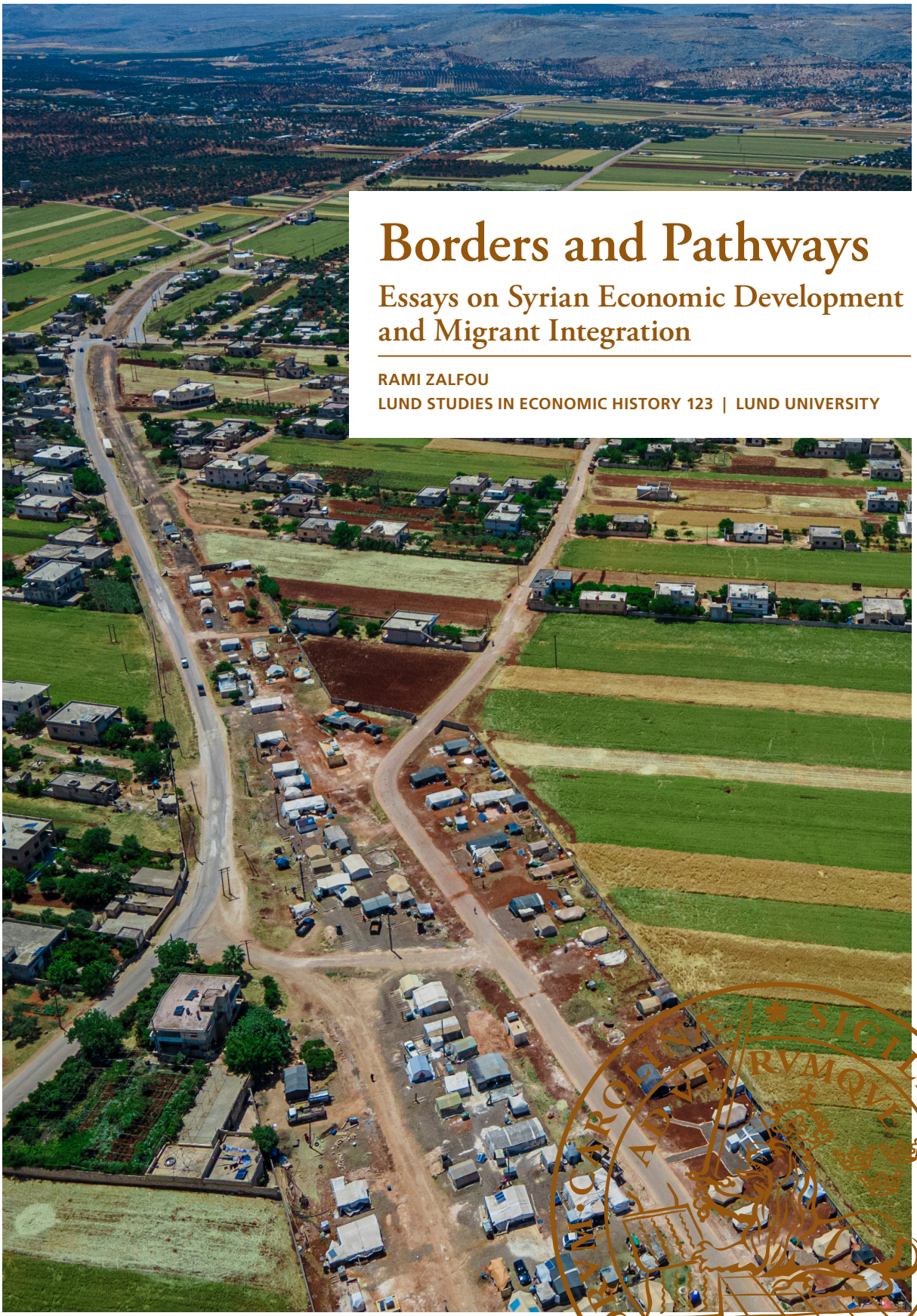


Borders and Pathways

Essays on Syrian Economic Development and Migrant Integration

RAMI ZALFOU

LUND STUDIES IN ECONOMIC HISTORY 123 | LUND UNIVERSITY

Borders and Pathways

Essays on Syrian Economic Development and Migrant Integration

Syria and the wider Middle East have witnessed dramatic demographic and economic change over the past century, but these transformations have unfolded alongside enduring inequalities. Regional divides, ethnic and religious cleavages, and gender disparities have proven remarkably persistent, shaping trajectories of development at home and influencing the experiences of millions abroad. From the Ottoman era to the refugee crises of the twenty-first century, the story of Syria is as much about fragmentation and inequality as it is about resilience and mobility.

This dissertation investigates these dynamics through four studies that connect history with the present and Syria with its diaspora. Two focus on Syria itself: the long-run consequences of statelessness for regional inequality, and the structural roots of ethnic and religious differences in women's labour force participation. The other two move outward, exploring how migrants from Syria and the broader Middle East integrated into new societies – Ottoman migrants and their children in early twentieth-century America, and MENA-origin youth navigating the transition to adulthood in contemporary Sweden.

The findings demonstrate the long shadow of history. Areas that lay beyond effective Ottoman state control remain disadvantaged in income, education, and infrastructure. Ethnic and religious gaps in female labour force participation are explained less by cultural norms than by uneven economic development, demographic structures, and access to public-sector jobs. In the diaspora, integration pathways diverge: while Ottoman-era migrants achieved rapid intergenerational mobility in the United States, today's youth of Middle Eastern origin in Sweden continue to face obstacles in achieving residential independence.

By bridging historical and contemporary perspectives, the thesis highlights how economic development, geography, and demography combine to produce both barriers and opportunities. It challenges explanations that rely solely on culture or identity, instead pointing to the interplay of structural conditions and historical legacies. Offering new empirical evidence from under-researched contexts, it contributes to broader debates in economic history, development, and migration studies, while also shedding light on Syria's enduring struggles and the diverse experiences of its people across borders.

Printed by Media-Tryck, Lund 2025 NORDIC SWAN ECOLABEL 3041 0903

Department of Economic History
School of Economics and Management

Lund Studies in Economic History 123
ISBN 978-91-989643-0-1
ISSN 1400-4860

Borders and Pathways

Borders and Pathways

Essays on Syrian Economic Development and Migrant Integration

Rami Zalfou

LUND
UNIVERSITY

DOCTORAL DISSERTATION

Doctoral dissertation for the degree of Doctor of Philosophy (PhD) at the School of Economics and Management at Lund University to be publicly defended on December 1st at 10.15 in building EC2 room 101, Department of Economic History, Lund.

Faculty opponent
Professor Mohamed Saleh

Organization: LUND UNIVERSITY

Document name: DOCTORAL DISSERTATION **Date of issue** 2025-10-30

Author(s): RAMI ZALFOU **Sponsoring organization:** NA

Title and subtitle: Border and Pathways: Essays on Syrian Economic Development and Migrant Integration

Abstract:

Over the past century, Syria and the broader Middle East and North Africa (MENA) region have undergone profound demographic and economic transformations, yet they remain marked by deep regional, ethnic, and gender-based inequalities. These divisions have shaped development trajectories and contributed to social fragmentation, while successive waves of migration carried these dynamics abroad, from Ottoman-era migration to the Americas to the mass displacement of the Syrian civil war. Despite their importance, the long-run roots of Syrian inequality and the integration of Syrian and MENA migrants remain underexplored in the empirical social sciences. This dissertation addresses these gaps by analysing the historical determinants of development within Syria and the economic and social integration of Syrian and Middle Eastern migrants in host societies.

The dissertation consists of four papers that combine economic history, demographic economics, and migration studies. Two focus on Syria, examining the long-term impact of historical statelessness on regional inequality and the drivers of ethnic and religious gaps in female labour force participation. The other two analyse migrant integration: the labour market outcomes of Ottoman migrants and their descendants in the United States during the early twentieth century, and the home-leaving behaviour of MENA-origin youth in contemporary Sweden. By linking historical and contemporary cases and applying methods ranging from spatial regression discontinuity and decomposition techniques to longitudinal analysis of census and register data, the dissertation provides novel quantitative evidence from contexts often overlooked in existing scholarship.

The findings from Syria show that historical legacies and structural inequalities have enduring effects on development and gender outcomes. Regions historically beyond Ottoman state control remain disadvantaged in income, education, and infrastructure, with settlement patterns transmitting these gaps over time. Ethnic and religious disparities in women's labour force participation are similarly rooted in structural conditions such as economic development, demographics, and access to public sector employment, rather than cultural norms alone.

In host societies, integration unfolds differently across contexts and generations. Ottoman migrants initially faced earnings penalties despite occupational prestige, but their children achieved rapid upward mobility through education and occupational advancement. In contemporary Sweden, however, MENA-origin youth display persistent gaps in social integration: they leave the parental home later than natives, with women especially likely to remain at home or exit through marriage, and neighbourhood context strongly shaping outcomes.

Key words: Syria, regional inequality, historical legacies, statelessness, gender inequality, migration, integration, social mobility, intergenerational outcomes.

Classification system and/or index terms (if any) **Number of pages:** 258

Language: English

ISSN and key title: 1400-4860 Lund Studies in Economic History 123

ISBN: 978-91-989643-0-1 (print) 978-91-989643-1-8 (PDF)

Borders and Pathways

Essays on Syrian Economic Development and Migrant Integration

Rami Zalfou

LUND
UNIVERSITY

Coverphoto by Ammar Alzeer / Shutterstock

School of Economics and Management
Department of Economic History

ISBN 978-91-989643-0-1 (print)
ISBN 978-91-989643-1-8 (PDF)
ISSN 1400-4860

Printed in Sweden by Media-Tryck, Lund University
Lund 2024

Media-Tryck is a Nordic Swan Ecolabel certified provider of printed material. Read more about our environmental work at www.mediatryck.lu.se
MADE IN SWEDEN

In memory of my father, Kamran

Table of Contents

Acknowledgments

This work would not have been possible without the generous support of many people. I am deeply grateful to the Centre for Advanced Middle Eastern Studies (CMES) at Lund University for funding this research as part of the strategic research area *MECW: The Middle East in the Contemporary World*. I particularly thank Karin Aggestam, director of CMES, for her encouragement and for giving me the freedom to pursue my research interests.

I owe special thanks to my supervisor, Martin Dribe, for his trust, invaluable advice, and steady support, and to my co-supervisor, Anna Tegunimataka, whose sharp eye and thoughtful critiques helped keep my work well-grounded.

My four years in the Department of Economic History passed quickly, thanks to the stimulating and diverse intellectual environment there, as well as the equally inspiring community at CMES. I am grateful to the professors, lecturers, and administrators in both institutions for their support throughout this journey.

I benefitted greatly from the Centre for Economic Demography environment. I am indebted to Jonas Helgertz for data provision; to Tommy Bengtsson for many insightful discussions; and to Björn Eriksson for feedback at my mid-way seminar. I want to also thank my final seminar discussants Jutta Bolt and Martin Nordin for reading a draft of the thesis and providing valuable comments and feedback, and Gunes Gokmen for feedback on the first two papers of the thesis. I also thank Ellen Hillbom for valuable guidance throughout the PhD. Many others offered helpful comments at seminars and conferences in Lund and elsewhere, and although too numerous to name, I am sincerely thankful to all of them.

I am especially grateful for the camaraderie of my PhD colleagues, in particular my cohort: Vinicius de Souza Maia, Liuyan Shi, Johanna Fink, John Brolin, Benjamin Chatterton, Anton Svensson, Jonathan Jayes, Philipp Jonas Kreutzer, and Youssouf Merouani. I would also like to thank Elien Dalman, a wonderful office mate, as well as Marcos Castillo and Louise Cormack for their kindness and always useful advice, as well as the many other PhD students whose company I valued throughout the last four years.

I owe the deepest gratitude to my wife, Anasuya Prabhu, my most devoted supporter. Her encouragement, willingness to accompany me across countries, and unwavering belief in me made this journey possible. I am also grateful to my sister, Ghalia, and my niece, Julia, whose summer visits to Sweden filled our home with warmth and joy. Above all, I am profoundly grateful to my mother, Reem AlMufti, whose gentle encouragement and unwavering support have been a constant source of strength.

List of Papers

I. The Historical State and Path Dependence: Evidence from Ottoman Syria's Stateless Frontier.
Unpublished manuscript.

II. Ethnic and Religious Differences in Female Labour Force Participation: Evidence from Syrian Census Data.
Published online in the Journal of Demographic Economics (2025), 1-33.

III. Premium or Penalty? Occupations and Earnings of Ottoman Immigrants and Their Offspring in the United States, 1900–1940. Co-authored with Martin Dribe.
Published in the European Review of Economic History (2025), 29(2).

IV. Home-Leaving Patterns Among MENA-Origin Youth in Sweden: The Influence of Origin, Generation, and Neighbourhood. Co-authored with Anna Tegunimataka.
Unpublished manuscript.

For the co-authored papers, I carried out the statistical analyses, while the framing and writing were developed jointly with the co-authors.

Introduction

1. Motivation and aim

Over the past century, the Middle East and North Africa (MENA) region has undergone profound economic and demographic transformations. Countries across the MENA embarked on industrialization at different moments in the 20th century, shaped by colonial legacies and natural resource endowments (Saleh 2021). These transformations were accompanied by rapid urbanization, shifting population structures, and uneven patterns of growth. Yet despite the scale of these changes, there remains limited systematic evidence on the deeper historical roots of development in the region, especially how economic transformations unfolded at the subnational level, and how they intersected with pre-existing ethnic and cultural divisions.

Syria represents a particularly compelling case for exploring these dynamics. As a historically diverse and regionally fragmented society, it has long faced difficulties integrating its various communities within a coherent national project. From its emergence as a modern state in the wake of Ottoman collapse, it struggled with state-building, uneven development, and social fragmentation (Hinnebusch 2004). These long-standing challenges were further magnified by the civil war that began in 2011, which exposed and deepened entrenched regional and sectarian inequalities. Although empirical research has begun to document patterns of regional economic inequality in Syria since before the war (El Laithy and Abu-Ismail 2005) and the relationship between ethnic cleavages and conflict (Mazur 2021), far less attention has been paid to the historical roots of these disparities. In particular, little is known about how the legacy of state formation – or its absence – shaped long-term economic trajectories, and how persistent socio-economic inequalities interacted with gender dynamics within Syria's complex social mosaic.

These domestic challenges did not remain confined within Syria's borders. They reverberate abroad through successive waves of migration, including the mass migration triggered by the Syrian civil war. Much of the scholarship on migrant integration has focused on turn of the century US and post-WWII Europe (Borjas 1994; Abramitzky and Boustan 2017), while migration from and within the MENA region received limited attention until the recent refugee crises. The 2015 displacement of millions of refugees brought new urgency to understanding how migrants from the MENA integrate into host societies – economically, socially, and

across generations. Significant gaps remain in our knowledge of their life-course transitions, intergenerational mobility, and the mechanisms of integration in receiving countries. Furthermore, studies of migrant integration that address the recent migrant flows from Syria and the broader MENA region typically focus on labour market outcomes, disregarding the equally important social and demographic integration. The recent migrant flow is also not the first of its kind, The early 20th century saw a large but often overlooked wave of emigration from Syria and the broader Ottoman Empire to the Americas (Naff 1993). This historical migrant flow provides important parallels to contemporary developments. And while it has been the subject of extensive qualitative and ethnographic work – particularly in cultural and diaspora studies (Naff 1993, Khater 2001, Gualtieri 2009) – it has received limited attention in the quantitative social sciences. Few studies have examined this migration using systematic data or sought to measure long-term integration outcomes across generations.

Syria thus provides a uniquely valuable case for addressing broader debates in economic history, development, and migration. Its experience as a historically fragmented frontier region speaks to global questions about how uneven state formation shapes long-run economic outcomes, offering rare subnational evidence on the persistence of regional inequality comparable to cases in sub-Saharan Africa, Latin America, and Central Asia. At the same time, Syria's combination of low female labour force participation with deep ethnic and religious diversity makes it a critical test case for understanding the structural and cultural drivers of gender inequality in low- and middle-income settings. Finally, Syria's migration history spans both the "age of mass migration" and the more recent refugee crises of the twenty-first century, allowing for direct comparison between Syrian and European migrant experiences in the Americas as well as insights into the challenges of refugee integration in contemporary Europe.

Taken together, these observations point to two significant gaps in the literature: the lack of evidence on the historical roots and consequences of regional economic inequality in Syria, and the limited understanding of the long-term integration of Syrian and broader MENA migrants. This dissertation addresses these gaps by combining historical and contemporary analysis across origin and destination contexts, drawing on new data and empirical methods to examine both regional development within Syria and the integration trajectories of Middle Eastern migrants abroad.

1.1. Scope of the study and contributions

This dissertation is a collection of four research papers that engage with themes of economic development and migrant integration. While each study explores a different dimension of these broad topics, they are linked by a shared focus on Syria's economic development and on migrant populations originating in Syria and

the broader MENA region. Across the four papers, recurring themes emerge around economic inequality – whether expressed through spatial disparities in development within Syria, inequalities between ethnic and religious communities, or differences in economic and social outcomes between migrant and native populations in host countries. Table 1 outlines the structure of the thesis and the four constituent papers including the main research question and data and analysis method used.

Table 1: Structure of the thesis

	Paper title	*Research question*	*Data and method*
I	The Historical State and Path Dependence Evidence from Ottoman Syria's Stateless Frontier	How does historical statelessness affect economic development in Syria today?	Geo-coded census data, Syria Regression discontinuity design
II	Ethnic and Religious Differences in Female Labour Force Participation Evidence from Syrian Census Data	What are the drivers of ethnic and religious differences in female labour force participation in Syria?	Geo-coded census data, Syria Gelbach decomposition
III *	Premium or Penalty? Occupations and Earnings of Ottoman Immigrants and Their Offspring in the United States, 1900–1940	How did Syrian and other Ottoman migrants and the second generation integrate into the US labour market?	Individual census data, US Panel regression
IV #	Home-Leaving Patterns Among MENA-Origin Youth in Sweden The Influence of Origin, Generation, and Neighbourhood	How do Middle Eastern-origin youth integrate into Swedish residential independence norms?	Population register data, Sweden Survival analysis

* co-authored with Martin Dribe
\# co-authored with Anna Tegunimataka

Paper 1 examines the long-term effects of historical statelessness in the Ottoman era on contemporary economic development in Syria. It contributes to two strands of research: the literature on impact of historical institutions on economic development (Acemoglu et al. 2001, Michalopoulos and Papaioannou 2013, Dell et al. 2018), and the literature on the role of path dependence in shaping long-run regional inequality (Krugman 1992, Bleakley and Lin 2012, Michaels and Rauch 2018, Dalgaard et al. 2022, Baerlocher et al. 2024). Previous studies have looked at the effect of historical state centralization on contemporary economic development outcomes (Michalopoulos and Papaioannou 2013, Dell et al. 2018) and linked these to the development of economic and political institutions. The paper takes this idea further

by analysing the effect of historical statelessness on economic development, looking at the impact of various mechanisms including population dynamics, economic institutions, and ethnic and cultural differences, in the Syrian context. The paper highlights the role of path dependence as a mechanism of persistence, resulting from the emigration of people from historically stateless areas, and the concentration of population in the areas that were under Ottoman state control in the centuries prior to the emergence of the modern state. The contribution is part of the broader debates on the persistence of historical institutions (Nunn 2020) and the role of state capacity in long-term economic growth (North et al. 2009). The paper also presents a novel view of the historical roots of regional economic inequality in Syria and contributes to our understanding of regional economic inequality in broader settings like the MENA region, and in agricultural-pastoral frontier regions more generally (Bai and Kung 2011, McGuirk and Nunn 2024).

Paper 2 studies the social consequences of uneven development by exploring the determinants of ethnic and religious differences in female labour force participation (FLFP) in Syria. It contributes to the literature on FLFP in developing countries, with particular relevance to Muslim-majority contexts, which continue to exhibit some of the lowest female participation rates globally (Ross 2008, Bayanpour-Tehrani and Sylwester 2013, Dildar 2015, Assaad 2014, Assaad et al. 2020). It also engages with broader research on the determinants of women's economic activity, including the influence of cultural norms, structural constraints, and development trajectories (Goldin 1995, Gaddis and Klassen 2014, Giuliano 2017). Although a handful of studies have explored religious differences in FLFP in Muslim-majority countries such as Turkey (Dildar 2015, Ugur 2018, Akyol and Ökten 2024), existing research remains geographically narrow and empirically limited, often due to the lack of data. By focusing on Syria, a country marked by significant ethnic and religious diversity and stark regional economic inequalities, this paper broadens the empirical scope of the literature and sheds new light on how macro-structural factors intersect with identity and shape gendered labour outcomes. The paper explores the factors associated with ethnic and religious gaps in FLFP using a rich set of variables including indicators of economic development and local labour market conditions, demographic structure, human and physical capital, economic institutions, gender norms, and geography. In doing so, the paper advances our understanding of how structural inequalities and regional development trajectories mediate gender inequality in the labour market in developing country contexts.

The next two papers move beyond Syria’s borders to examine issues of migrant integration in host societies. Paper 3 investigates the labour market integration of Ottoman-era migrants from Syria, Lebanon, and Turkey and their offspring in the US in the period 1900 to 1940. It examines the occupations and earnings trajectories of these individuals and compares their outcomes to those of US natives and other migrants groups. The paper contributes to the historical literature on migrant economic integration (Abramitzky et al. 2014, Abramitzky and Boustan 2017,

Connor 2020, Ward 2020, Escamilla-Guerrero et al. 2021), while addressing an important gap. Previous historical studies on migrant integration focus mostly on European immigrants in the US, while the economic integration of Middle Eastern migrants has not been studied using quantitative methods, even though the integration of Ottoman-era migrants in the US has received a lot of interest in qualitative and historical research (Naff 1985, Khater 2001, Gualtieri 2009). The paper provides new empirical evidence on the extent to which Middle Eastern migrants and their descendants were able to integrate economically, and whether their trajectories resembled or diverged from those of European migrants. The paper also explores the factors behind the successful assimilation of second-generation Ottoman immigrants, contributing to debates on the mechanisms behind migrant assimilation during the age of mass migration (e.g. Connor 2020, Ward 2020, Abramitzky et al. 2021).

Paper 4 turns to contemporary migration, analysing the social integration of Middle Eastern-origin youth in Sweden by looking at their transition to adulthood, and specifically at the transition out of the parental home and into independent living. The transition out of the parental home is considered a key life course transition in Western countries, and it reflects economic independence and social autonomy. Nordic countries like Sweden stand out in international comparisons for the early age at leaving the parental home (Dribe and Stanfors 2005). While some previous studies have looked at immigrant integration through the lens of this early life transition (Gillespie et al. 2020, Kleinepier and de Valk 2017, Zorlu and Mulder 2011), the literature remains limited in scope and focus. This study contributes to this research, and more broadly to the literature on migration social and demographic integration, by examining home-leaving behaviour among MENA-origin youth in Sweden, how it changes between the first- and second- generation and by age at arrival in the host country. The paper also examines the influence of individual and parental economic and contextual factors, as well as the influence of the local neighbourhood context.

Of the four papers, the first two papers focus on regional economic inequality in Syria and explore its historical roots and its impact on society, and the last two papers follow Syrian and other Middle Eastern migrants in the US and Sweden and analyse their economic and social integration in their host societies. By examining these facets, the dissertation sheds light on the roots, mechanisms and consequences of economic and social disparities in both origin and destination contexts. Overall, the thesis contributes to our understanding of social and economic inequalities in terms of region, ethnic and religious identity, gender, and migrant background.

The dissertation contributes to debates in economic history, development economics, and migration studies. It also introduces novel empirical evidence from contexts that remain underrepresented in these literatures, particularly the Syrian and broader Middle Eastern experience. In doing so, the thesis offers a corrective to the geographic concentration of much existing research on both economic

development and migrant integration, which tends to focus on Western settings or on a limited set of developing countries with available data.

The dissertation also makes methodological contributions by demonstrating the value of combining historical and contemporary data sources in the study of the MENA region, and by applying quantitative techniques to questions that are often addressed qualitatively in this context. This is particularly evident in the analysis of Ottoman-era migrants in the United States and in the examination of ethnic and religious diversity in Syria, where the use of disaggregated data and multivariate methods allows for new insights into longstanding debates.

