicated in their children's educational opportunities.

The findings highlight the urgent need for targeted educational policies that not only expand access, but also adapt the educational system to the specific needs of the population. In the case of Indigenous peoples, bilingual education should not simply consist of translating Spanish materials into local languages, but rather serve as a tool to preserve cultural heritage. This requires well-trained teachers who can properly support

doi:10.6342/NTU202504017

students in learning both their native languages and Guatemala's official language, Spanish. Equally important is the implementation of policies that ensure educational quality. As Jacir de Lovo emphasized, it is not enough to extend the years of schooling—it is also necessary to guarantee quality education at all levels in order to foster stronger human capital. Furthermore, our study underscores the critical importance of broader socioeconomic interventions, particularly policies aimed at reducing poverty, given its pervasive impact on access to and outcomes in education. Concurrently, strategies to enhance parental educational attainment are essential, as our analysis consistently demonstrates its significant role in explaining the educational gap and mitigating the intergenerational transmission of educational disadvantage.

doi:10.6342/NTU202504017

References

Alkire, S., & Santos, M. E. (2010). Acute multidimensional poverty: A new index for developing countries (Human Development Research Paper 2010/11).

Arriba González de Durana, A. (2002). El concepto de exclusión en política social (Documentos de trabajo, N.º 1). CSIC. Unidad de Políticas Comparadas.

Barro, R. J. (2000). Inequality and growth in a panel of countries. Journal of Economic Growth, 5(1), 5–32. https://doi.org/10.1023/A:1009850119329

Becker, G. S. (1975). Human capital: A theoretical and empirical analysis, with special reference to education (2a ed.). National Bureau of Economic Research.

Collins, P. H. (2019). Intersectionality as critical social theory. Duke University Press. https://doi.org/10.2307/j.ctv11hpkdj

Congreso de la República de Guatemala. (1991). Decreto Legislativo 12-91. Ley de Educación Nacional.

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. University of Chicago Legal Forum, 1989(1).

Cuervo Álvarez, D. B. (2016). La conquista y colonización española de América. Historia Digital, 16(28).

Feres, J. C., & Mancero, X. (2001). Enfoques para la medición de la pobreza: Breve revisión de la literatura.

Ferreira, F. G., Messina, J., Rigolini, J., López-Calva, L. F., & Vakis, R. (2013). La movilidad económica y el crecimiento de la clase media en América Latina.

Goldin, C. (2016). Human capital. In C. Diebolt & M. Haupert (Eds.), Handbook of Cliometrics. Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-642-40406-1_23

doi:10.6342/NTU202504017

Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. Journal of Economic Literature, 46(3), 607–668. http://www.aeaweb.org/articles.php?doi=10.1257/jel.46.3.607

Herrera Larios, A. C., Cifuentes Estrada, W. H., & Otzoy Chipix, B. L. (2023). Educación bilingüe intercultural: Estudio de los avances y desafíos en la educación superior estatal en Guatemala 2011–2020. Programa Universitario de Investigación en Educación y Estudios para la Paz, Escuela de Formación de Profesores de Enseñanza Media, EFPEM.

Holden, R. H. (2006). Review of The Last Colonial Massacre: Latin America in the Cold War, by G. Grandin. The Americas, 63(1), 143–145.

Instituto Centroamericano de Estudios Fiscales (ICEFI) & Save the Children. (2011). Educación bilingüe en Guatemala: Logros, desafíos y oportunidades.

Instituto Nacional de Estadística (INE). (2018). Resultados del censo 2018. INE. https://censo2018.ine.gob.gt/explorador

International Work Group for Indigenous Affairs (IWGIA). (2020). El mundo indígena 2020: Guatemala. https://iwgia.org/es/guatemala.html

Lee, C.-I., & Solon, G. (2009). Trends in intergenerational income mobility. The Review of Economics and Statistics, 91(4), 766–772.