1.2. Limitations

While the dissertation makes several novel contributions to the study of economic development and migrant integration, it is important to acknowledge the limitations inherent in the approach, particularly those related to data availability, scope of generalizability, and methodological constraints.

First, the empirical analysis relies heavily on observational data, both historical and contemporary, which limits the ability to make strong causal claims. While efforts have been made to address endogeneity - through the use of quasi-experimental variation in paper 1, and rich sets of control variables and robustness checks in the all the papers - there remains a risk of omitted variables and unobserved confounding factors, biasing the results.

Second, data availability and its limitations constrain the scope and focus of the analysis. Data is particularly sparse in the MENA context, which means that the analyses in papers 1 and 2 ultimately rely on cross-sectional data without variation across time or across countries to validate the results. In terms of papers 3 and 4, the analyses are again limited to one country in each study, ignoring the MENA migrant flows to other countries in Europe or the Americas, where the experiences of migrant integration may differ. This raises questions about the external validity of the analyses and their applicability in other contexts. Further research is needed on regional and ethnic inequality in other MENA countries, and on MENA migrant integration in other receiving countries to understand the commonalities and differences across contexts.

Third, the dissertation engages with topics that touch on identity and gender, but it is ultimately grounded in the methods of empirical economic history and economics. For this reason, a certain degree of simplification was necessary in the treatment of complex social and cultural processes. Concepts such as ethnic identity, religious affiliation, gender norms, and integration are deeply embedded in historical, political, and cultural contexts, and cannot be fully captured through quantitative proxies alone.

Despite these limitations, the thesis provides robust empirical findings and opens new avenues for future research to build upon and refine its findings.

2. Economic and historical context

The dissertation focuses on regional economic inequality in Syria, and on the economic and social integration of Syrian and other MENA origin migrants in host countries. This section provides a background on the modern economic history of Syria, and the migration flows from Syria and the broader MENA region.

Throughout its recorded history, the territory of present-day Syria has rarely existed as an independent political entity, instead forming part of a succession of larger imperial structures. Syria was integrated into such a system under Ottoman rule from the early 16th century until the empire's dissolution at the end of World War I. It was then incorporated into the 'French Mandate for Syria and the Lebanon' until achieving independence in 1946. During the four centuries of Ottoman governance, the region was divided into provinces that were only loosely managed by the central authority, with significant autonomy emerging in peripheral areas – especially those inhabited by ethnic and religious minorities (Ma'oz 2013).

Syria – like much of the MENA region – occupies a historically paradoxical position in the global economic landscape. Once a vital centre of commerce and scientific advancement, its cities played a central role in the pre-modern world economy. However, beginning in the early modern period and accelerating with the advent of industrialization, global economic power gradually shifted westward. As Western Europe industrialized and expanded its colonial and commercial reach, the MENA region, including Syria, experienced a relative economic decline. What had once been a core region of global trade and innovation became increasingly marginalized in the emerging global capitalist order, transformed from centre to periphery.

During the centuries of Ottoman rule, Syria experienced relative economic stagnation. While the empire maintained long-distance trade routes and urban centres in Syria like Damascus and Aleppo retained some regional importance, the economic structures remained largely agrarian and extractive, with limited investment in infrastructure, education, or industrial development. Local economies were shaped by a system of tax farming and administrative decentralization, which often hindered the formation of cohesive markets or state-led development. The decline of caravan trade and the rise of maritime commerce further marginalized Syria's position, while the recurrence of the plague kept the population stationary or declining (Owen 1981).

The historiography of Syria highlights rural decline in the late Ottoman period, linked to the empire's weakening ability to secure the countryside, particularly in

the face of tribal migrations from the Arabian Peninsula (Owen 1981, p. 6). The erosion of central authority was also evident in fiscal administration. The Ottoman government increasingly relied on powerful local families to manage taxation, most notably through the *malikane* tax-farming system. Families such as the Azms parlayed their wealth into political influence, purchasing support in Istanbul to secure appointments like the governorship of Damascus. These elites dominated both urban administration and landed wealth, extracting surplus from the rural economy through control of cereal, barley, and sheep production. Beyond them, independent chieftains maintained authority in peripheral regions – the Alawi mountains near the coast, the Druze Mountain in the south, and the tribal domains of the Syrian steppe in the east – limiting the effective reach of the Ottoman state across the regions of Syria.

From the mid-19th century, however, the Ottomans launched a series of centralizing reforms (the Tanzimat and subsequent provincial reorganizations) aimed at curbing such local autonomy, strengthening taxation, and projecting authority more directly into Syria. While important in reshaping state–society relations, these reforms were uneven in their impact and left many underlying tensions unresolved. The empire's collapse after World War I brought Syria under the French Mandate (1920–1946). French administrators built on Ottoman legacies while introducing new legal codes, roads, and schools, and they sought to impose central authority by cultivating loyalist elites. At the same time, they reinforced sectarian divisions by granting minorities like the Alawis and Druze separate institutions. The result was a paradox: the Mandate extended certain aspects of modern statehood, but primarily to serve imperial priorities rather than genuine developmental goals.

It was only after the end of the Mandate and the achievement of independence in 1946 that Syria began to pursue its own developmental path. The post-independence state adopted a more interventionist role in the economy, launching land reforms, nationalizing key industries, and experimenting with import-substitution industrialization. Though constrained by political instability and regional conflict, these policies marked Syria's first sustained effort to integrate into the global economy as a modern nation-state.

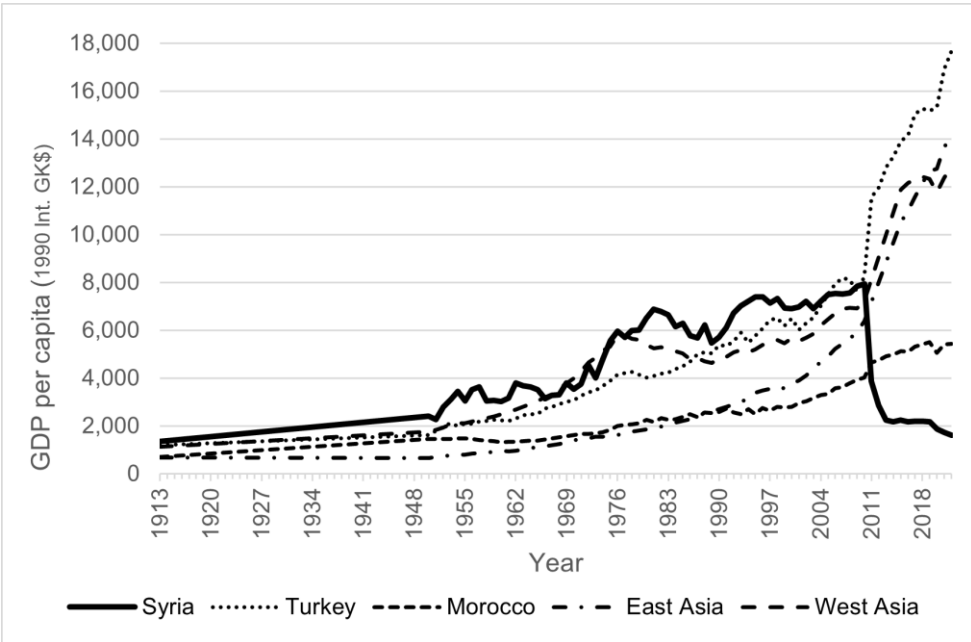
2.1. Economic development and inequality

Figure 1 traces Syria's per capita GDP between 1913 and 2022, measured in 1990 international dollars and compared with selected countries and regions[1]. Economic growth was modest before 1940, with per capita GDP rising by only about $1,000 during the first half of the 20th century. From independence in 1946 until 1970, growth remained sluggish, hampered by chronic political instability – including

[1] The comparator countries and regions used in the following figures were chosen pragmatically. The comparators vary by indicator, depending on data availability and relevance.

repeated military coups – and the pursuit of nationalization and import-substitution policies. A turning point came in 1970 with the rise of the Assad regime, which established political stability, adopted a more pragmatic stance in economic and foreign policy (Borazan 2021), and oversaw rapid growth through the 1970s, boosted by rising oil prices. This momentum unravelled in the 1980s as the economy contracted sharply. After 1990, further liberalization spurred substantial but volatile growth until 2010. The outbreak of the Syrian civil war in 2011 triggered a severe collapse in living standards: by 2015, per capita GDP had fallen back to its 1950 level of around $2,000, and it continued to decline as the conflict stretched into the 2020s[2].

Figure 1: GDP per capita in Syria and selected comparators 1931 - 2021



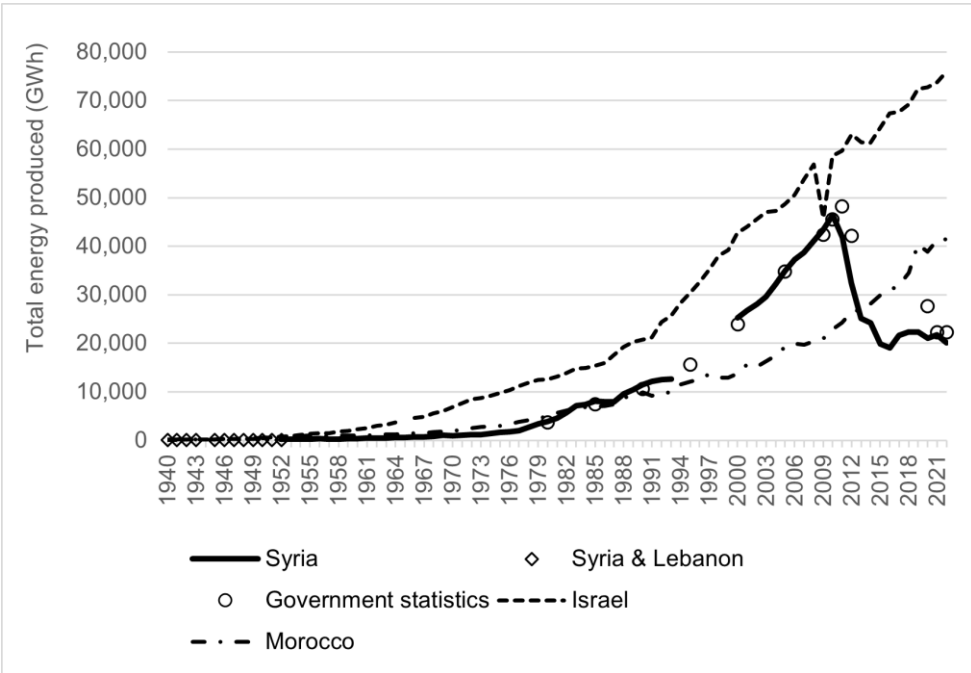
Source: Before 2011: Bolt and van Zanden (2014); 2011 – 2022; Bolt and van Zanden (2025).

[2] It is important to note that estimates of the level and trend in per capita GDP differ across sources. Later revisions of the Maddison database give Syria a lower level of per capita GDP while revising upwards the levels in comparators like Turkey and East Asia. Those estimates still put Syria in Middle Income territory. However, the 2020 and 2023 revisions show a peak in GDP per capita in 1980 and gradual decline thereafter, which is implausible and contradicts other sources such as Pamuk (2006). The 2013 version of the Maddison Database (Bolt and van Zanden 2014) is also used in Saleh (2021). Similarly, estimates of the fall of GDP per capita following the start of the civil war vary between studies (Kešeljević and Spruk 2024).

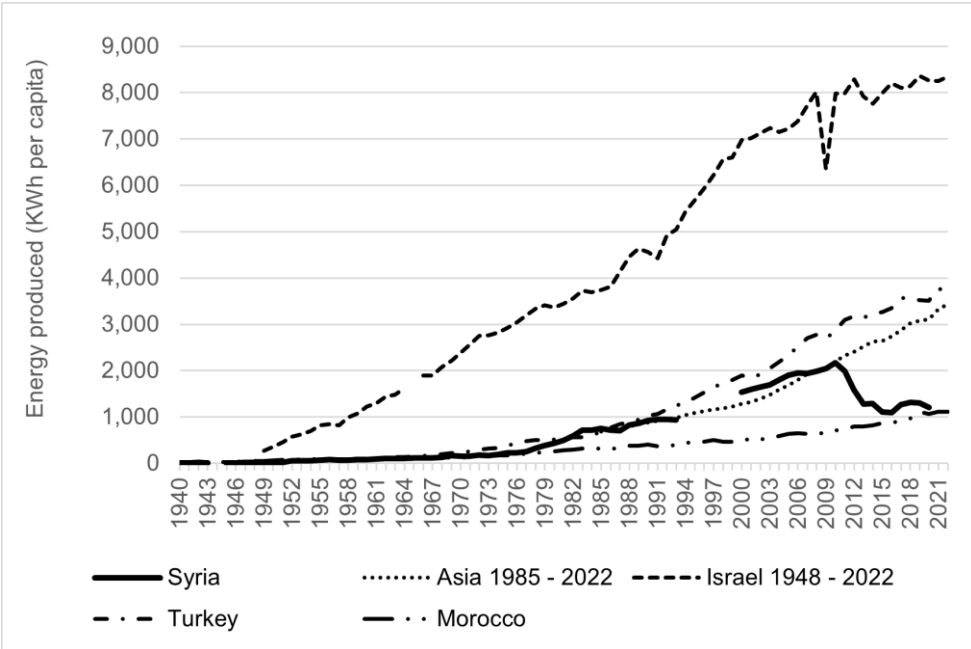
Another way to proxy for economic development is by looking at energy production, which can be seen in figure 2. Panel A shows the evolution of total electrical energy produced in Syria, with Israel and Morocco as comparators. It illustrates how Syrian energy production began to expand significantly after 1977, and it grew in parallel with energy production in Israel, before experiencing a collapse following the breakout of the civil war in 2011. Despite the favourable gross production picture for the period 1977-2011, panel B of figure 2 shows that per-capita energy production was much slower in Syria as compared to Israel. The former followed the growth trajectory of Asia in the latter part of the 20th century, and grew in step with neighbouring Turkey, while Israel's per capita production expanded much more rapidly to reach developed-country levels.

Figure 2: Energy production in Syria and selected comparators

A. Total energy produced 1940 - 2022



B. Per capita energy produced 1940 - 2022

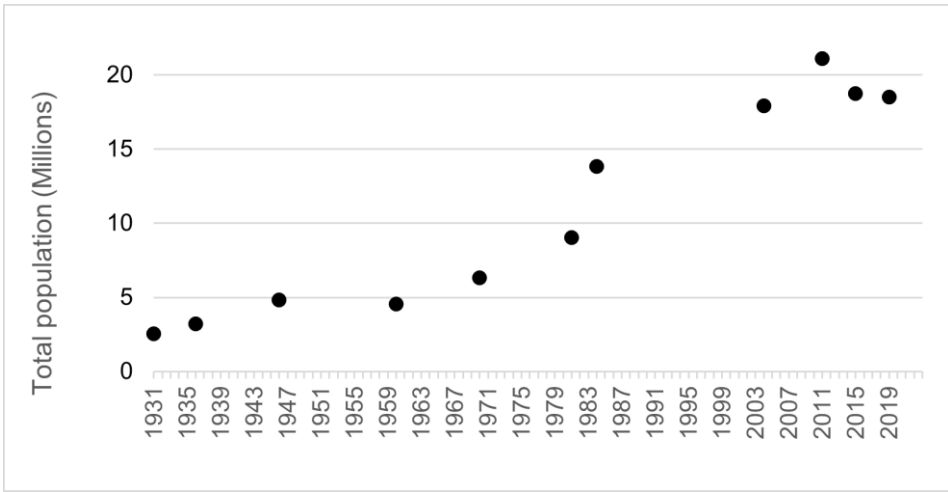


Source: Data until 1993 from B.R. Mitchell (1998). After 1993: IEA World Energy Statistics and Balances.

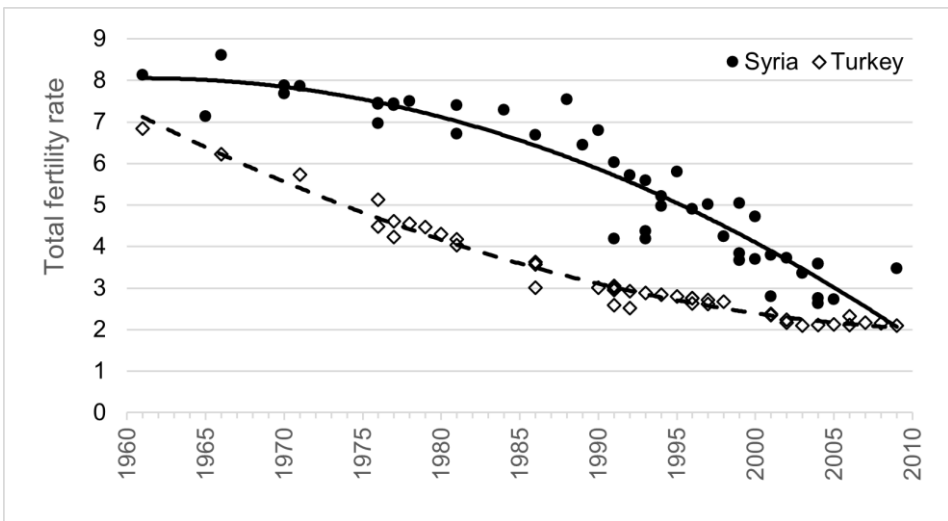
These trends must be understood in the context of Syria's rapid population growth throughout the 20th century. The population grew from about 2.5 million in 1931 to a peak of 21.1 million in 2011, before declining due to large-scale emigration caused by the civil war (Figure 3.A). This represents an eightfold increase in less than a century, with the most rapid growth occurring during the 1970s and 1980s – a period of demographic take-off. Figure 3.B presents UN estimates of fertility rates in Syria from 1960 to 2010, using Turkey as a comparator. While fertility began to decline in both countries in the mid-20th century, Syria's transition was considerably slower, declining from a TFR of 8 in 1960 to a TFR of 7 in 1990, with persistently high fertility rates throughout the 1970s, 1980s, and 1990s. The decline became much more rapid after 1990, In the post-independence era, Syrian government policies remained largely pro-natalist until the late 1980s, and even after these policies were reversed, authorities were hesitant to implement measures to curb population growth (Courbage 1994).

Figure 3: Population and fertility

A. Total population in Syria 1931 - 2022



B. Total fertility rate in Syria and Turkey 1960 - 2010

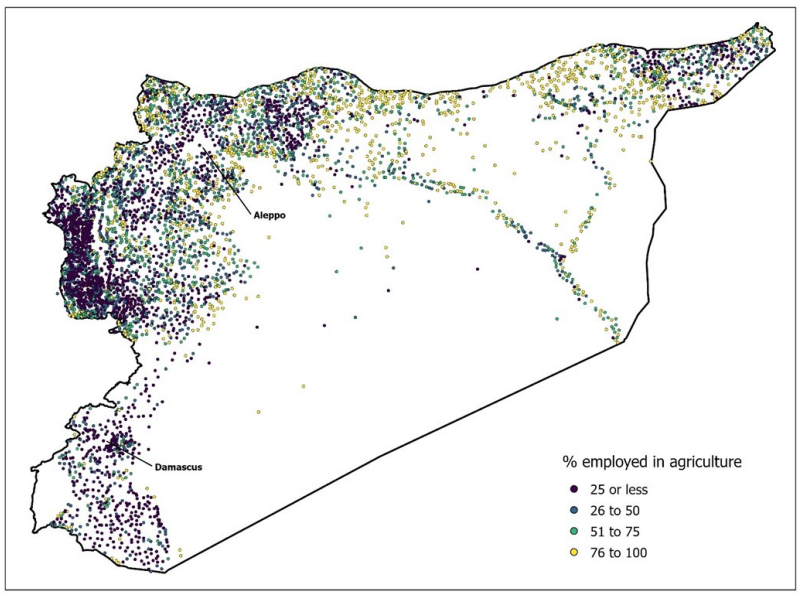


Source: United Nations (2019)

Country-level economic figures hide significant within-country differences. Studies undertaken on poverty in Syria highlight its uneven geographic spread, with the highest poverty rates found in the east of the country (El Laithy and Abu-Ismail 2005). These differences have deep historical roots, as the eastern region has been for centuries at the margins of centralized state authority. Under both Ottoman and

prior dynasties, much of this region was considered part of the "tribal zone" – a frontier sparsely populated by Bedouin tribes and largely beyond the reach of direct state control. Unlike the more densely populated western regions, such as the Orontes valley and the Damascus basin. As a result, successive states struggled to project authority over these lands, relying instead on negotiated relationships with tribal leaders, or at times, neglecting the region entirely. As a result, Syria is today highly geographically unequal, as can be seen in figure 4 which shows a map of the share of the workforce employed in agriculture. This share is lowest in the coastal region, the Damascus basin area, and increases towards the east of the country, reaching above 75% in many areas. The national-level share of the workforce employed in agriculture stood at 15% in 2004, down from around 70% in 1955 (International Bank for Reconstruction and Development, 1955).While the share employed in agriculture is not a perfect or complete measure of economic development, it provides a useful proxy for understanding spatial patterns. These pattern mirrors broader patterns of development in human capital and infrastructure across the geography of the country.

Figure 4: Share of the workforce employed in agriculture in Syria in 2004



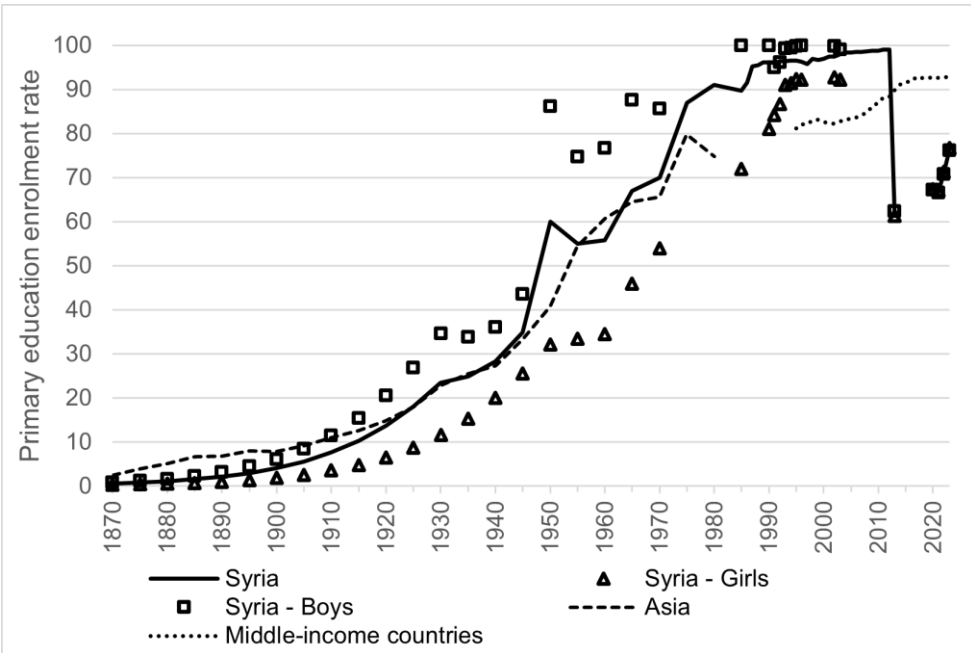
Source: The Syria population and housing census of 2004.

These geographic disparities are not the only axis along which economic development is uneven. Gender-based inequalities compound these regional divides, with girls – particularly in poorer and more remote areas – facing barriers to schooling. To understand the long-term trends in gender inequality in schooling,

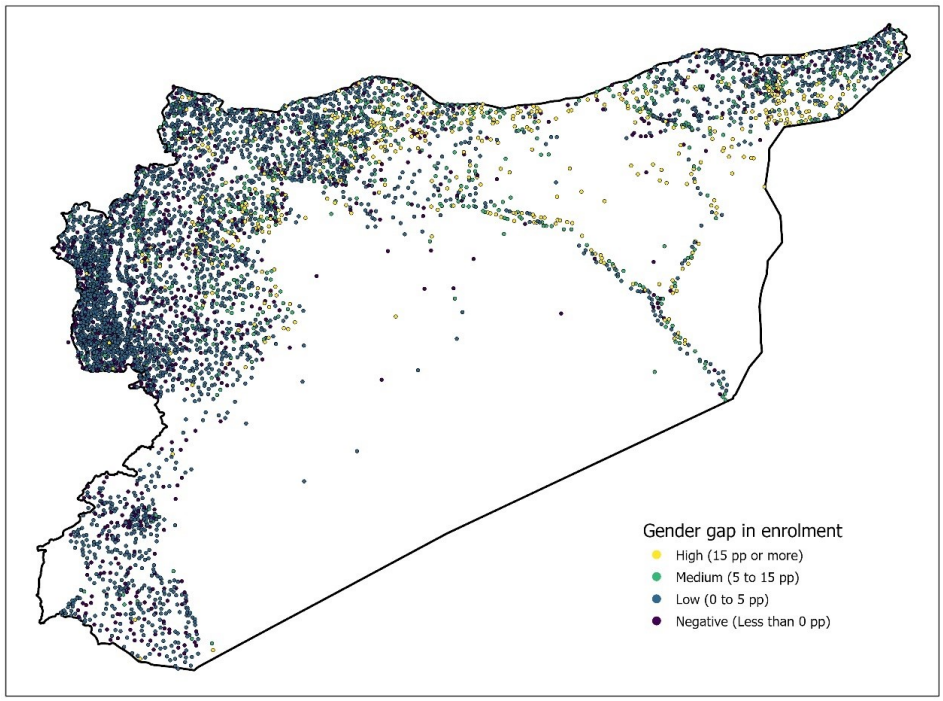
figure 5 shows male and female school enrolment rates in Syria across the 20th century. Panel A illustrates how primary enrolment starts growing in the late 19th century, with the gap between boys and girls growing with the growth of enrolment rates. Enrolment growth speeds up in the post-independence era, as does the enrolment gender gap, before rates converge towards the end of the 20th century. After that, we see the collapse in enrolment rates following the start of the civil war, which affects both genders. Looking at the geography of the gender gap in primary school enrolment using data from the 2004 census in panel B, we see that the poorer regions in the north and east of the country are the ones with the highest gender gaps in enrolment, reaching over 15 percentage points in some areas, despite the near convergence at the national level. A similar spatial pattern to that of overall regional economic inequality across the country.

Figure 5: Primary school enrolment rates in Syria

A. Primary school enrolment 1870 - 2020



B. Gender gap in primary school enrolment across geography in 2004 (ages 12 to 14)



Source: A: Ritchie et al. (2023); B: The Syrian population and housing census of 2004.

Scholarly work on the Syrian conflict connects deeply with the issue of inequality. Since the onset of the Syrian civil war, scholars have examined a wide range of factors behind its outbreak. Early research emphasized climactic factors like drought and its associated economic pressures, but later studies were critical of this theory (Angermayr et al. 2023). Barout (2022) instead underscores deep regional inequalities: the marginalization of Syria's smaller cities and their rural hinterlands from economic growth, mounting demographic pressures from a large youth population facing high unemployment, the erosion of subsidies, and the absence of democratic reforms capable of channelling popular grievances. Mazur (2021) shifts attention to the escalation of protests into an ethnic conflict, attributing this to entrenched ethnic and sectarian inequalities under the Assad regime. While the regime had long relied on trans-ethnic patronage networks to secure loyalty, these were undermined by the liberalizing reforms of the 1990s and 2000s and ultimately unravelled as the uprising unfolded. Haddad (2020) stresses the damaging consequences of Syria's economic liberalization, including increased inequality, weakening of the social contract, and pervasive corruption and rent-seeking. The

civil war itself has had severe consequences for Syria's economy, and the sanctions imposed on the Syrian government prolonged and protracted the negative effects of the war (Kešeljević and Spruk 2024). The post-Assad era, beginning in late 2024, has raised cautious hopes for recovery, particularly following the lifting of international sanctions. Yet deep ethnic polarization remains a central challenge, with Syria fragmented into de facto zones of control along ethnic lines. Sporadic outbreaks of intercommunal violence continue to threaten the fragile stability, raising fears of a relapse into civil war.

2.2. Migration

The Syrian civil war triggered one of the largest displacement crises in recent history. More than six million people became internally displaced, while more than five million left the country. While unprecedented in scale for the MENA region, this mass migration followed a longer historical pattern of both domestic and international population movements in and out of Syria.

Internal migration within Syria had already been rising in the late 20th century, driven by industrialization and uneven economic development that concentrated opportunities in urban centres. The civil war intensified this trend. For example, Damascus saw its population rise from about 194,000 in 1935 to 2.4 million in 2010, and further to 2.68 million by 2024 – nearly doubling the national rate of population growth despite having a lower fertility rate than other regions.

Syria has also long been a destination for refugees. In the second half of the 20th century, it hosted successive waves of Palestinian refugees (in 1948, 1967, 1970, and 1982) and, more recently, Iraqi refugees after 2003. By 2009, Syria was home to an estimated 1.33 million refugees (Mehchy and Doko 2011). This tradition of refugee reception dates back even further to the late Ottoman period, when Syria absorbed displaced populations from the Balkans, the Caucasus, and, during World War I, Armenian refugees from Anatolia.

At the same time, Syria has a long history of emigration. The first major wave occurred during the late Ottoman era, as part of the broader migration flow during the age of mass migration. Syrians, along with Lebanese and other Ottoman citizens, joined the flow of European emigrants heading to the Americas. This migration was not evenly distributed; it was particularly concentrated among Christian and Druze communities and in urban centres like Aleppo and Damascus. In Mount Lebanon, roughly a quarter of the population is estimated to have emigrated during this period.

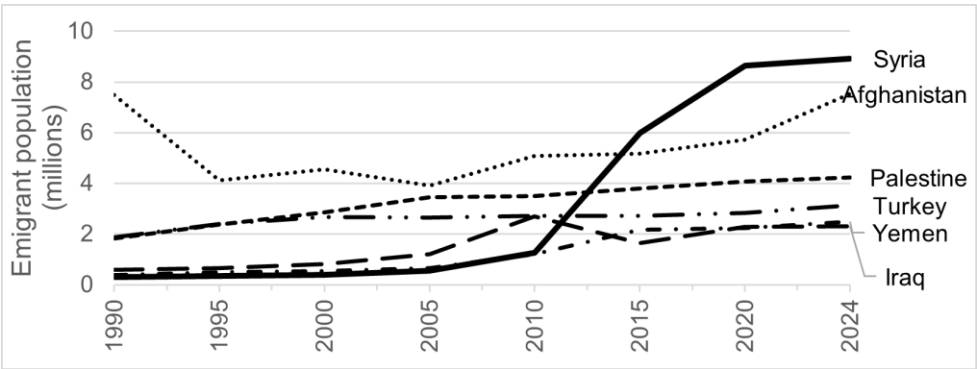
A second significant wave of emigration followed Syria's independence, driven by labour demand in the rapidly expanding oil economies of the Gulf. Countries like Saudi Arabia, Kuwait, and the United Arab Emirates became major destinations for both low-skilled and high-skilled Syrian workers, creating a circular migration

system that deeply shaped Syria's labour market and economy. Meanwhile, smaller but continuous flows of Syrian migrants also headed to the Americas and Europe.

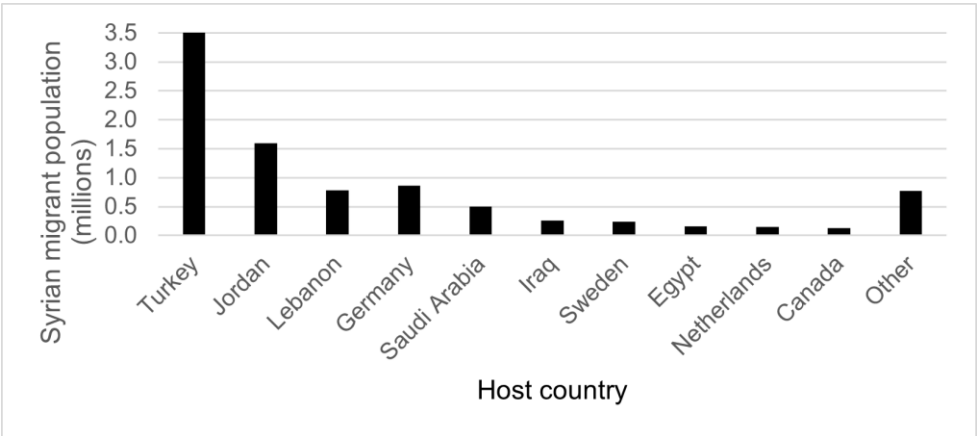
Figure 6.A displays the emigrant stock by country of origin for selected MENA countries. For Syria, the emigrant population grew from nearly 300,000 in 1990 to 550,000 by 2005. By 2010, it had more than doubled to approximately 1.26 million. Following the outbreak of conflict in 2011, the number surged dramatically - reaching nearly 6 million in 2015 and approaching 9 million by 2024. As shown in Figure 6.B, the vast majority of Syrian emigrants in 2024 resided in neighbouring countries such as Turkey, Jordan, and Lebanon, as well as European countries including Germany, the Netherlands, and Sweden.

Figure 6: Size of emigrant populations

A. Emigrant population by country of origin 1990 - 2024



B. Syrian emigrant population by host country in 2024



Source: United Nations Department of Economic and Social Affairs, Population Division (2024).

3. Theory and previous research

This thesis is structured around three central themes. First, it examines how historical state presence affects long-term economic development. Second, it explores the determinants of spatial and ethnic disparities in women's labour force participation. Third, it analyses the economic and social integration of immigrants in their host countries. In addition to these core areas, the individual papers also engage with a range of ideas within each area.

3.1. Historical institutions, path dependence, and economic development

Economic development is shaped by a diverse and interwoven set of factors, ranging from geography and natural resources to demography, culture, institutions, technology, trade networks, and geopolitics. Recent work has focused on the role of historical state institutions, and the way that they affect economic development over the long run. In their seminal work, Acemoglu et al. (2001) show that colonial institutions determine the long-run growth trajectories of the places where Europeans settled. In places where the geography was welcoming to the new inhabitants, Europeans migrated in great numbers and established inclusive institutions to govern their lives, including private property institutions and checks and balances against government expropriation. Places that were less welcoming and had valuable resources, received less European migration, and those that went there established 'extractive' institutions designed to exploit and expropriate the labour and resources of native populations. These institutions persisted in one way or another into the present day and have shaped and continue to shape economic development across the globe.

This theory has proved highly influential and has inspired a growing literature on the effect of historical institutions in various settings and their mechanisms. Several studies have explored the role of historical state centralization in explaining economic development in the present (Gennaioli and Rainer 2007, Michalopoulos and Papaioannou 2013, Dell et al. 2018), which all show that the presence of more centralized states in the past lead to better economic development outcomes today, whether measured through human and physical capital and the provision of public goods (Gennaioli and Rainer 2007), nightlight density (Michalopoulos and Papaioannou 2013, or household consumption (Dell et al. 2018). The theorized mechanism in this literature is largely institutional, in the form of stronger and more inclusive local governance institutions that determine long-run development outcomes.

At the same time, scholars have pointed to the importance of path dependence in explaining long-run economic trajectories. Several studies show that initial location

advantages can generate self-reinforcing agglomeration effects, which contribute to explaining long-term regional development patterns (e.g. Bleakley and Lin 2012, Baerlocher et al. 2024). Oto-Peralías (2020) demonstrates that the frontier zone of medieval Spain, once marked by chronic insecurity and warfare, continue to suffer from the effects of this historical locational disadvantage, with long-term consequences for patterns of population settlement and for economic development. By affecting population settlement patterns, locational advantages can shape long term development trajectories through the agglomeration of people and economic activity. A theoretical and empirical literature explores this idea by analysing the effect of population agglomeration on economic growth. Boserup (1965, 1981) argued that the choice of technology used in agriculture is affected by population density, and several studies have followed this line of thinking, theorizing a link between population size and population density on the one hand, and long-run technological growth on the other (Kremer 1993, Klasen and Nestmann 2006).

The institutional explanation for long-term economic development is often rooted in geography and path dependence as well. According to Acemoglu et al. (2001), the disease environment in African countries gave rise to extractive institutions, while Engerman and Sokoloff (2000) also theorize that initial factor endowments like climate, soil, and native population density in the new world gave rise to different types of institutions, which in turn determined long-term economic development paths. Therefore, initial conditions may directly affect economic development in the long term through agglomeration effects, or indirectly through their effect on institutions.

Geography has played a major role in shaping historical state institutions in different ways. What constitutes a “frontier zone” depends on the historical context. In the Spanish frontier studied by Oto-Peralías (2020), for instance, a natural defensive barrier – a river – combined with the region's location between two warring sides to create conditions of chronic warfare and insecurity. In the Middle East, geography similarly influenced state formation: centralized states developed around river systems, while stateless societies persisted in steppes, deserts, mountains, and other marginal areas. The geographic diversity of the MENA region not only fostered cultural variation but also led to significant differences in the presence, strength, and reach of state institutions. Syria offers a clear example. Unlike its historically centralized neighbours, Egypt and Mesopotamia, Syria rarely produced a unified territorial state due to its fragmented geography (Hinnebusch 2021). Even when large empires like the Ottomans ruled Syria, they struggled to exert control over its diverse terrain. Central authority, even in the most secure areas, continually faced resistance from stateless peoples and tribal and clan societies along the peripheries. This geographic fragmentation continues to shape modern patterns of economic development and remains a major challenge for nation-building efforts today.

The case of Syria thus illustrates how geographic fragmentation can produce long-lasting institutional weaknesses, with significant consequences for economic

development. In contrast to regions where centralized states emerged early and endured, Syria's fragmented geography fostered persistent political instability, weaker state capacity, and a patchwork of semi-autonomous local authorities. Over time, this institutional fragmentation shaped patterns of population settlement, economic organization, and infrastructure development in ways that continue to influence regional disparities today. Just as frontier regions in Spain experienced enduring disadvantages due to chronic insecurity, Syria's fragmented historical state structure has contributed to uneven development and limited the emergence of strong, centralized governance.

More broadly, Syria highlights how geography, institutions, and path dependence interact to shape long-run economic outcomes. Initial geographic conditions constrained the formation of strong states, leading to institutional weaknesses that became self-reinforcing over time. These dynamics mirror the mechanisms identified in the broader literature, where initial conditions – whether environmental, demographic, or geographic – shape the institutions and agglomeration patterns that ultimately determine economic development trajectories. The Syrian case thus underscores the importance of considering both geography and path dependence when analysing long-term effect of historical institutions on economic development.

3.2. Economic development and female labour force participation

A large and growing body of research examines the causes and consequences of gender inequality, with particular emphasis on FLFP as a key indicator of women's position in the economy and society more broadly. The onset of the modern era, especially with the rise of industrialization, marked a profound transformation in gender relations within economic life. Over the course of the twentieth century, Western societies experienced significant convergence between men and women in various dimensions of economic activity, including labour force participation, occupational attainment, and earnings (Goldin 2014). This shift has spurred a rich historical literature aimed at uncovering the drivers of these changes, highlighting factors such as education, fertility decline, changing social norms, technological innovation, and institutional reforms (Goldin 1995, Olivetti and Petrongolo 2016; Stanfors and Goldscheider 2017; Mancini 2018). In parallel, an extensive international literature has emerged that systematically compares patterns of female labour force participation across different countries and regions, seeking to understand the role of economic development, cultural factors, policy frameworks, and demographic transitions in shaping cross-national variation (e.g., Klasen 2019). Together, these literatures underscore the complexity and multifaceted nature of the processes underpinning changes in gender inequality, both within and across societies.

The Middle East and North Africa (MENA) region consistently ranks among the lowest in the world in terms of female labour force participation (FLFP), a pattern that has drawn sustained attention in international comparisons. This underrepresentation of women in the labour force poses a significant challenge for policymakers, as excluding half of the working-age population from economic activity hampers equity, productivity, and long-term development. Scholars focusing on the MENA region have sought to explain this persistent gender gap, proposing a range of explanations. Among the most prominent is the cultural argument: deep-seated social norms and traditional gender roles are believed to constrain women's participation in paid work. This perspective has been widely explored in cross-country analyses (e.g., Alesina et al. 2013), and it has also been the focus of numerous country-specific studies within the region, particularly in Turkey where individual-level data is available (Akyol and Ökten 2024; Atasoy 2017; Dildar 2015). These studies suggest that social and cultural attitudes toward gender, shaped by religion and religiosity, as well as family structures and historical legacies, play a central role in shaping women's economic opportunities.

Other studies have gone beyond the cultural explanation, emphasizing the importance of MENA-specific structural economic factors that contribute to the suppression of FLFP in the region. One such factor is the region's heavy reliance on oil revenues, which Ross (2008) argues reduces women's economic participation by generating rents that lessen the state's need to tax citizens and thereby reduce political pressure for gender-inclusive policies. In addition, the region has witnessed a gradual shrinking of the public sector - a sector that has historically provided relatively secure and socially acceptable employment opportunities for women, thereby limiting accessible employment channels for female workers (Assaad et al. 2020). Furthermore, the MENA labour market is characterized by pronounced dualism, with a clear divide between a limited formal sector, and a large informal one. Formal sector jobs, which tend to offer better pay and working conditions, are scarce, and women often face additional barriers to accessing them due to legal restrictions, social expectations, and hiring biases (Assaad 2014). Together, these structural constraints not only limit the availability of good-quality jobs but may also interact with prevailing gender norms to reinforce women's exclusion from the workforce. As İlkkaracan (2012) puts it, this results in the absence of a demand-side challenge to the gendered division of labour, meaning that the economic structure does not generate sufficient incentives or pressures to shift traditional gender roles or encourage greater female participation in paid employment.

The interaction between structural economic constraints and gender norms is not uniform across the population. Women from ethnic and religious minorities may experience these dynamics more acutely, as their labour market prospects are shaped not only by gender but also by the relative status of their group. Discrimination, regional inequalities, and disparities in access to public services and education can amplify the effects of economic dualism and public sector contraction, leading to

differences in FLFP rates between ethnic and religious groups. Differences in demographic structure can also contribute. The study of these gaps is further exacerbated by the presence of non-linearities, as suggested by the U-shaped relationship between economic development and FLFP (Boserup 1970, Goldin 1995). In early stages of development, economic necessity may drive women into low-wage informal work, while at middle-income levels, restrictive social norms and rising male wages can reduce FLFP. It is only at higher levels of development – with increased access to education, delayed marriage, and growth in the service sector – that FLFP tends to rise again. For minority women, however, these transitions may be delayed or disrupted by group-specific barriers. As such, aggregate trends may mask substantial heterogeneity within national populations.

In the post-independence era, many MENA governments placed women's emancipation at the heart of nation-building projects. It was widely believed that economic development and expanded access to education would naturally lead to greater gender equality. However, by the late 20th century, as states retreated from earlier progressive agendas, the ideal of gender equality came under increasing scrutiny, portrayed by some as incompatible with local values or national priorities. Women's rights became a symbolic battleground in broader ideological struggles: between secular nationalism and political Islam, and between imported models of modernity and claims to cultural authenticity. The rollback of earlier reforms, combined with rising authoritarianism and disillusionment with the promises of state-led development, fuelled growing scepticism toward state feminism. In this context, debates over gender justice were increasingly taken up by grassroots movements, NGOs, and religious actors – each promoting divergent visions of women's roles in society (Moghadam 2003).

Yet, these broad narratives were rarely accompanied by sustained empirical investigation into the complexity of women's economic involvement, including non-linear patterns of labour force participation, or regional, ethnic, and religious variation in outcomes. This has left many gender-focused policies based on overly generalized assumptions about a uniform female experience. Understanding the intersection of gender, ethnicity, and religion is therefore essential not only for diagnosing the roots of inequality but also for designing inclusive labour market policies that recognize the heterogeneity of women's lives and constraints across different communities in Syria and the broader region.

3.3. Migrant integration

The integration of migrants has been a longstanding focus in both historical and contemporary scholarship, driven by enduring concerns about the economic contributions of migrants, economic equity, and social cohesion. Scholars have approached migrant integration through multiple dimensions, often distinguishing between economic and social aspects. Economic integration generally refers to the

participation of migrants in the host country labour market and its rewards in line with their native counterparts, measured through employment rates, occupational attainment, income levels, and upward mobility (Abramitzky and Boustan 2017). In contrast, social integration encompasses a broader and more nuanced set of indicators. These include migrants' sense of belonging and self-identification, cultural preferences and values, levels of trust and social interaction with the majority population, civic and political participation, as well as patterns of demographic behaviour such as intermarriage and residential integration (Laurentsyeva and Venturini 2017). Together, these economic and social dimensions provide a comprehensive framework for assessing the extent to which migrants become part of the host society. The papers in this dissertation are concerned with the economic and demographic aspects of migrant integration.

Economic integration

In terms of economic integration, studies from the US during the age of mass migration at the turn of the 20th century show that immigrants typically experienced similar occupational and earnings growth to that of natives, with little convergence occurring in the first generation (Abramitzky and Boustan 2017). The degree of income and occupational assimilation depended on the country of origin, with factors like human capital stock determining the size and sign of the gap between immigrants and natives. On the other hand, children of immigrants experienced higher mobility as compared to their native counterparts (Boustan et al. 2025), with different factors contributing to this mobility including location choices (Abramitzky et al. 2021) and human capital accumulation (Connor 2020).

The existing literature on historical earnings assimilation faces several important limitations. Chief among these is the lack of reliable individual-level earnings data, particularly for historical contexts where such information was rarely recorded. As a result, researchers have often relied on indirect measures or proxies to estimate earnings. Notably, occupational income scores – such as those developed by Sobek (1995) and the more recent LIDO scores introduced by Saavedra and Twinam (2020) – have been widely used. These scores infer earnings based on an individual's occupation, sometimes incorporating additional factors like industry affiliation or estimated years of work experience to enhance precision. However, while these proxies provide a practical solution in the absence of direct data, they are not without shortcomings. A growing body of research has raised concerns about their validity, especially when used to analyse income disparities or intergenerational mobility. Studies by Inwood et al. (2019), Saavedra and Twinam (2020), and Ward (2023) highlight how occupational scores can obscure true income differences, potentially introducing bias or masking important heterogeneities. These critiques underscore the need for more refined methods or alternative data sources to accurately capture historical income dynamics.

Differences between historical and contemporary assimilation rates have often been a topic of debate. During the era of mass migration to the US, most migrants came from Europe, and the average migrant matched the income and occupations of the US born on arrival, despite considerable origin-country heterogeneity (Abramitzky and Boustan 2017). After the age of mass migration, and particularly after 1965, the composition of migrant origins shifted from Europe to the developing world, particularly Asia and Latin America. At the same time, most Western European countries transitioned into net migrant-receiving nations. Research shows that immigrants' earnings trajectories in the US differ markedly before and after 1965: pre-1965 immigrants tended to earn similarly to natives over time, while post-1965 immigrants started with lower earnings but experienced faster earnings growth (Duleep and Regets 2002; Duleep et al. 2025). A key finding from this literature is that earnings growth is inversely related to initial earnings. This pattern supports the immigrant human capital investment model: migrants from developing countries often arrive with fewer transferrable skills, but as they acquire host-country-specific skills, their earnings converge towards those of natives.

Social and demographic integration

A wide range of studies look at the social integration of immigrants through the lens of demography. Outcomes that capture demographic integration are strongly correlated with economic integration, like intermarriage between immigrants and natives (Dribe and Lundh 2008, Dribe and Nystedt 2015), as well as cohabitation (Elwert and Tegunimataka 2016). Fertility behaviour also provides insight into the extent of social integration (Scott and Stanfors 2010). Apart from union formation and childbearing, another important dimension of demographic integration lies in the timing and trajectory of the move out of the parental home, which marks a critical step toward adult independence and is often shaped by cultural norms, economic constraints, and social expectations. Recent studies have increasingly used this life event to assess the social integration of immigrants, comparing the timing and destinations of home-leaving between immigrant-origin youth and the native-born population, and the factors that contribute to differences between migrant and native outcomes (Gillespie et al., 2020; Kleinepier & de Valk, 2017; Zorlu & Mulder, 2011).