López, W. E. (2024, June 19). El horizonte de la desigualdad educativa: Navegar por la brechas y desafíos en el acceso a la educación en Guatemala. Plaza Pública. https://www.plazapublica.com.gt/ensayo/articulo/el-horizonte-de-la-desigualdad-educativa-navegar-por-la-brechas-y-desafios-en-el

Lucas, R. E., Jr. (1988). On the mechanics of economic development. Journal of Monetary Economics, 22(1), 3–42.

doi:10.6342/NTU202504017

Martínez Peláez, S. (1998). La patria del criollo: Ensayo de interpretación de la realidad colonial guatemalteca.

Ministerio de Desarrollo Social, Rosales, S., Lemus, I., & García, E. (2018). Índice de pobreza multidimensional. Ministerio de Desarrollo Social.

Ministerio de Relaciones Exteriores. (s.f.). Información sobre Guatemala. https://www.guatemalaun.com/guatemala/

Mó Isém, R. (2019). *¿Cómo se implementa la educación bilingüe en Guatemala?* (Pueblos Indígenas y Educación Nro. 66). Ediciones Abya-Yala.

Montejo, V. (2022). Indigenous threatened heritage in Guatemala. In J. Cuno & T. G. Weiss (Eds.), Cultural heritage and mass atrocities.

Neef, D. (2024). Echoes of the past: The struggle to save Mayan languages in Guatemala.

Organismo Naleb'. (2013). Informe evaluación de la Declaración de las Naciones Unidas sobre los Derechos de los Pueblos Indígenas en Guatemala (E/C.19/2013/CRP.3). https://www.un.org/esa/socdev/unpfii/documents/2013/CRP-3.pdf

Psacharopoulos, G., & Patrinos, H. A. (2018). Returns to investment in education: A decennial review of the global literature (Policy Research Working Paper No. 8402). Education Global Practice, World Bank. https://documents.worldbank.org/en/publication/documents-reports/documentdetail/260321523264062931/

Ravallion, M. (1998). Poverty lines in theory and practice (Living Standards Measurement Study, Working Paper No. 133). World Bank.

Romero Leyva, F. A., Romero Castro, M. del R., & Fernández Velázquez, J. A. (Coords.). (2024). Educaciones y pedagogías interculturales para repensar la educación indígena. Astra Editorial.

doi:10.6342/NTU202504017

Schultz, T. W. (1961). Investment in human capital. The American Economic Review, 51(1), 1–17.

Smith, A. (2015). The wealth of nations (Original work published 1776). Cambridge University Press. https://doi.org/10.1017/CBO9781107338296

Strasser, H. (1994). Review of The Constant Flux: A Study of Class Mobility in Industrial Societies, by R. Erikson and J. H. Goldthorpe. European Sociological Review, 8(3).

UNESCO. (2024, November 5). Entornos de aprendizaje seguros: Prevención y tratamiento de la violencia en la escuela y sus alrededores. https://www.unesco.org/es/health-education/safe-learning-environments

UNIR. (2024, May 22). ¿Por qué se da la exclusión social y qué consecuencias tiene? Revista UNIR, Ciencias Sociales. https://www.unir.net/revista/ciencias-sociales/exclusion-social/

Velásquez, A. M. (2022). La desigualdad social en Guatemala: Evolución y respuesta institucional (LC/MEX/TS.2022/9). Comisión Económica para América Latina y el Caribe (CEPAL).

Wiessner, S. (2020). Rights and status of Indigenous Peoples: A global comparative and international legal analysis. Harvard Human Rights Journal, 12.

World Bank Group. (2015). Agricultura para la prosperidad de los territorios rurales en Guatemala: Vincular el desarrollo agropecuario con la prosperidad del campo (Informe No. Aus7583). World Bank Group.

World Bank Group. (2025, April 17). The World Bank in Guatemala. https://www.worldbank.org/en/country/guatemala/overview

Zimmerman, K. (1997). Modos de Interculturalidad en la Educacion Bilingüe. Reflexiones acerca del caso de Guatemala. *Revista Iberoamericana de Educacion* , 113-127.

doi:10.6342/NTU202504017