The research on leaving the parental home highlights how this seemingly private decision – such as when and how to leave the parental home – is in fact deeply structured by broader social forces, including personal preferences, economic conditions, and cultural expectations (Marini 1984, Goldscheider & Goldscheider 1987, De Valk and Liefbroer 2007). In contexts such as Sweden, where early home-leaving is the norm and enabled by institutional support, deviations from this trajectory among immigrant-origin youth can signal barriers to integration, whether economic, cultural, or institutional. At the same time, the act of home-leaving is not only a reflection of adaptation to the host society, but also of the negotiation between

inherited family norms and the opportunities and expectations found in the receiving context. Studying home-leaving behaviour thus provides a window into the intersection of structural constraints and identity formation during a key life transition and offers a valuable tool for assessing the extent and nature of social integration across migrant generations.

4. Data and methods

The dissertation uses a diverse set of data sources, in line with the diversity of the topics addressed in the four papers. Both historical and contemporary sources are used, ranging from village-level census data in Syria to longitudinal population registers in Sweden. While the contexts differ substantially, the unifying principle is to leverage population-wide datasets that capture entire societies rather than limited samples. This allows for a systematic study of inequality and integration across space, time, and institutional settings. The combination of cross-sectional, longitudinal, and historical data also makes it possible to address questions of both persistence and change, as well as to place the Syrian case in comparative perspective.

Methodologically, the dissertation combines quasi-experimental designs with decomposition techniques, applying each where it is best suited to the research question. While the specific approaches vary by paper, they share a concern with causal inference, the role of structural constraints, and the interaction of demographic, geographic, and institutional factors. Together, the different methods highlight complementary dimensions of inequality and integration and underscore the value of a pluralistic empirical approach. The methodological approaches in these papers include spatial regression discontinuity designs (RDD), decomposition techniques, panel regression models, and event-history models for competing risks. Each method is chosen to address the specific research questions posed by the paper, while collectively illustrating the value of combining quasi-experimental, decomposition-based, and longitudinal approaches in the study of inequality and integration.

Papers 1 and 2 rely on the Syrian population and housing census of 2004. This census dataset is aggregated at the level of village, town, and city, and offers a rich set of variables related to demographic composition, local economic conditions, physical capital, as well as educational composition and labour force composition. This dataset facilitates the study of regional economic inequality in a data-poor region, with no other countries in the MENA region having produced a publicly accessible dataset at this level of aggregation, to my knowledge. The census data is supplemented by the ethnic composition data produced by Khaddour and Mazur (2018), which provides estimates of the ethnic, religious, sectarian composition of

each village and town in the country, allowing for the analysis of ethnic and religious economic and social gaps and their correlates. Together, these datasets offer a unique view of economic inequality across the geography of Syria, and across its constituent ethnic and religious communities. Papers 1 and 2 also make use of a dataset of geographic and climatic conditions at the village level, constructed by the author, to account for geographic variability across the country in the analysis of regional economic inequality. This data was sourced through Google Earth Engine and created to capture the geographic and climatic diversity across the country at the time of the census enumeration in 2004. Paper 1 further makes use of secondary historical sources, in particular the study of 18th century settlement patterns in Syria by Norman Lewis (1987).

Paper 1 employs a spatial regression discontinuity design (RDD) to estimate the long-term causal effect of historical state presence on contemporary economic development. The method leverages a sharp spatial boundary – the historical boundary separating Ottoman-administered areas from stateless tribal territories – as an exogenous source of variation in historical governance. By comparing villages located just to the west and east of this line, the analysis identifies the discontinuity in present-day outcomes attributable to differences in historical state incorporation, holding geography constant. The empirical specification follows standard RDD conventions, fitting linear and polynomial functions of longitude and latitude on either side of the boundary (following Dell et al. 2018). Spatial autocorrelation is accounted for using spatial autoregressive lag models, while multiple placebo boundaries and falsification tests ensure that results are not artefacts of unobserved geographic gradients. This quasi-experimental design allows for credible inference on the persistence of historical state structures despite the absence of longitudinal microdata.

Paper 2 applies decomposition methods, specifically the Gelbach (2016) decomposition, to quantify the relative contribution of structural, demographic, and institutional factors to observed ethnic and religious gaps in female labour force participation (FLFP). The Gelbach approach extends the Oaxaca–Blinder decomposition by partitioning the change in the estimated group coefficient when controls are added into components attributable to each variable or set of variables. This makes it possible to identify which structural dimensions – such as education, income, demographic age structure, or access to public-sector employment – drive the observed disparities. The analysis proceeds hierarchically, estimating a baseline model with no controls, followed by the inclusion of variables grouped by domain (e.g., development, demographics, labour market, geography). By estimating the relative contribution of each set of variables to FLFP gaps, the method allows for highlighting the extent to which ethnic and religious FLFP gaps are structural rather than purely cultural.

Paper 3 makes use of individual-level full-count data from the US decennial censuses for the years 1900 to 1940. This data covers the entire population resident

in the US at the time of the census and captures the demographic characteristics of individuals, their country of origin, as well as their labour market outcomes (Ruggles et al. 2021). The data is supplemented by the IPUMS-MLP panel which allows for tracking the outcomes of individuals across census years (Helgertz et al. 2020), thus distinguishing assimilation over time from compositional changes in the migrant stock. The US census data is exceptional in its coverage of the entire population of the US at the height of the era of mass migration and has been used in prior studies of migrant integration in the US (Abramitzky and Boustan 2017). The analysis leverages multiple earnings measures, including occupational income scores, adjusted LIDO scores, and actual wage earnings. The empirical models estimate earnings and occupational outcomes as a function of years since migration, cohort, and generation, controlling for age, education, and location fixed effects. The analysis employs both fixed-effects estimators to verify robustness and uses decomposition of the earnings gap to assess the relative importance of human capital accumulation and geographic effects.

Paper 4 makes use of the Swedish population register data, which covers the entire population resident in Sweden in the years 1998 to 2022. The data contains information on the living conditions of individuals, allowing for tracking the transition from the parental home, alongside detailed yearly information on education, work, and income, as well as information on each individual's parents. The analysis uses event-history (survival) models with competing risks, suited to analysing time-to-event data where multiple outcomes are possible. The models estimate the hazard of leaving the parental home, distinguishing between two competing exits: independent living and marriage. Following Fine and Gray (1999), the subdistribution hazard approach allows simultaneous estimation of the relative risk of each transition while accounting for censoring and the mutually exclusive nature of the outcomes. Time is measured in single-year intervals from age 17 until either event occurs or censoring at age 35. Covariates include individual and parental socio-economic characteristics, education, and income, as well as contextual neighbourhood indicators such as ethnic composition and unemployment rate. The analysis is further stratified by gender and generation, providing insights into how cultural norms, economic opportunities, and residential contexts interact to influence integration trajectories. The methodological approach uses competing risks models to estimate the likelihood and timing of leaving the parental home, distinguishing between independent living and marriage as mutually exclusive pathways. The models incorporate both individual- and parental-level controls and exploit neighbourhood-level data, including the ethnic composition and unemployment rates of local areas, to assess how context shapes home-leaving behaviour.

Each dataset comes with its own limitations – such as the aggregate nature of the Syrian census, potential misreporting in historical US data, or the absence of attitudinal measures in Swedish registers. These limitations are addressed through

robustness checks and the use of multiple indicators. By combining sources that differ in scope, scale, and period, the dissertation is able to balance depth with breadth and to demonstrate that the patterns observed are not artefacts of a single dataset or method.

5. Summary of the papers

Paper 1

The Historical State and Path Dependence: Evidence from Ottoman Syria's Stateless Frontier

This paper explores how historical statelessness has shaped long-term economic development in Syria. It focuses on a historical boundary which marked the eastern limit of effective Ottoman state control prior to the mid-19th century (Lewis 1987). Areas west of the boundary were incorporated into the Ottoman state from the early 16th century, while areas to the east remained under tribal control, suffering from insecurity, land abandonment, and limited state investment. Using a spatial regression discontinuity design (RDD), the paper compares contemporary economic outcomes on either side of this boundary to evaluate the long-term effects of historical state absence.

The observed disparities are substantial. Incomes in historically stateless areas are lower by about half a standard deviation, and these areas have worse access to infrastructure such as sanitation and water networks. They also show lower levels of university attainment and higher illiteracy. Labour force characteristics similarly reflect underdevelopment, with fewer managerial workers and a larger share of agricultural employment. These differences persist after controlling for geography, proximity to major cities, and ethnic composition, suggesting that they are not driven by environmental or cultural factors.

A key contribution of the study is to show that the main mechanism of persistence lies in population density. Areas with a history of statelessness have lower settlement density, both in terms of the size of individual villages and the number of surrounding settlements. These density gaps explain roughly half of the differences in income and education, and even more of the disparity in infrastructure provision. The paper links these outcomes to a path-dependent process: historical insecurity and lack of state protection discouraged permanent settlement and economic investment, setting in motion long-term patterns of outmigration and underdevelopment. This is further supported by demographic data showing a persistent deficit in the working-age male population, indicating ongoing male-selective emigration from historically stateless areas.

The paper also examines alternative explanations. One possibility is that weaker private property institutions in stateless regions have contributed to underdevelopment. The data show higher levels of informality in property ownership in stateless areas, but these institutional differences have limited explanatory power for the economic gaps observed. Similarly, controlling for ethnic and religious differences – including the presence of Arab tribal communities and religious minorities – does not eliminate the estimated effects. The results hold even when excluding tribal areas or minority-majority towns, indicating that the developmental gap is not simply a reflection of cultural or identity-based factors.

By focusing on a case of historical state absence rather than variation in state quality or centralization, this paper extends the literature on institutional persistence and long-run development. It also connects with research on economic geography and path dependence, showing how early state presence shaped population distribution, which in turn influenced the spatial structure of economic opportunity. The study sheds new light on why certain areas in Syria, and by extension in other parts of the Middle East, remain underdeveloped even in the absence of major geographic constraints.

The findings have broader implications for development policy. If lower density and historical neglect continue to constrain development, then spatially targeted interventions – particularly investments in infrastructure, connectivity, and human capital – may be necessary to overcome these historical legacies. As climate change and pastoral-agricultural conflict increasingly shape frontier regions in the Global South, the Syrian case offers a cautionary example of how early patterns of governance and settlement can cast a long shadow over future economic potential.

Paper 2

Ethnic and Religious Differences in Female Labour Force Participation: Evidence from Syrian Census Data

Published online in the Journal of Demographic Economics (2025), 1-33.

This paper investigates the extent and underlying causes of ethnic and religious disparities in female labour force participation (FLFP) in pre-war Syria, challenging the dominant narrative that cultural or religious norms are the principal drivers of gaps in FLFP. Instead, it shows that structural factors – particularly economic development, demographic profiles, and public sector employment – play a larger role in explaining these disparities.

Using data from the 2004 Syrian population and housing census, the study documents substantial differences in FLFP across Syria's diverse ethnic and religious communities. While FLFP averaged only 10% in Sunni Arab majority areas, it reached 28% in Alawi-majority areas and 26% in Christian-majority areas.

These gaps appear at least as large between Muslim sects as between Muslims and Christians, suggesting the importance of examining sub-national variation within the Muslim majority. The paper uses a decomposition method (Gelbach decomposition) to isolate and quantify the contribution of various explanatory factors – including income, education, age structure, and gender parity in education – to these differences.

The analysis reveals that structural factors account for most of the observed disparities. Among the most influential factors are income levels, the age distribution of the population, and the share of public sector employment. Public sector jobs - often tied to political privilege in Syria - play a significant role in elevating FLFP among Alawis, who were politically dominant during the Assad era. Education levels are also important, particularly among Christian communities, which historically had earlier access to missionary and private schools. However, the paper finds that gender gaps in school enrolment and attainment – used as proxies for gender bias or restrictive norms – have only a limited impact in explaining FLFP differences across groups. Even where educational gender gaps exist, they explain only around 10% of the observed FLFP advantage in Christian and Alawi communities.

The study also highlights the importance of demographic structure, particularly the share of working-age population. Groups with higher proportions of the population in prime working age – such as Alawis and Christians – exhibit higher FLFP, suggesting a demographic dividend effect. Conversely, Sunni Muslim communities, which have higher dependency ratios and younger populations, show lower FLFP. Notably, higher fertility rates do not appear to play a major role in explaining group differences.

Contrary to expectations that higher education universally promotes women's workforce participation, the study finds a U-shaped relationship between development and FLFP. In poorer, agrarian regions (e.g. Kurdish and tribal areas), FLFP is relatively high due to women's involvement in family-based agricultural work. In middle-income areas, FLFP drops sharply, especially among Sunni families. In more developed areas with strong service and public sectors, FLFP rises again, especially among Alawis and Christians. This mirrors the U-feminization hypothesis (Boserup 1970, Goldin 1995) which posits that women's workforce participation first declines and then rises as economies shift from agriculture to services.

The analysis also considers the potential role of property rights, infrastructure, and geography. The contribution of these factors is modest. Housing informality and infrastructure gaps are somewhat relevant in poorer regions but do not significantly explain FLFP gaps across groups. Geographic factors such as elevation and agricultural suitability are examined as potential sources of historically embedded

gender norms, but they do not account for much of the variation in women's workforce participation.

The study's findings challenge the notion that cultural or religious conservatism is the primary explanation for low FLFP in MENA societies. Instead, they show that ethnic and religious disparities in FLFP in Syria reflect long-standing inequalities in development, public sector access, and demographic transitions. By comparing outcomes across homogenous localities (i.e. towns and villages with a clear ethnic or religious majority), the study provides a conservative estimate of group differences. It is likely that within-city comparisons, where group interactions are more frequent, would reveal smaller gaps due to acculturation and similar socioeconomic profiles.

Finally, the paper considers the broader relevance of these findings in the context of Syria's civil war and future reconstruction. Although FLFP in Syria increased significantly during the war - likely due to economic necessity and male displacement - the underlying structural patterns are expected to persist. Unless regional inequalities in income, public sector access, and educational attainment are addressed, FLFP is likely to remain uneven across Syria's ethnic and sectarian landscape. The study offers broader insights into gender inequality in the Middle East and reinforces the importance of economic policy and demographic factors over essentialist assumptions about culture or religion.

Paper 3

Premium or Penalty? Occupations and Earnings of Ottoman Immigrants and Their Offspring in the United States, 1900–1940

Co-authored with Martin Dribe.

Published in the European Review of Economic History.

This paper examines the economic integration of immigrants from Ottoman Syria and Turkey and their offspring in the United States from 1900 to 1940. It contributes to the literature on migrant integration during the age of mass migration by analysing an overlooked group using large-scale linked census data. The study investigates whether these Middle Eastern immigrants experienced economic premiums or penalties, and how their labour market outcomes evolved over time and across generations, compared to both US natives and European immigrant groups.

Findings reveal that Ottoman immigrants, particularly those from Syria-Lebanon, held substantial occupational score premiums upon arrival. These premiums persisted or even increased slightly with time spent in the US. However, when actual earnings are used instead of occupational scores, the picture shifts: real earnings

data show substantial and growing penalties for Ottoman immigrants, especially after 10 or more years in the US. For instance, by 1940, immigrants from Syria-Lebanon earned $3,700–$5,500 less than their native-born counterparts, and those from Turkey earned $3,400–$4,200 less. These findings highlight a divergence between occupational prestige and actual income, which is explained by self-selection into high-reward sectors like trade and retail.

The second generation showed marked upward mobility. They closed the earnings gap with natives and, in some cases, even achieved parity or slight premiums. This generational convergence aligns with broader findings in the migration literature suggesting high intergenerational mobility, especially when immigrants settle in economically vibrant regions. While immigrants often entered the labour market with lower levels of formal education, the second generation had educational attainment similar to natives, which played a crucial role in their economic advancement.

The analysis reveals that structural factors such as occupational selection and education explain much of the observed gaps in earnings. Ottoman immigrants were heavily concentrated in retail and self-employment, which inflated their occupational scores but did not necessarily translate into high earnings. Within occupation comparisons confirm this: Ottoman immigrants consistently earned less than natives within the same occupational category, with the largest penalties among managers and sales workers. In contrast, earnings were more equal – or even favourable – for Ottoman immigrants working as labourers or operatives.

A decomposition analysis shows that years of education explain around 40–45% of the earnings gap for Ottoman immigrants, highlighting the role of human capital. Italian immigrants showed similar patterns of disadvantage and convergence, while Irish immigrants had consistent earnings premiums. These comparative insights help position the Ottoman case within the broader spectrum of European and non-European migration.

Ultimately, the study underscores the complexity of immigrant economic integration. Ottoman immigrants were positively selected into commerce-related occupations but faced structural barriers in the labour market that suppressed their actual earnings. Nevertheless, their children made significant strides, supporting the narrative of second-generation assimilation through education and occupational mobility. By focusing on a non-European group, the paper challenges the Eurocentric bias of much of the migration literature and adds nuance to the discussion of ethnicity, occupation, and intergenerational mobility in the American labour market.

Paper 4

Home-Leaving Patterns Among MENA-Origin Youth in Sweden: The Influence of Origin, Generation, and Neighbourhood

Co-authored with Anna Tegunimataka.

This study investigates the transition to independent living among Middle Eastern and North African (MENA)-origin youth in Sweden, analysing how origin, generation, gender, and neighbourhood characteristics influence the timing and pathway of leaving the parental home. Using rich longitudinal population register data from 1998 to 2022, it tracks individuals aged 17 to 35, applying competing risks models to examine the transition to independent living, treating marriage as a competing event. The paper reveals stark generational and gendered differences in home-leaving behaviour, with persistent integration gaps that remain even after controlling for socio-economic background.

In Sweden, where the welfare state supports early transitions to adulthood through accessible housing and education, native youth tend to leave the parental home relatively early. However, youth of MENA origin leave significantly later, especially females. While about 98% of native Swedes transition to independent living by age 35, this share drops to around 78% for second-generation MENA youth. The differences are most pronounced in direct transitions to marriage: only 1.5% of native youth leave home to marry, compared to up to 14% among MENA second-generation youth. Meanwhile, a substantial share of MENA-origin youth, particularly women, remain in the parental home well into adulthood, even when accounting for education, employment, and income.

The study categorizes individuals into six generational groups (e.g., 1st generation, 1.5 generation, 2nd generation, etc.) and compares MENA-origin youth to both native Swedes and Eastern European immigrants. Findings show that first-generation MENA youth who arrive in Sweden before age 6 (gen 1.75) exhibit more similar patterns to natives than those arriving at later ages or born to two foreign-born parents (gen 2). Interestingly, the largest gaps emerge not in the first generation, but in the second generation, those born in Sweden to immigrant parents, indicating that cultural norms around co-residence and marriage may persist or even become more entrenched for the second generation, though this effect may also be explained by differential selection patterns across the first and second generation.

Gender differences are stark. Women from MENA backgrounds are significantly less likely to leave home independently than men, and more likely to leave directly into marriage. This pattern may reflect the influence of traditional gender roles emphasizing family control and female modesty, which can discourage early autonomy for women. Even in high-income households, second-generation MENA women remain significantly less likely to leave home independently than their

native-born counterparts, suggesting that the persistence of cultural norms outweighs economic barriers.

The role of socio-economic status, particularly parental income, is surprisingly limited. Even when MENA-origin youth come from families in the top income quartile, they are still significantly less likely to leave home independently compared to natives from the bottom income quartile. This pattern holds across multiple generations and persists after controlling for education, employment, and occupation. Thus, structural integration in terms of economic resources does not translate into convergence in social integration milestones like home-leaving.

The paper also examines how neighbourhood context shapes home-leaving behaviour. Youth from MENA backgrounds who grow up in immigrant-dense neighbourhoods with low shares of native Swedes and high unemployment are substantially less likely to leave the parental home. In high-unemployment, low-native share areas, where many second-generation MENA youths live, the effect is particularly large, especially for women. These neighbourhoods may reinforce community norms of intergenerational co-residence and delay marriage or union formation, while also limiting access to housing, jobs, and mainstream social networks. Although these effects could partly reflect selection as families who value cultural retention may cluster in such areas. The results remain robust after controlling for parental income, employment, and education, indicating an independent role of neighbourhood-level influences.

Despite these persistent gaps, the majority of MENA-origin youth do eventually leave home. By age 35, over 75% of second-generation MENA youth live independently, and fewer than 10% remain in the parental home. This marks a clear departure from norms in many MENA countries, where home-leaving – especially for women – is almost exclusively tied to marriage. The findings therefore suggest partial integration into Swedish norms of early independence, particularly among those in mixed-origin families (generation 2.5), or those raised in neighbourhoods with higher native shares and lower unemployment.

This study expands the literature on migrant integration by showing that social integration outcomes, such as the timing of leaving the parental home, are shaped by a complex interplay of cultural norms, generational change, gender roles, and neighbourhood context. It emphasizes that integration is not simply a matter of income or education, but also of embedded social expectations and environments. The delayed transitions observed among MENA-origin youth, and especially among women, underscore the need for more nuanced policy approaches that address the interaction of cultural and structural barriers to full social inclusion.

6. Concluding discussion

Syria experienced profound transformations throughout the 20th century, marked by drastic shifts in its economic structure and demographics. The country's diversity was both a source of richness and a source of tensions, intersecting with entrenched inequalities and the burdens of historical legacies. Stark disparities between urban and rural areas, uneven development across regions, and divisions along ethnic and confessional lines not only complicated the project of nation-building but also contributed to instability and conflict. At the same time, migration, both internal and international, emerged as a crucial avenue for social mobility and economic security. Yet migration also created new cleavages and reproduced inequalities both in Syria and the host societies that received Syrian immigrants.

These dynamics form the backdrop to the central questions of this dissertation: what explains the persistence of regional and social inequalities in Syria, and how do migrants from Syria and the broader Middle East integrate across generations and contexts? By situating these questions within Syria's long history of uneven development and migration, the dissertation brings together four papers that examine these themes from different angles. In doing so, it provides new empirical answers and connects them to wider debates in economic history, development economics, and migration studies.

The Syria papers demonstrate the importance of historical legacies and structural inequalities. Paper 1 shows that areas lying beyond Ottoman state control continue to lag behind in income, education, and infrastructure. These disparities cannot be reduced to culture or geography but are instead explained by population settlement patterns, where settlement was suppressed by insecurity in stateless frontier regions. The result is a clear case of path dependence: historical state presence shaped population geography, which in turn structured long-run economic concentration and opportunities. Paper 2 shows that ethnic and religious differences in female labour force participation (FLFP) similarly reflect structural rather than purely cultural factors. Regional economic development, demographic composition, and access to public sector employment explain the bulk of observed gaps, while gender norms as proxied by school enrolment differentials play a much smaller role.

The migration studies reveal different dynamics across time and place. Paper 3 shows that Ottoman immigrants to the US arrived with high occupational prestige but faced significant earnings penalties. Yet their children achieved rapid upward mobility through education and occupational change, closing the gap with natives. Paper 4 shows that in contemporary Sweden, MENA-origin youth lag behind natives in leaving the parental home, especially women, and that these differences are driven less by parental income and education than by gender norms and neighbourhood context. In contrast to the US case, integration here is slower and

more uneven, suggesting that structural opportunity is necessary but not sufficient when cultural continuity and residential segregation reinforce social distance.

The evidence from Syria contributes to the literature on institutions and long-run development (Acemoglu et al. 2001; Michalopoulos & Papaioannou 2013; Dell et al. 2018). While prior studies emphasized state centralization, Paper 1 highlights the enduring consequences of state absence, mediated through population density and settlement patterns. This mechanism links the institutional persistence tradition (Nunn 2020) with the economic geography of agglomeration (Bleakley & Lin 2012; Baerlocher et al. 2024). It also resonates with frontier studies in other contexts, such as Oto-Peralías' (2020) work on Spain, by showing how historical insecurity generated disadvantages that persist today.

The analysis of women's labour force participation in paper 2 challenges culturalist accounts of women's low participation in MENA economies (Ross 2008; Alesina et al. 2013) by showing that structural inequalities are central. It supports a strand of work that emphasizes the interaction of economic development, labour market structure, and demographic transitions (Goldin 1995; Assaad et al. 2020; Klasen 2019). By demonstrating that ethnic and religious gaps within a single country mirror the U-shaped development trajectory of women's work, The analysis situates Syria within broader comparative patterns, while also highlighting the role of political privilege (through public sector access) in shaping gendered outcomes.

Paper 3 contributes to the literature on immigrant assimilation during the age of mass migration. Existing studies have focused overwhelmingly on European groups (Abramitzky & Boustan 2017; Connor 2020; Ward 2020). By examining Ottoman migrants, the analysis adds a Middle Eastern perspective and shows both the limitations of occupational income scores as proxies (Saavedra & Twinam 2020; Ward 2023) and the importance of intergenerational human capital accumulation in overcoming first-generation disadvantages. The findings reinforce the narrative of high second-generation mobility (Boustan et al. 2025) while expanding its applicability beyond European migrants.

Paper 4 contributes to the literature on migrant social and demographic integration. Studies have examined intermarriage, fertility, and home-leaving as indicators of integration (Goldscheider & Goldscheider 1987; Zorlu & Mulder 2011; Gillespie et al. 2020). The analysis in this paper shows that even in a generous welfare state with extensive support for youth independence, cultural norms and neighbourhood context produce persistent gaps. The study underscores the importance of considering not only economic resources but also the social environments in which migrants grow up, echoing insights from research on segregation and neighbourhood effects.

Taken together, these findings point to three broader lessons. First, economic geography and state capacity matter deeply: historical inequalities of settlement and governance are not easily erased and continue to structure opportunity. Second,

social norms and structural factors interact, neither culture nor economics alone accounts for gender and integration outcomes; their interplay must be recognised. Third, integration is dynamic and conditional: while barriers can persist across generations, they can also be overcome when institutions enable access to education, employment, and housing.

The findings also carry policy implications. For Syria and the MENA region, they highlight the importance of spatially targeted investment in historically disadvantaged areas and of reducing structural barriers to women's participation in the labour force. For host societies, they suggest that labour market policies alone are insufficient: housing, neighbourhood composition, and gendered pathways to adulthood must be part of integration strategies.

Although the lessons that emerge from this thesis are grounded in the specific topics and themes addressed in each paper, several cross-cutting themes stand out that warrant deeper reflection. These themes resonate within the MENA region and the broader developing world and also carry broader relevance for countries that host migrant populations from developing countries.

One recurring insight across the thesis is the central role of economic geography in shaping patterns of development and inequality. Paper 1 underscores how historical variations in state presence have left an enduring imprint on economic development, primarily by influencing the spatial distribution of populations. This legacy continues to shape the diverse economic trajectories observed across different parts of the country, reflecting part of a complex and layered historical landscape.

Similarly, Paper 2 sheds light on the determinants of gender inequality across ethnic and religious lines, showing that these disparities are closely linked to regional economic conditions. The paper reveals that where women live - and the historical and structural economic dynamics of those areas - matters significantly for their opportunities and constraints. Both papers, in their distinct ways, highlight how geographic disparities in development are not merely background conditions but active forces shaping social and economic inequalities.

This thesis, therefore, contributes to a growing body of scholarship that re-centres economic geography in analyses of social and economic outcomes, particularly in developing country contexts and in the MENA context. While much contemporary social science emphasizes the normative role of institutions and cultural norms in shaping behaviour and outcomes, the findings here remind us that these forces often operate atop a more fundamental economic base. Regional disparities in infrastructure, access to education, labour market opportunities, and state capacity are not just outcomes of political choices but also reflect deeper, structural forces that have accumulated over generations.

The interplay of economic development and geography does not matter solely for developing countries. Migrants that arrive in host countries bring the legacies of

their origin countries with them, and migrant workers from developing countries are often at a disadvantage due to their lower human capital endowment. Paper 3 shows this clearly in the case of Ottoman immigrants in US, and it also shows how the second generation surpasses these legacies and achieves parity with native workers by accumulating human capital. This suggests that, while initial conditions shape early integration outcomes, opportunities for intergenerational mobility remain significant, particularly when access to education and supportive institutions are in place. It also underscores that migrant integration is not a static process but one that evolves over time, shaped by both the background characteristics of migrants and the broader structural opportunities in the host country.

Emphasizing structural factors does not mean that social norms are unimportant. On the contrary, Paper 4 reveals that the integration trajectories of MENA-origin youth in Sweden are deeply intertwined with gender. The findings suggest that economic factors - such as employment or income – are not the primary drivers of integration outcomes in this context. Instead, gender gaps are particularly important, which suggest that gender-specific attitudes and preferences toward independent living appear to exert a significant influence. This highlights the need to consider how social norms interact with economic structures, shaping the pathways through which young people integrate into their host societies. It also points to the importance of policies that not only address material barriers to integration but also engage with the social and cultural dimensions that underpin them.

The findings from this thesis, while drawn from distinct case studies and analytical approaches, collectively highlight the complex interplay between economic structures, geographic disparities, and social norms in shaping economic and social outcomes. The thesis illuminates recurring themes and insights that resonate across contexts. The evidence points to the enduring significance of regional economic inequality, the intergenerational potential of human capital accumulation for migrant populations, and the persistent influence of social and gender norms in shaping life course transitions for young people.

Future research can build on these findings by expanding beyond Syria to other MENA contexts, by further integrating historical and contemporary data, and by analysing how migrants' origin-country inequalities interact with their host-country outcomes.

In conclusion, the evidence presented here shows how borders – whether the frontier zones of past states, ethnic and religious divides, or the dividing lines of national origin – shape inequality. These borders constrain opportunity, but pathways exist which can bridge divides, whether through economic development, inclusive labour markets, access to education, or changing community norms.

References

Abramitzky, R., Boustan, L., Jácome, E., & Pérez, S. (2021). Intergenerational mobility of immigrants in the United States over two centuries. *American Economic Review, 111*(2), 580-608.

Abramitzky, R., & Boustan, L. (2017). Immigration in American economic history. *Journal of Economic Literature, 55*(4), 1311-1345.

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2014). A nation of immigrants: Assimilation and economic outcomes in the age of mass migration. *Journal of Political Economy, 122*(3), 467-506.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review, 91*(5), 1369-1401.

Akyol, P., & Ökten, Ç. (2024). The role of religion in female labor supply: evidence from two Muslim denominations. *Journal of Demographic Economics, 90*(1), 116-153.

Alesina, A., Giuliano, P., & Nunn, N. (2013). On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics, 128*(2), 469-530.

Angermayr, G., Anthony, B., Dinc, P., & Eklund, L. (2023). *The Syrian climate-migration-conflict nexus: an annotated bibliography.* CMES-Center for Advanced Middle Eastern Studies, Lund University.

Assaad, R., Hendy, R., Lassassi, M., & Yassin, S. (2020). Explaining the MENA paradox: Rising educational attainment, yet stagnant female labor force participation. *Demographic Research, 43*, 817.

Assaad, R. (2014). Making sense of Arab labor markets: The enduring legacy of dualism. *IZA Journal of Labor & Development, 3*(1), 6.

Atasoy, B. S. (2017). Female labour force participation in Turkey: The role of traditionalism. *The European Journal of Development Research, 29*(4), 675-706.

Baerlocher, D., Firmino Costa da Silva, D., Lambais, G., Reis, E., & Veras, H. (2025). Old but gold: historical pathways and path dependence. *SSRN Working Paper.*

Bai, Y., & Kung, J. K. S. (2011). Climate shocks and Sino-nomadic conflict. *Review of Economics and Statistics, 93*(3), 970-981.

Baroutt, M. J. (2022). *Past decade in Syria: the dialectic of stagnation and reform (Part 1 of 5).* Arab Center for Research and Policy Studies.

Bayanpourtehrani, G., & Sylwester, K. (2013). Female labour force participation and religion: a cross-country analysis. *Bulletin of Economic Research, 65*(2), 107-133.

Bengtsson, T., Lundh, C., & Scott, K. (2005). From boom to bust. The economic integration of immigrants in post war Sweden. In *European migration: What do we know*, 15-58. Springer.

Bleakley, H., & Lin, J. (2012). Portage and path dependence. *The Quarterly Journal of Economics, 127*(2), 587-644.

Bolt, J., & Van Zanden, J. L. (2025). Maddison-style estimates of the evolution of the world economy: A new 2023 update. *Journal of Economic Surveys, 39*(2), 631-671.

Bolt, J., & Van Zanden, J. L. (2014). The Maddison project: collaborative research on historical national accounts. *The Economic History Review, 67*(3), 627-651.

Borazan, A. (2021). The Political Economy of Rural Syria: Narrative, Class, and Conflict (1970-2011). *Ekonomi Politika ve Finans Araştırmaları Dergisi, 6*(1), 82-100.

Borjas, G. J. (1994). The economics of immigration. *Journal of Economic Literature, 32*(4), 1667–1717

Boserup, E. (1981). *Population and Technology.* Oxford: Blackwell.

Boserup, M. (1965). Agrarian structure and take-off. In W. W. Rostow (Ed.), *The Economics of Take-Off into Sustained Growth* (pp. 201-224). Palgrave Macmillan UK.

Boserup, E. (1970). *Woman's role in economic development.* London: George Allen & Unwin.

Boustan, L., Jensen, M. F., Abramitzky, R., Jácome, E., Manning, A., Pérez, S., ... & Zohar, T. (2025). Intergenerational Mobility of Immigrants in 15 Destination Countries (No. w33558). *National Bureau of Economic Research Working Paper.*

Card, D., DiNardo, J., & Estes, E. (1998). The More Things Change: Immigrants and the Children of Immigrants in the 1940s, the 1970s, and the 1990s. In *Issues in the Economics of Immigration.* University of Chicago Press.

Connor, D. S. (2020). Class background, reception context, and intergenerational mobility: A record linkage and surname analysis of the children of Irish immigrants. *International Migration Review, 54*(1), 4-34.

Courbage, Y. (1994). Fertility transition in Syria: From implicit population policy to explicit economic crisis. *International Family Planning Perspectives,* 142-146.

Dalgaard, C. J., Kaarsen, N., Olsson, O., & Selaya, P. (2022). Roman roads to prosperity: Persistence and non-persistence of public infrastructure. *Journal of Comparative Economics, 50*(4), 896-916.

Dell, M., Lane, N., & Querubin, P. (2018). The historical state, local collective action, and economic development in Vietnam. *Econometrica, 86*(6), 2083-2121.

Dildar, Y. (2015). Patriarchal norms, religion, and female labor supply: Evidence from Turkey. *World Development, 76*, 40-61.

Dribe, M., & Nystedt, P. (2015). Is there an intermarriage premium for male immigrants? Exogamy and earnings in Sweden 1990–2009. *International Migration Review, 49*(1), 3-35.

Dribe, M., & Lundh, C. (2008). Intermarriage and immigrant integration in Sweden: An exploratory analysis. *Acta Sociologica, 51*(4), 329-354.

Dribe, M., & Stanfors, M. (2005). Leaving home in post-war Sweden: A micro-level analysis of leaving the parental home in three birth cohorts. *Scandinavian Economic History Review, 53*(2), 30-49.

Duleep, H., Dowhan, D., Liu, X., Regets, M., & Gesumaria, R. (2025). A Historical Note on the Assimilation Rates of Foreign-born Men and Women in the US (No. 1596). *GLO Discussion Paper.*

Duleep, H. O., & Regets, M. C. (2002). The elusive concept of immigrant quality: evidence from 1970-1990 (No. 631). *IZA Discussion Papers.*

El Laithy, H., & Abu-Ismail, K. (2005). *Poverty in Syria: 1996-2004: Diagnosis and Pro-Poor Policy Considerations.* United Nations Development Programme.

Elwert, A., & Tegunimataka, A. (2016). Cohabitation premiums in Denmark: Income effects in immigrant–native Partnerships. *European Sociological Review, 32*(3), 383-402.

Engerman, S., & Sokoloff, K. L. (2000). Institutions, factor endowments, and paths of development in the new world. *Journal of Economic perspectives, 14*(3), 217-232.

Escamilla-Guerrero, D., Kosack, E., & Ward, Z. (2021). Life after crossing the border: Assimilation during the first Mexican mass migration. *Explorations in Economic History, 82*, 101403.

Gaddis, I., & Klasen, S. (2014). Economic development, structural change, and women's labor force participation: A reexamination of the feminization U hypothesis. *Journal of Population Economics, 27*(3), 639-681.

Gillespie, B. J., Bostean, G., & Malizia, S. (2020). Timing of departure from the parental home: Differences by immigrant generation and parents' region of origin. Hispanic *Journal of Behavioral Sciences, 42*(2), 165-190.

Giuliano, P. (2017). Gender: An historical perspective. *IZA Discussion Paper Series* (No. 10931).

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review, 104*(4), 1091-1119.

Goldin, C. (1994). The U-shaped female labor force function in economic development and economic history. *NBER Working Paper Series* (No. 4707).

Gualtieri, S. (2009). *Between Arab and White: Race and ethnicity in the early Syrian American diaspora* (Vol. 26). Univ of California Press.

Haddad, B. S. (2020). *Business networks in Syria: The political economy of authoritarian resilience.* Stanford University Press.

Hammarstedt, M., & Palme, M. (2012). Human capital transmission and the earnings of second-generation immigrants in Sweden. *IZA Journal of Migration, 1*(1), 4.

Ritchie, H., Samborska, V., Ahuja, N., Ortiz-Ospina, E., and Roser, M. (2023). *Global Education*. Our World in Data. https://ourworldindata.org/global-education

Hinnebusch, R. A. (2021). *Authoritarian Power and State Formation in Bathist Syria: Army, Party, and Peasant.* Routledge.

Hinnebusch, R. (2004). *Syria: Revolution From Above.* Routledge.

Inwood, K., Minns, C., & Summerfield, F. (2019). Occupational income scores and immigrant assimilation. Evidence from the Canadian census. *Explorations in Economic History, 72*, 114-122.

İlkkaracan, İ. (2012). Why so few women in the labor market in Turkey? *Feminist Economics, 18*(1), 1-37.

Kešeljević, A., & Spruk, R. (2024). Estimating the effects of Syrian civil war. *Empirical Economics, 66*(2), 671-703.

Khaddour, K., & Mazur, K. (2018). *Syria town database.* Harvard Dataverse, V1. https://doi.org/10.7910/DVN/YQQ07L

Khater, A. F. (2001). *Inventing home: Emigration, gender, and the middle class in Lebanon, 1870-1920.* University of California Press.

Klasen, S. (2019). What explains uneven female labor force participation levels and trends in developing countries? *The World Bank Research Observer, 34*(2), 161-197.

Kleinepier, T., & Valk, H. A. D. (2017). Leaving home of migrant and Dutch young adults: Parent-child and peer relations. *Journal of Comparative Family Studies, 48*(1), 15-39.

Krugman, P. (1992). *Geography and Trade.* MIT press.

Laurentsyeva, N., & Venturini, A. (2017). The social integration of immigrants and the role of policy–A literature review. *Intereconomics, 52*(5), 285-292.

Ma'oz, M. (2013). The emergence of modern Syria. In *Syria Under Assad (RLE Syria)* (pp. 9-35). Routledge.

Mancini, G. (2018). Women's Labor Force Participation in Italy, 1861-2011. *Rivista di Storia Economica, 34*(1), 3-68.

Mazur, K. (2021). *Revolution in Syria: Identity, networks, and repression.* Cambridge University Press.

McGuirk, E. F., & Nunn, N. (2025). Transhumant pastoralism, climate change, and conflict in Africa. *Review of Economic Studies, 92*(1), 404-441.

Mehchy, Z., & Doko, A. M. (2011). General Overview of Migration into, through and from Syria. *CARIM Analytic and Synthetic Notes* 2011/41.

Michaels, G., & Rauch, F. (2018). Resetting the urban network: 117–2012. *The Economic Journal, 128*(608), 378-412.

Michalopoulos, S., & Papaioannou, E. (2013). Pre-colonial ethnic institutions and contemporary African development. *Econometrica, 81*(1), 113-152.

Mitchell, B. R. (1998). Education. In *International Historical Statistics: Africa, Asia & Oceania, 1750–1993* (pp. 965-1007). London: Palgrave Macmillan UK.

Moghadam, V. M. (2003). *Modernizing women: Gender and social change in the Middle East.* Lynne Rienner Publishers.

Naff, A. (1993). *Becoming American: the early Arab immigrant experience.* SIU Press.

Lewis, N. N. (1987). *Nomads and settlers in Syria and Jordan, 1800-1980.* Cambridge University Press.

North, D. C., Wallis, J. J., & Weingast, B. R. (2009). *Violence and social orders: A conceptual framework for interpreting recorded human history.* Cambridge University Press.

Nunn, N. (2020). The historical roots of economic development. *Science, 367*(6485), eaaz9986.

Olivetti, C., & Petrongolo, B. (2016). The evolution of gender gaps in industrialized countries. *Annual Review of Economics, 8*(1), 405-434.

Oto-Peralías, D. (2020). Frontiers, warfare and economic geography: The case of Spain. *Journal of Development Economics, 146*, 102511.

Owen, R. (1981). *The Middle East in the world economy, 1800-1914.* London and New York: Methuen.

Pamuk, Ş. (2006). Estimating economic growth in the Middle East since 1820. *The Journal of Economic History, 66*(3), 809-828.

Ross, M. L. (2008). Oil, Islam, and women. *American Political Science Review, 102*(1), 107-123.

Saavedra, M., & Twinam, T. (2020). A machine learning approach to improving occupational income scores. *Explorations in Economic History, 75,* 101304.

Saleh, M. (2021). The Middle East: Decline and resurgence in West Asia. In S. Broadberry & K. H. O'Rourke (Eds.), *The Cambridge economic history of modern Europe: Volume 2, 1870 to the present* (pp. 213–250). Cambridge University Press.

Scott, K., & Stanfors, M. (2010). Second Generation Mothers-Do the children of immigrants adjust their fertility to host country norms? In *Demographic Aspects of Migration* (pp. 123-152). Wiesbaden: VS Verlag für Sozialwissenschaften.

Stanfors, M., & Goldscheider, F. (2017). The forest and the trees: Industrialization, demographic change, and the ongoing gender revolution in Sweden and the United States, 1870–2010. *Demographic Research, 36,* 173-226.

Sobek, M. (1995). The comparability of occupations and the generation of income scores. *Historical Methods: A Journal of Quantitative and Interdisciplinary History, 28*(1), 47-51.

United Nations, Department of Economic and Social Affairs, Population Division. (2019). *World fertility data 2019 (POP/DB/Fert/Rev2019).* United Nations.

Ward, Z. (2020). The low return to English fluency during the Age of Mass Migration. *European Review of Economic History, 24*(2), 219-242.

Zorlu, A., & Mulder, C. H. (2011). Ethnic differences in leaving home: Timing and pathways. *Demography, 48*(1), 49-72.

Paper I

The Historical State and Path Dependence: Evidence from Ottoman Syria's Stateless Frontier

Rami Zalfou

Abstract

This paper investigates the long-term effects of statelessness in Ottoman Syria by analysing the region divided by the "desert line", a boundary that separated Ottoman-controlled areas from those subject to tribal raids. Using a spatial regression discontinuity design, I estimate the impact of historical statelessness on contemporary economic outcomes. The results show that historically stateless areas have lower incomes, less developed human and physical capital, and a higher share of workers in the primary sector. These effects persist primarily due to lower population density, which explains a significant portion of the economic gap. This persistence reflects path-dependent consequences: emigration and low population density, initiated by statelessness, continue to hinder economic development in these areas today. The results are robust to geographic controls and testing for competing hypotheses including the roles of institutions and culture.

1. Introduction

A large and growing literature examines the various factors influencing economic development. Recent studies highlight the significance of institutional persistence as a key determinant of economic development (Acemoglu et al. 2001, Michalopoulos and Papaioannou 2013, Dell et al. 2018). There is also a growing focus on the role of path dependence in sustaining regional inequality, emphasizing how historical shocks and initial conditions influence economic outcomes even centuries later (Krugman 1992, Bleakley and Lin 2012, Michaels and Rauch 2018, Dalgaard et al. 2022, Baerlocher et al. 2024). Together, this research highlights the need to examine how historical legacies, particularly those tied to state formation and governance, continue to shape patterns of development today. At the same time, no studies have isolated the causal effect of the absence of state control itself, disentangling it from geography, culture, or endogenous processes of institutional evolution.

This paper fills this gap by examining the long-run economic consequences of historical statelessness in Ottoman Syria. Specifically, I estimate the causal impact of statelessness on modern economic outcomes using a spatial regression discontinuity design (RDD) that compares communities located on either side of the "desert line"—a historical boundary that delimited the eastern edge of effective Ottoman control until the mid-19th century.

The desert line which separated Ottoman Syria from the stateless areas to the east did not denote the actual extent of the desert, but rather the area where land was largely abandoned due to Bedouin raids and lack of security and rule of law. The term was used by European explorers and cartographers who travelled across Syria in the 18th and 19th centuries (Lewis 1987). Areas west of the border line were under Ottoman state control and formed the economic and demographic centre of Syria, whilst areas east of the border only came under state control in the middle of the 19th century. The expansion of the Ottoman state in this period came as part of a larger shift towards the east precipitated by the loss of Ottoman territories in Europe and the Caucasus, and as part of a broader set of reforms that sought to centralize and modernize the Ottoman state.

The credibility of the causal design rests on two key assumptions: First, the location of the desert line, as reconstructed from 18th- and 19th-century explorer accounts (Lewis 1987), is plausibly exogenous to later economic development - it reflected the limits of Ottoman security provision and military reach, not any systematic economic or cultural cleavage. Second, in the immediate vicinity of the line, geographic conditions - including rainfall, elevation, and agricultural suitability - vary smoothly. This supports the continuity assumption necessary for RDD identification: that any discontinuity in outcomes at the boundary can be attributed to the treatment of historical statelessness, rather than to underlying spatial trends.

Guided by this design, I use village-level measures of contemporary income, human and physical capital, and sectoral employment from the 2004 Syrian Population and Housing Census, as well as data on the geographic and climatic features of these villages which support the identifying assumption of the econometric design. The analysis results show large and persistent economic disadvantages in historically stateless areas: lower incomes, lower levels of education and infrastructure, and a greater reliance on agriculture. These differences remain robust when controlling for fine-grained geographic variables, spatial trends, and alternative specifications including spatial autoregressive lag models.

To assess mechanisms, I test competing hypotheses related to institutional persistence and cultural differences. While historically stateless regions display somewhat higher informality in property ownership, this does not explain the observed economic gaps. Nor do differences in ethnic and religious composition. Instead, the results point to path dependence in settlement patterns as the dominant mechanism: population density explains a substantial share of the differences in income and human capital, consistent with theories of agglomeration and self-reinforcing economic geography (Krugman 1992; Bleakley and Lin 2012).

These findings advance the literature in several respects. First, while prior work has emphasized variation in the strength or centralization of historical states (Michalopoulos and Papaioannou 2013, Dell et al. 2018), this study isolates the effect of prolonged statelessness relative to adjacent governed areas. Secondly, by demonstrating that lower historical state capacity translated into persistent underdevelopment primarily through the channels of demography and economic geography, the paper underscores the importance of considering path-dependent processes alongside institutional legacies, further contributing the literature on the roles of population agglomeration and path dependence in shaping economic development (Boserup 1965, Krugman 1992, Bleakley and Lin 2012, Oto-Peralías 2020). The contribution is part of the broader debates on the persistence of historical institutions (Nunn, 2020) and the role of state capacity in long-term economic growth (North et al., 2009). The paper also relates to the literature on agricultural-pastoral frontier zones (Bai and Kung 2011, McGuirk and Nunn 2024). Whereas previous literature has focused on the climactic drivers of conflict in similar settings, this paper contributes to this line of research by showing the long-term implications of statelessness in a frontier zone.

2. Historical background

2.1. The Desert Line

Following the Ottoman conquest of Syria in 1516, state authority remained spatially uneven. Ottoman control was strongest in the western provinces stretching from

Aleppo to Damascus, which served as administrative and economic hubs. By contrast, the steppe regions to the east experienced only intermittent oversight, as their open terrain and the mobility of Bedouin tribes limited Ottoman reach (Ma'oz 2013).

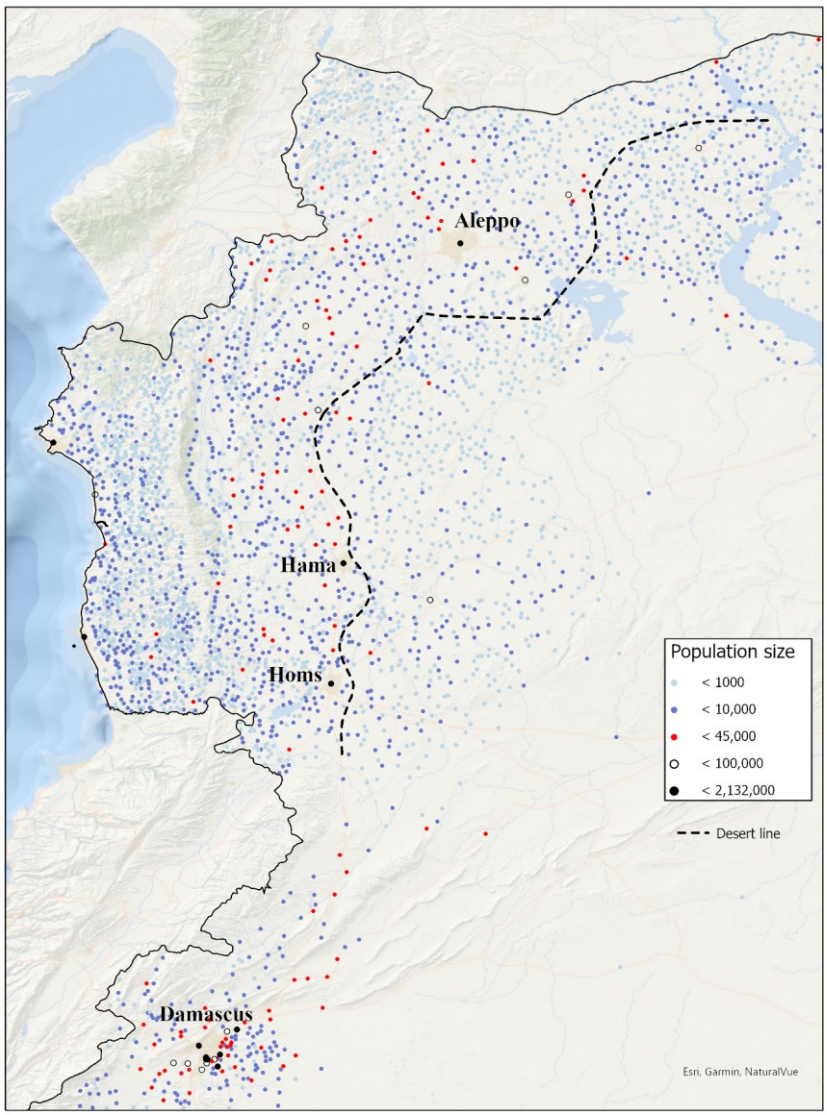
The boundary termed the 'desert line' marked the effective limit of Ottoman protection and taxation before the mid-19th century. European travellers and cartographers documented this zone with unusual precision. Carsten Niebuhr's 18th-century expeditions produced systematic surveys of the area, documenting village abandonment (Hopkins 1967; Lewis 1987). Niebuhr and his successors consistently observed that villages east of this line were abandoned due to tribal raids, whereas settlements just to the west were subject to regular Ottoman administration.

Importantly, this line did not reflect climatic desert boundaries. Rather, it emerged as a frontier shaped by the state's military capacity and the incentives to project power over taxable agricultural production. This historical contingency is critical: the desert line was a product of Ottoman state-building constraints, rather than of any natural division in soil quality or rainfall. This supports the assumption that, in the absence of Ottoman expansion, settlement conditions on either side would have been comparable.

The region that surrounds the boundary is a transitional zone between the steppe or semi-desert of the interior and the well-watered lands towards the coast. Lewis (1987) describes the transitional zone in the 19th century as a 'debatable area between the steppe and the settled farming country of western Syria, and often a zone of contention between nomads and peasants'. The Syrian desert, located to east of this zone, was the heartland of nomadic tribes. In the present day, the majority of the transitional zone is under cultivation, with seasonal grazing by herders during the summer, though settled villagers constitute the dominant population. Describing the relationship between the state and economic development in the area after the middle of the 19th century, Lewis states that 'as the state extended and strengthened its hold on the countryside, the economy of the country developed, peasants and landlords moved into the transitional zone, and the nomads gradually changed their way of life.'

Figure 1 shows the desert line as mapped by Lewis (1987). Areas south and east of the line were outside of state control until the middle of the 19th century. Lewis's study places the desert line in the middle of the transitional zone.

Figure 1: Map of the desert line in the west of Syria with surrounding cities, towns, and villages in 2004 according to population size



Source: Author's own map based on Lewis (1987) and the Syrian Population Census of 2004. Background ESRI OpenStreetMap.

Carsten Niebuhr first used the expression 'desert line' in his 1774 book which records his travels through Syria and Palestine as part of the 'Danish Arabia expedition'. Lewis uses the account of Niebuhr and other explorers of the era to

pinpoint the location of the line from their recorded observations. The observations were systematic enough to allow Lewis to create a map of the line, which is consistent with Niebuhr himself being a trained cartographer and mathematician, and who created some of the first precise maps in parts of the Middle East using modern survey methods (Hopkins 1967). The traveller accounts used by Lewis were unanimous in singling out the tribes, including specific tribes like the Mawali, for depredations and exactions which resulted in deserted farmlands and villages (Lewis 1987, pp. 8). The Ottoman state also received blame for the poor state of public security, and for the exactions of tax collectors. Earlier accounts of the conditions in Ottoman Syria suggest similar dynamics, with the Ottoman state constantly faced with tribal incursions as recorded in the Ottoman provincial yearbooks *(Salname)* as early as the 1500's (Bayat 2017).

The desert line closely aligns with the location of the major urban centres of Aleppo, Hama, and Homs. These cities, with Damascus further south, form the political and demographic centre of Syria. Cities in this region were highly fortified to defend against raids and invasions, including those from the nearby desert. It is likely that to location of these cities was an important determining factor that shaped the location and extent of state authority in the region. The line also follows the location of the strategically important 'sultanic road' which connected the Syrian cities together with the rest of the empire. Areas west of the line are bounded in the west and north by mountain ranges, which makes holding this territory easier as compared to areas east of the desert line, which are open in the east and south to the desert steppe. This fits with the circumscription theory of state formation (Carneiro 1970), which posits that the emergence and persistence of centralized political authority is more likely in ecologically or geographically circumscribed areas where populations cannot easily disperse. In areas west of the line, the combination of defensible mountain terrain to the west and the presence of walled cities to the east made the area well circumscribed, in contrast to areas east of the line. Therefore, areas to the west were easier for the Ottoman state to control and offered greater rewards than the lands to the east. This cost-benefit calculation may have influenced the Ottoman policy of leaving the lands east of the desert line under tribal control until the middle of the 19th century.

The shifting frontier of settlement and the challenge of defending it against incursions from the steppe have been recurring themes in the region's history. In the 2000's, an archaeological team discovered the remains of a defensive wall, termed the 'Very Long Wall' (Geyer et al. 2010). The wall was built in the late third millennium BC, extends for 220 km in the area east of the transitional zone, situated east of the desert line and reaching well into the steppe region. The Amorite kings of Hama likely built it to delimit their territory and protect agricultural development against roaming nomads. A later fortified network of structures was also found in the same area dating to the Middle Bronze Age, which is associated with a recession of the frontier of settlement towards the west (Rousset et al. 2020). During antiquity,

the frontier of settlement in the transitional zone moved significantly across time, reaching its furthest extent towards the east under Byzantine rule in the 5th and 6th centuries (Geyer 2011), as a part of a cycle of intensification and abatement common to agricultural-pastoral frontier regions in the MENA (Wickham 2006, p. 19). Similar to the Ottomans but with more success, Romans and Byzantines relied on sedentarized nomads to rule inland Syria, with the city of Palmyra emerging as an important seat of power in late Antiquity (Liebeschuetz 2015).

The cities and towns on the Syrian Euphrates, which were ruined and abandoned by the time of Ottoman rule, had prospered from Antiquity well into the Middle Ages, and continued to play a significant role under the Islamic Caliphates. However, the entire area of present-day Syria was devastated by the Mongol invasions in the middle of the 13th century. In the present, the only cities in the transitional zone east of the desert line are Manbij and Salamiyah, both minor cities re-established by migrant groups in the 19th century (Circassians and Ismailis respectively) with support from the Ottoman state.

While Lewis (1987) framed his reconstruction of the desert line as a study of settlement abandonment, the historical sources he used make clear that this pattern was a direct manifestation of the operational reach of Ottoman authority. In Ottoman Syria, the viability of sedentary agricultural settlements depended on the state's ability to provide security from raids. Contemporary traveller accounts, Ottoman provincial reports, and local narratives repeatedly identify the same causal link: settlements east of the line were abandoned not due to environmental unsuitability, but because they lay beyond the effective protection of the state; those to the west persisted precisely where garrisons, tax collection, and administrative oversight were continuous. This tight coupling of settlement persistence and governance capacity means that the mapped frontier of abandonment is also the frontier of state presence. Lewis's line is therefore not simply a demographic boundary, but a historically grounded, spatially precise measure of the limits of Ottoman rule in the pre-reform period.

2.2. Geographic continuity across the desert line

A key requirement for the regression discontinuity design is that geography changes smoothly across the desert line. This condition is well supported by the region's physical geography. The study area forms a continuous whole, bounded in the west by the coastal mountain range and opening eastward toward the desert. It is one of Syria's main agricultural zones and contains three of the country's four largest cities (Aleppo, Homs, and Hama) highlighting its agricultural potential. Although soil fertility and rainfall decline gradually from west to east, there is no sharp environmental break along the desert line.

Settlements immediately to the west and east of the line are similarly close to major urban centres. The strategic Sultanic Road, which linked these cities to the wider empire, ran parallel to the frontier, ensuring trade access on both sides. With no major natural barriers at the boundary, and in the absence of differences in security or governance, communities near the line would have enjoyed broadly similar prospects for agricultural development and market integration.

Travellers that documented the abandonment of villages in stateless areas were aware of the agricultural potential of these abandoned areas (Robinson and Smith 1856; Lewis 1987). For example, the district of Salamiyah was described by Robinson and Smith (1856) as 'entirely deserted' even though it was described by locals as 'exceeding even the neighbourhood of Homs and Hama, in the fertility of its soil.' Thus, the observed spatial discontinuity in later development outcomes cannot be plausibly attributed to differences in geographic features across the border line.

To ensure that continuity holds in geographic variables across the border line, Table 1 presents balance checks for the two sides of the border line using a set of geographic variables.

Table 1: Balance checks

	Dependent variable:						
	Slope	Elevation	Temperature	Precipitation	Flow Accum.	Wheat suitability	Cotton suitability
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Unadjusted mean values							
Stateless	1.568	436.949	17.895	123.184	78.237	8991.516	8266.816
Control	3.353	433.248	17.202	206.730	271.769	8423.155	7728.057
Difference	-1.785	3.701	0.694	-83.546	-193.532	568.361	538.759
Panel B: Regression estimates (cubic polynomial)							
Stateless	0.718**	-38.759***	0.125***	4.127**	139.573	-256.537	357.065
	(0.300)	(10.754)	(0.045)	(1.795)	(233.060)	(198.060)	(278.609)
Observations	1,908	1,908	1,907	1,907	1,908	1,908	1,908

*p<0.1; **p<0.05; ***p<0.01

Note: Panel A shows unadjusted mean values for the geographic variables across the stateless and control areas and the difference between them, Panel B shows the RD cutoff effect using a cubic polynomial in latitude and longitude for each variable.

The unadjusted mean values in Panel A show that the stateless and control areas are similar on measures like slope (ruggedness), elevation, temperature, and wheat and cotton suitability, but diverge in measures of precipitation and flow accumulation. This can be explained by the proximity of stateless areas to the desert, with reduced

precipitation levels in the east. Panel B shows the coefficients on the stateless dummy variable in the regression of each geographic variable on the RDD model detailed in the methods section, with a cubic polynomial in longitude and latitude. The resulting coefficient show that the differences in precipitation and flow accumulation are not discontinuous at the border line. The coefficient is small and positive for the precipitation variable, and statistically insignificant for flow accumulation. The difference in temperature is also reduced to a very low level of one tenth of a degree Celsius. The magnitude of these effects suggests that they do not represent a threat to identification. Additionally, the analysis below shows that adding these geographic variables as controls in the main econometric model does not shift the effect sizes significantly.

2.3.Ottoman state expansion and reform

The abandonment of villages in the stateless areas of Ottoman Syria fit into a broader pattern of de-population and demographic decline. Population numbers in Syria had reached their peak under the early Islamic empires, and despite growing during Ottoman rule, vast regions of the empire remained well below the population peaks reached in prior eras. Contributing factors included the black death and other plague epidemics and natural disasters, the declining role of Mediterranean trade in favour of the Atlantic trade, as well as the Bedouin raids and rural insecurity. According to Williams (1981), rural depopulation was a chronic problem from the turn of the 17th century until the middle of the 19th century, and it figured prominently in Ottoman government reports, as well as in the writings of European travellers and consuls (Lewis 1987, Williams 1981). Walker (2012) refers to a ‘nomadization’ of the rural hinterland and the disappearance of villages in the southern Levant from the start of the Ottoman era.

In response to this problem, the Ottoman state implemented various measures to stabilize rural populations. One strategy involved replacing tax farming tenancies with life leases, giving tax farmers a long-term stake in the well-being of peasants. Additionally, the government undertook settlement projects across Anatolia and Syria to bring more land into cultivation and expand the tax base. These projects took various forms, including land grants to incoming refugees from the Balkans and the Caucasus and the forced sedentarization of nomadic populations, particularly in eastern Syria. Ma'oz (1968, Chapter 9) describes this settlement activity as sporadic, although efforts were more sustained in the Aleppo region, where authorities incentivized settlement by granting land, seeds, tools, and tax exemptions. However, these efforts were often undermined by the government's failure to protect newly sedentary populations from nomadic raids.

On the eastern fringes of the Syrian provinces, The Ottoman state allowed the Bedouin tribes considerable autonomy in managing their affairs. The relationship between the Bedouins and the Ottoman state was a complex one, with the Ottomans

attempting to co-opt the Bedouins by appointing a Bedouin governor (Emir of Badia) who was at the head of a tribal confederation and was tasked with keeping order in the frontier areas and protecting trade routes. Still, regions close to the desert were subject to constant tribal raiding. Ma'oz (1968, Chapter 9) suggests that the Ottoman co-optation policy was ineffective, with Bedouin tribes in government employment frequently attacking villages and trade caravans. The populations of stateless regions paid protection tax to the tribes (Khuwah), yet the tribes did not ensure order and rule of law. According to Ma'oz (1968), the protection duty paid out to a dominant tribe rarely secured a village from extortion or attack from other tribes, or even from the 'protector' tribe. Villages outside of state control were also required to pay tax to Ottoman officials, representing a double burden of taxation with little security in exchange. The Ottoman state did not maintain any presence deep in the Syrian desert, which formed the centre of power for the largest tribes. Migration of tribes from the Arabian Peninsula in the Syrian desert also led to conflict over grazing rights and may have pushed some tribes towards raiding.

Travelers in the region south-east of the desert line in the 18th and 19th centuries report that many villages were abandoned to the nomads, while the remains of medieval towns and cities were deserted and in ruins (Lewis 1987). Comparatively few villages remained inhabited on the south and east side of the desert line as compared to those lying on the western side, despite the abundance of cultivable land and their proximity to the urban centres of Aleppo, Hama, and Homs.

The first significant steps undertaken by the Ottoman state to extend its control over the east of Syria came in the late 1850's and early 1860's with the creation of new administrative districts and the building of military outposts in the Euphrates valley east of Aleppo, followed by the establishment of a garrison at Deir ez-Zor. Arab tribes occupying these areas were subdued in the process and agreed to pay tax to the Ottoman government, while the protection tax paid by peasants to the tribes was abolished. By the 1870's the Euphrates valley tribes were 'completely under control' and paid taxes regularly (Lewis 1987). The nomadic tribes settled new villages and expanded existing villages, towns, and cities throughout the rest of the 19th century and into the 20th century.

The Provincial Reform (Vilayet) law of 1864 was a crucial part of the Tanzimat reforms which re-defined the central government's authority over the provinces, with Syria one of the first regions to be re-organized (Rogan 1995). The aim of the law was to centralize and modernize the administrative system, enhancing the state's control over its territory. The regions controlled by Bedouin tribes were not the only ones affected by this development, with the Alawi coastal region and the Druze areas in the south experiencing similar centralization. Certain social classes such as urban elites and merchants played a key role in extending state authority in this period according to Rogan (1995), with the state providing favourable conditions for trade and economic activity. Barakat (2015) shows that tribal leaders as well as middling groups of nomads obtained positions as low-level bureaucrats by assisting

in the modernization of land administration during the Hamidian period (1876-1909).

The Ottoman state modernization and reforms of the 19th century were likely due to outside pressures on the empire. The loss of Ottoman territory in Europe, the modernization efforts of Muhammad Ali of Egypt and his occupation of Syria in the 1930's, as well as western pressures for governmental reform and the strong pull of the western modernization. State centralization in the Syrian provinces also came with reforms that promoted legal equality for religious minorities and sought to reduce the power of local elites as well as the influence of European powers within Ottoman lands. The reforms were not particular to the study area, and were largely exogenous to its internal dynamics, even if their implementation had profound implications for the region.

2.4. Historical parallels in the MENA region

Parallels to the Ottoman state centralization in Syria can be found across the MENA region. Several studies of the late Ottoman empire focus on frontier regions including Trans-Jordan, Kurdistan, the Persian Gulf, Eastern Arabia, and Yemen (Reinkowski 2001). Recent historical work highlights the history of governance in the Ottoman peripheries, as in the case of the Kurdish principalities (Özok-Gündoğan 2014) and the inter-imperial borderlands of the Balkans (Esmer 2014). While most countries in the region came under colonial control by the early 20th century, the pre-colonial era was characterized by diversity in terms of state centralization. In many cases, the pre-colonial empires, kingdoms, and sheikhdoms exercised weak control over their territories, especially in marginal areas and frontier regions. Scholars recognized this division in studies of pre-colonial North Africa (Hoffman 1967). The terms Bled Al-Makhzan (land of the treasury[1]) and Bled Al-Siba (land of anarchy) were used to denote the land held under government control and those which remained outside direct government rule respectively. Geography clearly shapes the Makhzan and the Siba in the setting of North-African, with the coastal areas and cities situated near the mediterranean coast being the centres of state control, while mountainous and desert regions further south were the centres of tribal presence and control. An earlier parallel is that of the border region between Byzantines and Seljuks in Anatolia which was occupied by Turkic tribes, and which experienced depopulation due to tribal raiding on peasant communities according to Lindner (2017).

A pattern of state formation in the Middle East region involves tribes on the margins of imperial states. This is most clearly seen in the case of Saudi Arabia and other

[1] The Arabic word makhzan, meaning "warehouse", came to signify the state in western North Africa. This usage likely stems from the term's historical association with the treasury and the presence of state tax collectors in regions under government authority.

states in the Gulf region, as well as some early modern examples like the tribal Kurdish confederacies in eastern Anatolia and the tribal confederacies of North Africa (Khoury and Kostiner 1990), and even the Ottoman state itself which originated in a tribal setting in Anatolia (Lindner 2017). In the Syrian case, the Kurdish and Arab tribes that settled the north-east of Syria were initially operating outside the bounds of the state but were eventually integrated into Ottoman state structures in the mid-19th century, with the aim of safeguarding the frontier region and creating sources of tax revenue. The Ottoman state supported the settlement and sedenterization of tribes in the north-east and even assigned a tribal leader as governor of Raqqa (Winter 2006), a pattern that followed in south-eastern Anatolia as well. Early on, the relationship between the tribes and the state was marked by conflict, with tribal authorities often undermining the Ottoman rule of law. But the tribes in the Syrian north-east did not create alternative state structures as was the case in the Arabian Peninsula. Proximity to the well-established Ottoman state in the west of Syria may have played a role, with little opportunity for the tribes to control urban areas and establish separate governance structures, as in the case of the Arabian Peninsula.

3. Historical statelessness and path dependence

An established line of research examines the role of pre-colonial state history and institutions in shaping present-day economic inequalities. Studies typically measure historical state centralization and find strong correlations with contemporary economic outcomes across and within countries. Gennaioli and Rainer (2007) show that pre-colonial state centralization in Africa is associated with better public goods provision today, arguing that hierarchical chiefdom structures foster greater accountability to state authorities and more effective policy implementation. Michalopoulos and Papaioannou (2013) extend this approach using ethnic group-level data, demonstrating a robust relationship between historical state centralization and economic development, proxied by nightlight density. Dell et al. (2018) further investigate this effect in Vietnam, exploiting a historical boundary between the centralized Dai Viet kingdom and the less centralized southern region under Khmer influence. Their findings suggest that state centralization is linked to improved economic outcomes through the persistence of village governance institutions. Another set of studies use macro-level data and analyse the effect of time under state institutions and find that it correlates with a range of contemporary economic and political measures (Bockstette et al. 2002, Chanda and Putterman 2007, Borcan et al. 2018).

Recent work has also explored how historical state centralization interacts with geography, conflict, and economic persistence. Oto-Peralías (2020) examines how historical frontier warfare in Spain influenced state formation and long-run

economic geography. The author argues that insecure frontier regions developed weaker state institutions and lower economic activity due to persistent violence and instability, shaping regional economic patterns into the present. This perspective highlights how security conditions influence the long-term concentration of people and economic activity, a theme that intersects with the literature on path dependence and economic geography.

Path dependence plays a central role in explaining regional economic disparities. Initial advantages in geographic conditions, resource endowments, or political stability can generate self-reinforcing agglomeration effects, whereby economic activity and population remain concentrated over long periods. Krugman (1992) illustrates this by noting that one-third of the U.S. population still resides within the original thirteen colonies, despite centuries of economic shifts. Similarly, Bleakley and Lin (2012) show that historical portage sites in the U.S. continue to influence population density, even after their original economic function has become obsolete. This persistence reflects the role of increasing returns to scale, which can sustain economic activity long after initial advantages have disappeared. Baerlocher et al. (2024) provide further evidence from Brazil, demonstrating how Portuguese-era road networks, originally constructed to serve gold mines, continue to shape contemporary economic density.

The effect of historical state centralization on economic geography is also linked to the role of population density in development. Densely populated areas are more likely to urbanize, facilitating specialization, investment, and technological advancement. Boserup (1965) argued that agricultural productivity is shaped by population density, while Kremer (1993) posited that larger populations foster faster technological growth. Klasen and Nestmann (2006) further show that population density spurs technological change, particularly in low-technology environments. High-density regions also benefit from greater returns to public goods investments, reinforcing development advantages over time.

Based on the theoretical and empirical literature reviewed above, I expect to find that historically stateless regions are less economically developed today than areas with a longer history of statehood. This effect is anticipated to persist due to path-dependent mechanisms, particularly those related to population density. Insecure frontier zones and areas with weak historical governance likely experienced lower long-run settlement density, limiting the emergence of agglomeration economies, specialization, and public goods provision. Accordingly, I expect the long-run developmental disadvantage of stateless areas to correlate more strongly with differences in population density than with cultural or institutional persistence. This would suggest that the legacy of historical statelessness operates primarily through its influence on spatial population patterns, rather than through the direct transmission of institutional norms or cultural traits.

4. Data and methods

The data used in this study comes from the Syrian population and housing census of 2004. The census data records a rich array of information on the population, labour force, and infrastructure and housing, aggregated at the level of the city, town, and village. Information on the historical boundary of the desert line comes from the work of Norman Lewis (1987) which uses Ottoman government reports and traveller accounts to pinpoint the boundaries of state control prior to the mid-19th century. In this context, settlement persistence was inseparable from state presence, making Lewis's (1987) mapped line a direct proxy for the historical boundary between governed and stateless areas. Because the line reflects a security frontier rather than a natural or cultural divide, its location is plausibly exogenous to later economic outcomes and satisfies the continuity assumption required for the regression discontinuity design.

The analysis exploits the discontinuous change in exposure to historical state control, comparing towns and villages in areas outside the control of the Ottoman state in the period before the middle of the 19th century with those that were incorporated previously when the Ottomans invaded Syria in 1516. I use the methods employed in Dell (2010) and Dell et al. (2018) and treat the desert line as a two-dimensional discontinuity in longitude-latitude space. The regression model takes the form:

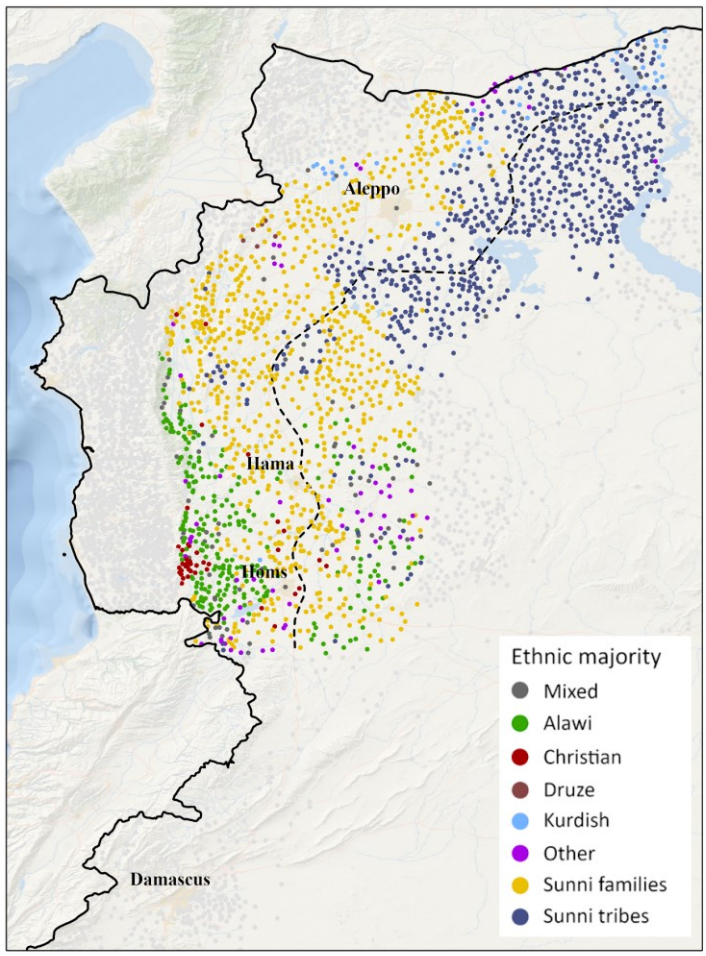
$$outcome_i = \alpha + \gamma\, Statless_i + f(geographic\ location_i) + \beta\, to\_Aleppo_i + \epsilon_i$$

Where $outcome_i$ is the outcome variable in village i, and $Stateless_i$ is an indicator equal to 1 if village i was on the south-east side of the desert line (i.e. outside of Ottoman state control before the middle of the 19th century) and equal to zero otherwise. $f(geographic\ location_i)$ is the RD polynomial which controls for smooth functions of the geographic location. to_Aleppo_i is the log distance to the centre of Aleppo city, which is the largest city in the region and an important trading centre. The analytical sample is limited to towns and villages within 50 kilometres of the desert line and excludes towns or cities with populations over 10,000. The reason for excluding major urban centres is that the location of the line itself is likely endogenous to the location of these centres, but towns and villages can be treated as if randomly distributed across the line, with towns and villages on either side of the border being in close proximity to the major cities and having similar geography as will be shown in the next section.

The estimation framework relies on the assumption that all relevant factors besides treatment vary smoothly at the boundary. Geographic characteristics may vary across the two sides of the desert line, such as precipitation, soil quality, or ruggedness of the terrain. The variation in outcomes that is due to these factors will be captured by the included geographic function.

Figure 2 shows a map of the study area according to ethnic group. The analysis controls for ethnic majority, as the estimation framework is also sensitive to selective sorting across the boundary.

Figure 2: Map of the study area with ethnic majorities



Source: Author's own map based on ethnic majority data from Khaddour and Mazur (2018). Background ESRI OpenStreetMap.

The regression model controls for ethnic composition based on the ethnic majority of each village or town. The ethnicity data comes from the work of Khaddour and Mazur (2018), where the ethnicity variable captures ethnic identity, religious and sectarian identity, as well as belonging to an Arab tribe. Controlling for tribal belonging is particularly relevant, as areas east of the desert line were settled by

Arab tribes whose culture may differ from that of the population west of the desert line and the analysis accounts for this fact.

Another threat to identification is due to spatial correlation in the error terms, which can lead to inflated t-values as detailed in Kelly (2019). To address this issue, I estimate spatial autoregressive lag models that relax the assumption of no correlation in the error terms.

To estimate gaps in incomes across areas, an income index is created using data on the local composition of the labour force in each village and town. Incomes are imputed based on the national-level average wages are used which are conditional on occupation, along with the local mix of sectors. The wage data comes from Syrian Labour Force Survey of 2007. The sectoral income variable reflects within-country differences in the composition of the economy but does not account for other sources of income inequality such as differentials in payoff to the same work across areas. The imputation formula takes the form:

$$Income_i = \sum_{j=1}^{7} Sector\ wage_j\ .\ Sector\ share_{ij}$$

Where i refers to the city, town, or village and j to the sector. Sector wage refers to the national-level average wage of sector j, and Sector share is the share of workers in the city/town/village employed in that sector. The set of sectors used to construct the sectoral income measure are agriculture, industry, construction, hotels and restaurants, transportation, finance and real estate, and other services. An alternative specification of the income measure is used in addition. Income is imputed in this case using educational attainment measures and corresponding average wages conditional on educational attainment using the following formula:

$$Income_i = \sum_{j=1}^{7} Mean\ wage_j\ .\ Educational\ attainment\ share_{ij}$$

Where the educational attainment share refers to the share of the adult population in attainment categories: illiterate, literate, elementary school, middle school, high school, middle academy, and university educated. Mean wage refers to the national-level average wage conditional on the educational attainment category. Both imputed incomes are normalized as z-scores using population-weighted mean and standard deviation across all cities, towns, and villages in Syria. The resulting measures represent deviations from population-weighted mean income.

This imputed income measures should be interpreted an index of expected local average earnings given the composition of the workforce, rather than observed local wages. By applying national-level average wages (by sector or educational attainment) to the local distribution of employment or education, the measure

captures within-country differences in income potential arising from variation in economic structure and human capital. It does not reflect local wage differentials for the same occupation or education level, nor other sources of income such as informal work, self-employment, remittances, or returns from local resources. As such, the measure is best understood as a proxy for structural income potential associated with the local labour force, abstracting from location-specific pay differences and cost-of-living variation.

In addition to imputed income, the analysis makes use of other data that proxy local economic development. Measures of the share of households connected to each of the electricity network, sanitation network, and water network are used, as well as measures that directly capture different development in the local labour force, including the share of white-collar workers (managers) and the share of workers in agriculture. Finally, direct measures of human capital are included which are the share of adults with a university education and the illiteracy rate in the adult population. Additional data is used in the analysis, such as geographic data that capture information on the climate, agricultural suitability, and land ruggedness. The data available in the census also allows for addressing mechanisms. The analysis makes use of data on informality in property ownership to examine whether private property institutions play a role in explaining the economic gaps across the desert line.

5. Results

5.1. Descriptive statistics

Table 2 shows descriptive statistics for the main variables used in the analysis. The income z-scores measure at -1.26 and -0.80 for the stateless and Ottoman areas respectively. Villages and towns in both areas are below the Syria national average in terms of income, though areas with a history of statelessness stand out. The alternative measure of income imputed using education levels gives slightly lower values at -1.20 and -0.72 respectively. The historically stateless region has worse infrastructure overall. Fewer households are connected to the sanitation network (17.8%) as compared to those in the Ottoman state areas (35.3%). A similar gap is found in the share of households connected to the freshwater network (39.7% and 65.8% respectively). The only exception is the proportion of households connected to the electricity network, with both sides of the border line reaching over 90% and with a small gap of less than 3%.

Table 2: Descriptive statistics

	Stateless	Ottoman state	Difference
Income index	-1.26	-0.80	-0.46
Education-based income index	-1.20	-0.72	-0.48
Electricity network	90.6	93.4	-2.88
Sanitation network	17.8	35.3	-17.43
Water network	39.7	65.8	-26.12
Managers	3.5	7.7	-4.19
Farmers	39.1	43.4	4.28
University educated	0.6	1.2	-0.59
Illiteracy rate	34.2	25.0	9.21
Population	1,030	1,644	-614
Population within 5 km	5,668	13,322	-7,654
Population within 10 km	30,797	75,512	-44,715
Observations	797	1111	

Note: The table shows mean values of outcomes used in the analysis according to the placement of each town or village in relation to the Ottoman state border line. Income indices are standardized z-scores. Other outcomes are in percentages (except for population sizes).

Areas with a stateless legacy also have worse labour market outcomes and lower levels of education. The proportion of workers in managerial positions is at 3.5% and 7.7% in stateless and Ottoman areas respectively, while the proportion of farmers is at 39.1% and 43.4% respectively. Education levels are also lower in areas with a stateless history, with the proportion of university educated adults at 0.6% and 1.2% in areas with historical state control and areas under Ottoman control, while the illiteracy rates are at 34.2% and 25% respectively.

Areas with a stateless legacy are less urbanized, with lower mean population in the towns and villages as compared to areas with a legacy of Ottoman state control, and a lower number of towns and villages. The overall gap is best captured by the measures of population within 5km and 10km radius, as these measures include the population size of the village or town itself, as well as any villages or towns centred within the radius. According to this measure, the population density is much higher in areas with a legacy of Ottoman state control, reaching 75,512 within a 10 km radius as compared with 30,797 for towns and villages with a stateless legacy.

5.2. Regression results

Table 3 shows the results for the main regression discontinuity models (full model estimates are available in appendix tables A3 to A6).

Table 3: Regression discontinuity effects

	Dependent variable:								
	Income index	Education-based income index	Electricity network	Sanitation network	Water network	Managers	Farmers	University educated	Illiteracy rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Linear model									
Stateless	-0.532***	-0.438***	-3.410**	-21.942***	-22.491***	-5.010***	7.098***	-0.530***	9.216***
	(0.061)	(0.037)	(1.326)	(2.047)	(2.529)	(0.407)	(1.853)	(0.075)	(1.039)
Panel B: Cubic polynomial model									
Stateless	-0.577***	-0.431***	-7.754***	-2.006	-27.177***	-3.648***	10.098***	-0.471***	13.676***
	(0.113)	(0.066)	(2.659)	(4.073)	(4.289)	(0.797)	(3.446)	(0.154)	(1.956)
Panel C: Added geographic controls									
Stateless	-0.500***	-0.409***	-7.338***	-0.903	-26.805***	-3.589***	7.743**	-0.421***	13.541***
	(0.112)	(0.067)	(2.673)	(4.080)	(4.313)	(0.806)	(3.379)	(0.157)	(1.959)
Panel D: Added ethnicity controls									
Stateless	-0.645***	-0.523***	-8.703***	-3.232	-28.155***	-4.279***	11.210***	-0.653***	14.533***
	(0.111)	(0.062)	(2.729)	(4.122)	(4.289)	(0.735)	(3.448)	(0.145)	(1.977)

*p<0.1; **p<0.05; ***p<0.01

Note: The table shows the RD cutoff effect for different model specifications and outcomes. Standard errors clustered by district in parentheses (92 clusters). All regressions control for log distance to Aleppo. The basic linear model includes controls for latitude and longitude on each side of the regression discontinuity line. The cubic polynomial model includes controls for latitude and longitude on each side of the discontinuity of the form $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$ where x and y denote longitude and latitude. The third and fourth rows include the cubic polynomial model and controls for geographic variables (elevation, slope, temperature, precipitation, and flow accumulation) and for ethnic majority composition respectively. Full regression outputs are included in Appendix tables A3 to A6.

The cells show the RD effect of being located in the region with a stateless legacy, for different outcome measures in the different columns. The panels represent the different model specifications. The model in Panel A is linear in longitude and latitude, and the effects are all negative and statistically significant (except for the illiteracy rate, which is higher in the area with a stateless legacy). For the income index, the gap is at -0.532 while the gap in education-based income is estimated at -0.438. The infrastructure effects vary, with both the sanitation network and the freshwater network being significantly less developed in the south-east (effect sizes of -21.8 and -22.5 respectively), while the effect for the electricity network is much smaller at -3.4 and significant only at the 0.05 level. When looking at the labour force variables, the proportion of managers in the south-east is lower by 0.5, while the proportion employed in agriculture is 7 percentage points lower. The proportion of adults with a university education is 0.53 percentage points lower, and the illiteracy rate is 9.2 percentage points higher. Overall, the estimated effects resemble the unadjusted gaps presented in the descriptive statistics table earlier, which suggests that these gaps are mainly driven by the desert line discontinuity and are not primarily due to geographic trends across the study region.

Panel B of Table 3 provides the effect estimates for the model controlling for cubic polynomial terms in longitude and latitude on each side of the border line. The effects sizes are similar level as in the previous model, except for the gap in the sanitation network which becomes small and statistically insignificant. The robustness of the results to the cubic polynomial controls suggest that they are not driven by geographic trends. In some cases, the effects are even larger in this model, as in the income index, the electricity and freshwater networks, and the share of farmers in the labour force as well as the illiteracy rate.[2]

To further ensure that geographic factors are not the drivers behind the results, Panel C adds controls for elevation, slope, temperature, precipitation, and flow accumulation. In this model, the effect sizes are also on a similar level to the above model and are equally significant.

Controlling for ethnic groups in the fourth row in addition to the cubic polynomial model does not shift the effect estimates much either. The income effects appear slightly larger at -0.65 and -0.52 for the income index and education-based income respectively, and the other effects are slightly higher each but remain broadly similar

[2] The results for these base models are replicated using 25km bands around the border line with similar results (see Online Appendix Table A1). Effect size sensitivity to bandwidth changes is shown in Appendix Figure A1 using a cubic polynomial model, and bandwidth sensitivity with a 5km donut around the desert line is shown in Appendix Figure A2. Appendix Table A2 shows the same results adding line segment fixed effects. The effect sizes are overall robust and become larger when adding line segment fixed effects.

and statistically significant. The role of ethnicity is explored further in the next sub-section.

The regression results are graphed in figure 3, which shows a map of the study area split by the desert line with each town and village plotted on the longitude on the x-axis and the latitude on the y-axis. Predicted values are estimated from the cubic polynomial model for a grid of longitude-latitude values for each outcome (excluding the sanitation network outcome). The background colours are synonymous with the typical two-dimensional curve in RD plots, while the dots show the actual outcomes as measured in the census. The maps show both considerable variation in the outcome variables across space but also capture the discontinuity in outcomes across the border line.

Figure 3: Regression discontinuity graphs

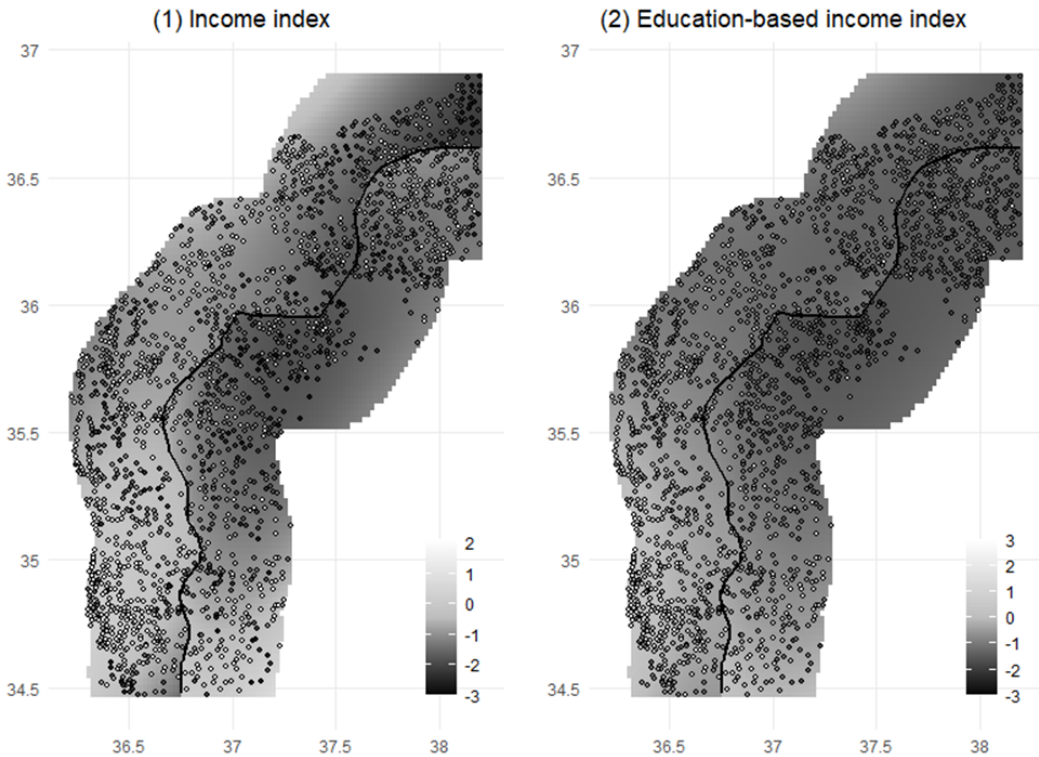


Figure 3 (continued)

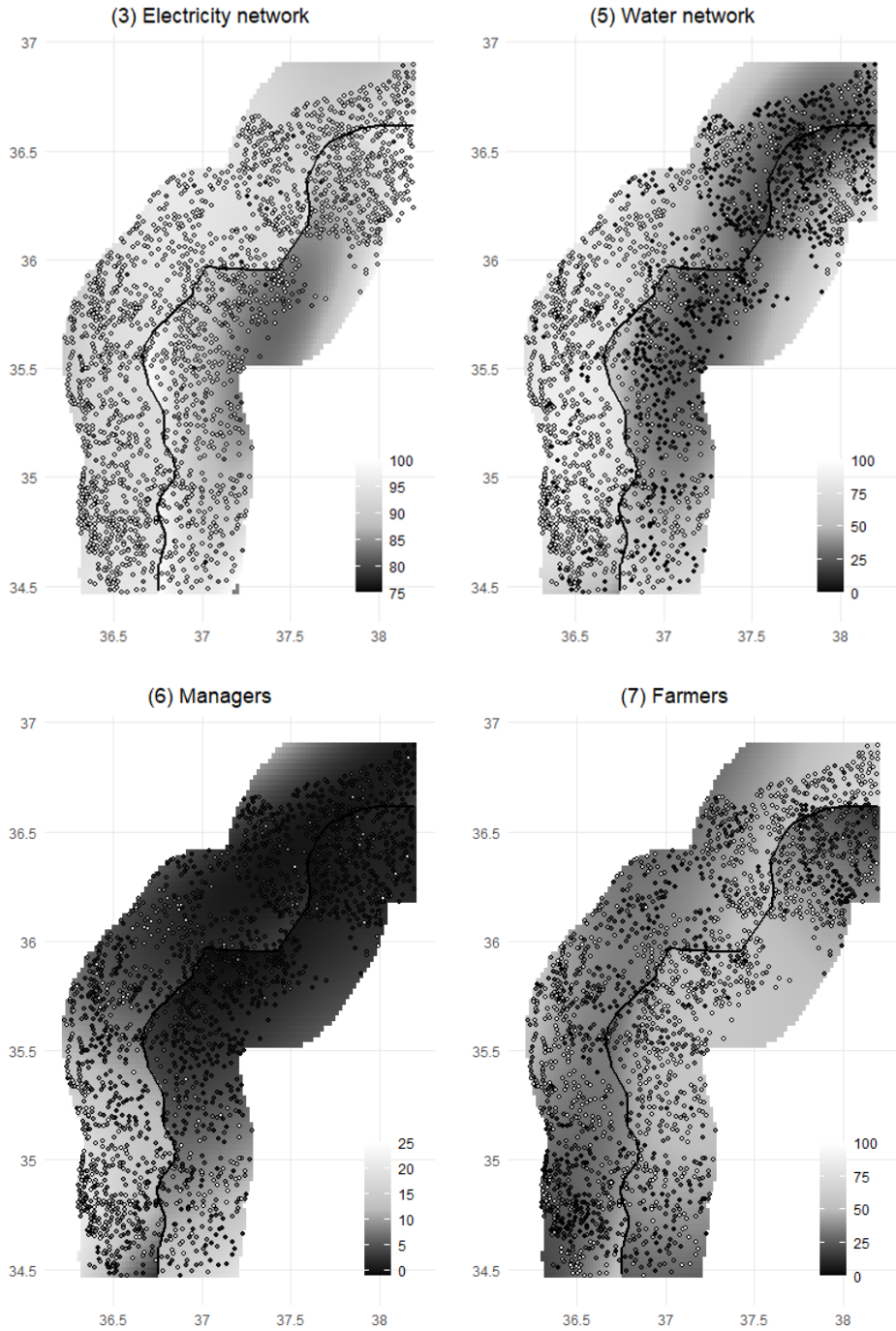
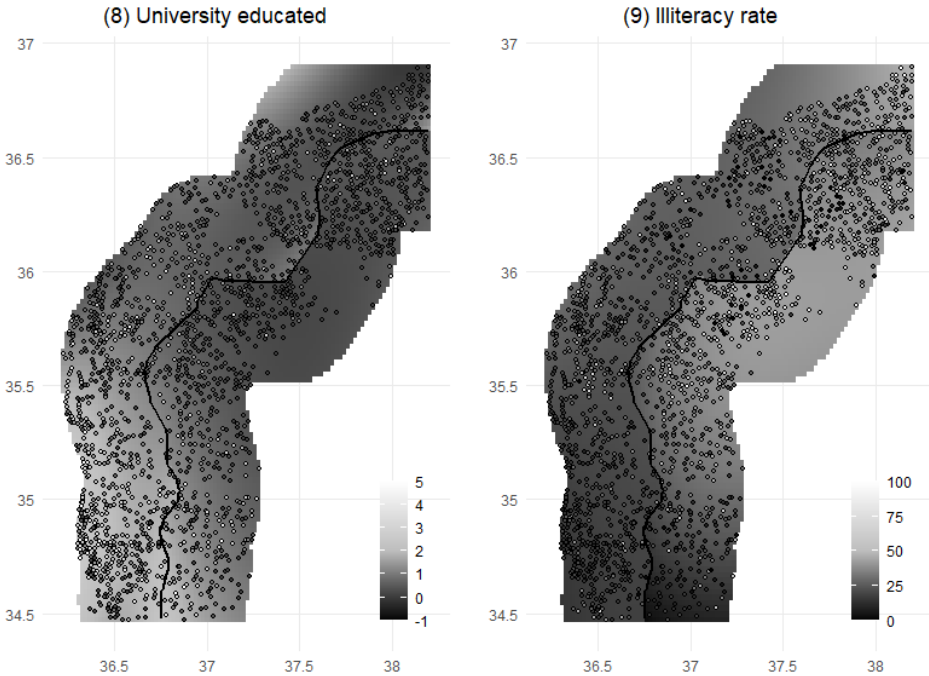


Figure 3 (continued)



Note: The x-axis and y-axis represent longitude and latitude. The data value is shown using an evenly spaced monochromatic colour scale. Actual data is shown as dots. The background shows predicted values, for a finely spaced grid of longitude-latitude coordinates, from a regression of the outcome variable using the main model detailed in the data and methods section with a cubic polynomial function in longitude and latitude.

To further evaluate the robustness of the results, I estimate spatial autoregressive lag models which address the Kelly critique (Kelly 2019) by relaxing the assumption of uncorrelated error terms. The models included in Table 4 allow each error terms to covary with that of the eight nearest observations in the sample (keeping the cubic polynomial functional form). This ensures that the effect estimates are not due to observations being correlated geographically.

In this case, the effects are overall smaller but remain statistically significant, and some remain on a similar level to those reported in the base models. The income effect is measured at -0.21 and that of education-based income at -0.19, while the effect on the electricity network and water network measures at -6.14 and -13.95 respectively (with the sanitation network effect not statistically significantly different from zero). The effect on the proportion of managers in the labour force is at -1.63, while the effect on the share of the labour force in agriculture is not statistically significantly different from zero. Finally, the effects on human capital

measures appear to be robust, measuring at -0.28 for the share with a university education, and 8.19 for the illiteracy rate.

Table 4: RD cutoff effects from spatial autoregressive lag models

	Dependent variable:								
	Income index	Education-based income	Electricity network	Sanitation network	Water network	Managers	Farmers	University educated	Illiteracy rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Stateless	-0.211**	-0.191***	-6.144**	-2.441	-13.954***	-1.625*	3.751	-0.283**	8.189***
	(0.103)	(0.066)	(2.460)	(3.885)	(4.032)	(0.895)	(2.819)	(0.143)	(1.767)
Observations	1,908	1,908	1,908	1,908	1,908	1,908	1,908	1,908	1,908

*p<0.1; **p<0.05; ***p<0.01

Moran's I statistics are reported to examine correlation in the error terms of the estimated models. Moran's I statistic takes values between -1 and +1, where 0 corresponds to random spatial patterns without spatial autocorrelation, and a value of 1 suggests perfect clustering where locations with similar values are adjacent to each other. Table 5 reports Moran's I statistics for the residual terms in the base cubic polynomial models and the spatial lag models. Overall, Moran's I for the base cubic polynomial models are reported at values between 0.1 and 0.3, which suggests some level of clustering in the residuals. Clustering is not pronounced as the values are on the lower end of the scale. In the case of the spatial lag models, all Moran's I statistics are not statistically significantly different from zero.

Table 5: Moran's I statistics

	Dependent variable:								
	Income index	Education-based income	Electricity network	Sanitation network	Water network	Managers	Agriculture	University educated	Illiteracy rate
Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Cubic polynomial model									
Moran's I	0.280***	0.246***	0.071***	0.177***	0.205***	0.255***	0.295***	0.139***	0.171***
	[25.64]	[22.583]	[6.565]	[16.188]	[18.764]	[23.48]	[26.994]	[12.886]	[15.706]
Panel B: Spatial autoregressive lag model									
Moran's I	-0.011	-0.008	-0.005	-0.008	-0.011	-0.014	-0.015	-0.007	-0.011
	[-0.98]	[-0.676]	[-0.413]	[-0.675]	[-0.916]	[-1.239]	[-1.299]	[-0.590]	[-0.976]

*p<0.1; **p<0.05; ***p<0.01

Note: The above table reports the Moran's I statistics for the cubic polynomial models and the spatial autoregressive lag models based on the 8 nearest neighbours of each point. Z-scores are reported in brackets.

Panel B of Table 6 reports results for a set of placebo outcomes estimated using linear models in longitude and latitude. These variables—unemployment rate, female labour force participation (FLFP), characteristics of dwellings in the local area (share of dwellings under refurbishment, share rented, and share vacant), and the gender gap in educational attainment—are chosen because they are not expected to be shaped by the region's historical state presence and do not serve as close proxies for contemporary economic development.

Unemployment and FLFP capture labour market conditions, rather than the long-run structural advantages linked to state history. The selected housing characteristics primarily reflect current real estate dynamics and local demand–supply conditions and thus are unlikely to have been persistently influenced by pre-modern political structures. Similarly, the gender gap in education is a product of modern schooling policies and social norms, not the historical institutional environment.

The estimates confirm these expectations: the effects on the placebo outcomes are statistically insignificant, except for a very small and marginally significant effect (at the 10 % level) on the share of vacant dwellings. This pattern supports the validity of the main causal estimates by showing no systematic differences for outcomes unrelated to long-run economic development.

Table 6: RD effects using placebo outcomes

	Dependent variable:					
	Unemployment	FLFP	Dwellings in refurbishment	Rented dwellings	Empty dwellings	Gender gap in schooling
	(1)	(2)	(4)	(5)	(6)	(7)
Stateless	0.845	-1.232	-0.132	-0.224	1.270*	0.223
	(0.974)	(1.091)	(0.441)	(0.571)	(0.753)	(0.365)
Mean Y (control)	12.855	15.837	4.839	65.831	1.540	9.624
SD	15.271	16.701	6.108	39.261	6.669	10.437
Observations	1,908	1,908	1,908	1,908	1,908	1,908

*p<0.1; **p<0.05; ***p<0.01

Note: The table shows the regression discontinuity effects using a set of placebo outcomes with a linear function in longitude and latitude. The outcomes considered are the unemployment rate, female labour force participation rate (FLFP), ratio of males to females, ratio of dwellings undergoing refurbishment, ratio of rented dwellings, ratio of empty dwellings, and the percentage points gap in secondary education completion rates between females and males.

6. Mechanisms

6.1. Ethnicity and religion

The previous section showed that the results are robust to controlling for ethnic and religious identity. This factor is explored further in Table 7, which shows the regression discontinuity effects for different sub-groups. Panel A excludes all area with non-Sunni Muslim majorities. The results remain broadly similar and statistically significant, and in many cases the effects are larger than those found in the full sample. Panel B further excludes areas with tribal majorities. In this case, the effects are larger for the income and education outcomes and are more than doubled for the percentage of farmers in the labour force, though they are statistically insignificant for the infrastructure variables.

Table 7: RD effects according to ethnic group

	Dependent variable:								
	Income index	Education-based income	Electricity network	Sanitation network	Water network	Managers	Farmers	University educated	Illiteracy rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Sunni Muslims									
Stateless	-0.624***	-0.563***	-9.876***	-5.631	-28.045***	-3.885***	10.260***	-0.806***	14.511***
	(0.124)	(0.068)	(3.101)	(4.603)	(4.807)	(0.642)	(3.923)	(0.152)	(2.115)
Mean Y (Control)	-1.066	-1.006	93.736	30.316	61.615	3.831	42.332	0.849	27.304
Observations	1,438	1,438	1,438	1,438	1,438	1,438	1,438	1,438	1,438
Panel B: Sunni Muslims excluding tribes									
Stateless	-1.151***	-0.799***	14.797	-13.331	-5.628	-2.220	23.947**	-1.360***	14.057**
	(0.386)	(0.219)	(9.144)	(14.423)	(15.667)	(2.283)	(11.190)	(0.459)	(6.115)
Mean Y (Control)	-0.826	-0.867	94.187	36.507	74.914	4.932	37.299	1.004	24.254
Observations	745	745	745	745	745	745	745	745	745
Panel C: Sunni tribes									
Stateless	-0.132	-0.128	-7.481	-13.781	1.739	-1.425	3.485	-0.322	4.318
	(0.364)	(0.160)	(7.317)	(13.355)	(17.025)	(1.217)	(11.791)	(0.326)	(8.036)
Mean Y (Control)	-1.565	-1.296	92.798	17.452	33.981	1.544	52.792	0.528	33.640
Observations	693	693	693	693	693	693	693	693	693

*p<0.1; **p<0.05; ***p<0.01

Note: The table shows the regression discontinuity effect for three subsamples. Panel A includes only locations with majority Sunni Muslim populations. Panel B further excludes

locations with majority tribal populations. Panel C includes only locations with majority tribal populations. Mean outcomes are shown for areas with an Ottoman state legacy.

Overall, the results are robust to excluding minority groups such as Christians, Alawis, Druzes, and Ismailis whose outcomes may differ from that of the Sunni Muslim majority for factors unrelated to the border line (such as discrimination or differences in cultural norms). The results also do not simply reflect a penalty held by the tribal population of North-East Syria either, as the effects are larger when this group is excluded (except for the infrastructure outcomes).

Panel C of Table 7 shows the effects for the subsample of areas with a Sunni tribal majority. The effects on income and education level are reduced to very low levels and none of the effects are statistically significant at the 0.1 level. Though the mean outcomes for tribal areas are overall closer to those of areas with a stateless legacy (in non-tribal areas). The lack of an effect in tribal areas is consistent with the stateless legacy of these communities, with settlement in Ottoman controlled areas occurring in the 19th century precluding a state legacy for the affected communities even if the land itself had been under state control prior to the 19th century.

6.2. Property rights

To analyse the role of institutions as a mechanism for persistence, I use data on informality in home ownership. The data records the share of housing deeds according to the type of deed, where different deed types correspond to different levels of informality. Three types are recorded in the data - formal deeds, agricultural deeds, and public notary contracts. Formal deeds correspond to the share of formally owned homes. Agricultural deeds correspond to the share of homes built on formerly agricultural land without a government permit. And public notary contracts correspond to the share of homes where owners possess a notarized contract verifying their purchase of the property from the previous owner, but without holding an official deed to the property.

Table 8 shows that informality in home ownership is higher in historically stateless areas. The share of homes owned with an official deed is significantly lower, and the proportion of homes owned via an agricultural deed higher, in historically stateless areas (columns 1 and 3 respectively). Similar effects are present when controlling for ethnic and religious identity (columns 2 and 4). However, no effect is found on the proportion of homes owned via notary contract. The models in this case control for linear terms in longitude and latitude, but similar effects hold when controlling for cubic polynomial terms, even if they mostly become statistically insignificant (see Appendix Table A7).

Table 8: RD effects on informality in property ownership

	Dependent variable:					
	Official deed		Agricultural deed		Public notary	
	(1)	(2)	(3)	(4)	(5)	(6)
Stateless	-8.376***	-6.710***	5.235**	4.754**	-0.431	-0.388
	(2.169)	(2.197)	(2.314)	(2.392)	(0.351)	(0.358)
Ethnicity	-	Y	-	Y	-	Y
Mean Y (Control)	37.894	37.894	20.615	20.615	1.105	1.105
SD	36.521	36.521	34.773	34.773	5.460	5.460
Observations	1,908	1,908	1,908	1,908	1,908	1,908

*p<0.1; **p<0.05; ***p<0.01

Note: The table shows the regression discontinuity effect on the share of properties owned according to official deeds, agricultural deeds, and public notary deeds, using a linear function in longitude and latitude.

Despite the observed effect of statelessness on property rights, Table 9 shows that differences in property rights across the desert line do not contribute significantly to explaining the gaps in economic development across the line. The regression discontinuity effects remain largely the same as in the unadjusted model in Table 3 (panel B). Overall, the effects of the property ownership variables on income and human capital are very small, but those for the infrastructure outcomes are more significant, yet they do not shift of the effect of historical statelessness. The effect sizes for the institutional variables remain largely similar when dropping the regression discontinuity variable and treating the study area as one region (See Appendix Table A8).

Table 9: RD effects on development outcomes controlling for institutional variables

	Dependent variable:								
	Income index	Education-based income	Electricity network	Sanitation network	Water network	Managers	Farmers	University educated	Illiteracy rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Stateless	-0.584***	-0.441***	-7.974***	-3.476	-28.096***	-3.628***	10.353***	-0.500***	13.889***
	(0.113)	(0.066)	(2.581)	(3.794)	(4.194)	(0.800)	(3.435)	(0.154)	(1.938)
Official deed	0.002**	0.001***	0.150***	0.294***	0.210***	-0.00002	-0.039**	0.005***	-0.010
	(0.001)	(0.0005)	(0.016)	(0.026)	(0.027)	(0.007)	(0.020)	(0.001)	(0.011)
Agricultural deed	0.001	-0.001*	0.126***	0.028	0.036	0.001	0.001	-0.0002	0.042***
	(0.001)	(0.0004)	(0.017)	(0.021)	(0.026)	(0.006)	(0.021)	(0.001)	(0.012)
Public notary	0.001	0.008***	0.194***	0.778***	0.665***	0.029*	-0.017	0.007*	-0.213***
	(0.002)	(0.002)	(0.042)	(0.171)	(0.113)	(0.015)	(0.102)	(0.004)	(0.075)
Observations	1,908	1,908	1,908	1,908	1,908	1,908	1,908	1,908	1,908

*p<0.1; **p<0.05; ***p<0.01

Note: The table shows the regression discontinuity effects and the effects of institutional variables on development outcomes, using a cubic polynomial in longitude and latitude.

6.3. Population density

Table 10 shows the regression discontinuity effects on three different measures of population density - the population size of the town or village itself, population size within a 5 km diameter, and population size within a 10 km diameter. The effects are generally large, on the scale of one standard deviation of the mean, and robust to the addition of ethnicity controls. This points to an important effect for historical statelessness in shaping population density, which can have important consequences for economic development. The evidence here is in line with the historical narratives that emphasize land abandonment and emigration from stateless areas.

Table 10: RD effect on population density

	Dependent variable:					
	Total population		Population within 5 Km diameter		Population within 10 Km diameter	
	(1)	(2)	(3)	(4)	(5)	(6)
Stateless	-1,360.842***	-1,409.517***	-17,343.250***	-18,178.850***	-99,027.280***	-99,859.590***
	(171.637)	(173.711)	(3,760.468)	(3,983.053)	(11,083.590)	(10,891.420)
Ethnicity	-	Y	-	Y	-	Y
Mean Y (Control)	1,644	1,644	13,323	13,323	75,512	75,512
SD	1,464	1,464	20,515	20,515	109,010	109,010
Observations	1,908	1,908	1,908	1,908	1,908	1,908

*p<0.1; **p<0.05; ***p<0.01

Note: The table shows the regression discontinuity effect on the total population in the village or town, total population within a 5 Km diameter, and total population within a 10 Km diameter, controlling for a cubic polynomial function in longitude and latitude.

To explore whether population density can explain the effect of stateless legacy on economic development today, Table 11 shows the regression discontinuity effects on economic development outcomes controlling for the population density measures. The last row compares the regression discontinuity effects here to those in the unadjusted model (Table 3, Panel B). The inclusion of the population size and density controls accounts for around half of the effect on incomes and human capital, and an even larger share of the effects on physical capital and the share of farmers. The coefficients on the density variables are generally significant but show slightly different patterns. For income and human capital, the population of the village or town itself matters most, whereas population density within a larger diameter matters more for infrastructure development outcomes. Being embedded in a wide dense network of towns and villages appears to matter most for infrastructure outcomes, and accounts for most of the effects observed for historical

statelessness. The regression discontinuity effect on the sanitation network (which is small and insignificant in the unadjusted cubic polynomial model) becomes positive and statistically significant, suggesting that historically stateless areas have an advantage in this outcome once low-density constraints are taken into account.

Table 11: RD effects controlling for population density

	Dependent variable:								
	Income index	Education-based income	Electricity network	Sanitation network	Water network	Managers	Farmers	University educated	Illiteracy rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Stateless	-0.297**	-0.190***	0.592	19.027***	-7.949*	-3.483***	2.837	-0.236	6.398***
	(0.122)	(0.070)	(2.835)	(4.225)	(4.582)	(0.928)	(3.765)	(0.155)	(2.083)
ln population	0.181***	0.152***	4.057***	11.728***	8.433***	1.013***	-4.841***	0.214***	-2.754***
	(0.025)	(0.016)	(0.596)	(0.844)	(0.868)	(0.232)	(0.699)	(0.035)	(0.420)
ln population within 5 Km	0.064***	0.030**	0.898*	1.106	1.351*	0.348***	-1.806***	0.014	-1.143***
	(0.018)	(0.013)	(0.462)	(0.763)	(0.753)	(0.134)	(0.519)	(0.024)	(0.309)
ln population within 10 Km	0.008	0.030	1.989***	4.888***	5.799***	-0.881**	0.014	-0.004	-1.993***
	(0.034)	(0.026)	(0.729)	(1.286)	(1.315)	(0.344)	(1.056)	(0.067)	(0.550)
Observations	1,908	1,908	1,908	1,908	1,908	1,908	1,908	1,908	1,908
% explained by population	49	56	100	-	71	5	72	50	53

*p<0.1; **p<0.05; ***p<0.01

Note: The table shows the regression discontinuity effect controlling for the size of the local population and the population of the neighbouring area within diameters of 5- and 10-kilometers distance (in log form). The last row displays the percentage of the unadjusted RD effect explained by adding controls for population density, comparing the RD effects in this table with those in the unadjusted model in Table 3 (Panel B).

6.4. Emigration

The evidence above is in line with a role for path dependence as a mechanism for persistence. Historical statelessness can shape outcomes today through its effect on population density, by causing persistent emigration flows. While data on historical population movements is not currently available, it is possible to detect contemporary emigration through its effect on the age structure of the population. If we assume that high emigration rates disproportionally affect males in working

age (age 15 to 64), then there should be a detectable effect on the share of that group in the overall population in historically stateless areas.

Table 12 shows the results from estimating the effect of historical statelessness on the proportion of each age group within the male population (columns 1 to 3) and on the ratio of males to females within each age group (columns 4 to 6). The first set of measures show that the stateless legacy reduces the proportion of working-age males within the overall male population. Areas with a stateless history experience a gap of -2.5 percentage points in the working age male population, as compared to control areas. Meaning that historical statelessness reduces the working-age male share by 4.7% relative to the control mean value. This effect is mirrored in the male-to-female ratio within the same age group, which declines by 0.048 in historically stateless areas, also representing a 4.7% relative reduction in gender balance within the working-age group. These consistent effects strongly indicate male-selective out-migration. The effect is also robust to controlling for fertility rates (see Appendix table A9), and to excluding areas with non-Sunni majorities (See Appendix Table A10).

Table 12: RD effects on age structure in the male population and the sex ratio according to age

	Dependent variable:					
	Males under 15	Males 15 to 64	Males 65 and over	Males/females under 15	Males/females 15 to 64	Males/females 65 and over
	(1)	(2)	(3)	(4)	(5)	(6)
Stateless	2.890***	-2.513***	-0.377***	0.026**	-0.048**	0.079
	(0.375)	(0.346)	(0.120)	(0.011)	(0.021)	(0.068)
Mean Y (Control)	43.012	53.174	3.813	1.081	1.027	1.412
SD	7.314	6.174	2.459	0.153	0.311	0.832
Observations	1,872	1,872	1,872	1,872	1,872	1,856

*p<0.1; **p<0.05; ***p<0.01

Note: The table shows the RD effects on the proportion of each age group within the male population divided by the proportion of the age group in the total population (columns 1 to 3), and the RD effects on the sex ratio within each age group, controlling for linear terms in longitude and latitude.

7. Discussion and conclusion

The results above provide evidence of the long-term economic consequences of historical statelessness in Ottoman Syria. The findings reveal a persistent economic gap between regions that experienced prolonged state control and those that remained outside its reach until the mid-19th century. These results contribute to the broader literature on institutional persistence, economic geography, and path dependence by highlighting how the absence of stable governance in the past continues to shape economic outcomes in the present.

The analysis demonstrates that historically stateless areas have lower incomes, less developed infrastructure, lower levels of human capital, and a less industrialized economy. These gaps remain robust even when controlling for potential confounders such as geography and ethnic differences. Importantly, the results suggest that these economic disparities are not primarily driven by differences in private property institutions or ethnic composition but are instead strongly correlated with population density. The lower population density in historically stateless regions appears to be a major mechanism of persistence, with historical insecurity leading to emigration and reduced long-term settlement in these areas.

The role of population density in sustaining regional economic disparities aligns with existing research on path dependence in economic development (Bleakley and Lin 2012; Baerlocher et al. 2024). The historical insecurity and lack of state protection in stateless areas created conditions that discouraged permanent settlement and economic investment. Over time, these conditions led to lower population densities, which in turn limited the potential for human and physical capital accumulation. The regression results indicate that differences in population density account for a significant portion of the observed income and human capital gaps, reinforcing the argument that path dependence is a key mechanism through which historical state presence influences contemporary economic development.

The findings are further corroborated by evidence on contemporary migration patterns. The persistent out-migration of working-age males from historically stateless regions suggests that these areas continue to struggle with retaining their labor force, further exacerbating their economic disadvantage. This persistent demographic pattern aligns with historical narratives of rural insecurity and land abandonment due to tribal incursions and weak state authority. These dynamics highlight how frontier conditions can have enduring effects on local economic structures, a pattern observed in other historical contexts as well (Oto-Peralías 2020).

The findings have important implications for policy efforts aimed at reducing regional inequalities and promoting economic development in historically marginalized areas. Policies that address population density constraints—such as targeted infrastructure investments, incentives for urbanization, and improvements

in local governance—may help mitigate some of the long-term disadvantages faced by historically stateless regions. In particular, investments in transportation networks, educational institutions, and economic diversification strategies could help attract and retain populations in these areas, fostering more sustainable long-term development.

These challenges are deeply rooted in historical patterns of settlement and conflict. The tension between settled agrarian communities and nomadic groups has shaped the economic geography of the MENA region for centuries and continues to influence the politics of many developing countries today, particularly in Sub-Saharan Africa. In a world affected by climate change, such frontier zones remain sites of recurring conflict, as worsening climatic conditions increase pressures on pastoralist populations (McGuirk and Nunn 2024, Bai and Kung 2011). The analysis here highlights how these conflicts can have lasting consequences for economic development, shaping patterns of human settlement and urbanization in ways that persist over time. Addressing these historical legacies through forward-looking policies may be key to fostering more inclusive and resilient development trajectories.

References

Abu-Ismail, K., Abdel-Gadir, A., & El-Laithy, H. (2016). Poverty and Inequality in Syria. *Arab Development Challenges Report Background Paper 2011* (15).

Baerlocher, D., Firmino Costa da Silva, D., Lambais, G., Reis, E., & Veras, H. (2025). Old but gold: historical pathways and path dependence. Available at SSRN 4513384.

Bai, Y., & Kung, J. K. S. (2011). Climate shocks and Sino-nomadic conflict. *Review of Economics and Statistics, 93*(3), 970-981.

Barakat, N. E. (2015). *An Empty Land? Nomads and Property Administration in Hamidian Syria* (Doctoral dissertation, UC Berkeley).

Barout, M.J. (2013). *Contemporary History of the Syrian Jazeerah: Challenges of Urban Transition for Nomadic Communities.* Arab Centre for Research and Policy Studies.

Bayat, F.M. (2017), *البلاد العربية في الوثائق العثمانية.* Research Centre for Islamic History, Art and Culture (IRCICA).

Bleakley, H., & Lin, J. (2012). Portage and path dependence. *The quarterly journal of economics, 127*(2), 587-644.

Bockstette, V., Chanda, A., & Putterman, L. (2002). States and markets: The advantage of an early start. *Journal of Economic growth, 7*(4), 347-369.

Borcan, O., Olsson, O., & Putterman, L. (2018). State history and economic development: evidence from six millennia. *Journal of Economic Growth, 23*(1), 1-40.

Boserup, E. (2014). *The conditions of agricultural growth: The economics of agrarian change under population pressure.* Routledge.

Chanda, A., & Putterman, L. (2007). Early starts, reversals and catch-up in the process of economic development. *Scandinavian Journal of Economics, 109*(2), 387-413.

Dalgaard, C. J., Kaarsen, N., Olsson, O., & Selaya, P. (2022). Roman roads to prosperity: Persistence and non-persistence of public infrastructure. *Journal of Comparative Economics, 50*(4), 896-916.

Dell, M. (2010). The persistent effects of Peru's mining mita. *Econometrica, 78*(6), 1863-1903.

Dell, M., Lane, N., & Querubin, P. (2018). The historical state, local collective action, and economic development in Vietnam. *Econometrica, 86*(6), 2083-2121.

Esmer, T. U. (2014). Economies of violence, banditry and governance in the Ottoman empire around 1800. *Past & Present, 224*(1), 163-199.

Gennaioli, N., & Rainer, I. (2007). The modern impact of precolonial centralization in Africa. *Journal of Economic Growth, 12*(3), 185-234.

Geyer, B. (2011). The steppe: human occupation and potentiality, the example of northern Syria's arid Margins. Syria. *Archéologie, art et histoire* (88), 7-22.

Geyer, B., Calvet, Y., Awad, N., Al-Dbiyat, M., & Rousset, M. O. (2010). Un «Très Long Mur» dans la steppe syrienne. *Paléorient*, 57-72.

Hatton, T. J. (2020). Asylum migration to the developed world: Persecution, incentives, and policy. *Journal of Economic Perspectives, 34*(1), 75-93.

Hinnebusch, R. (2012). Syria: from 'authoritarian upgrading' to revolution? *International Affairs, 88*(1), 95-113.

Hoffman, B. G. (1967). *The structure of traditional Moroccan rural society.* Mouton.

Hopkins, I. W. (1967). The maps of Carsten Niebuhr: 200 years after. *The Cartographic Journal, 4*(2), 115-118.

Khaddour, K. and Mazur, K. (2018), *Syria town database.* Harvard Dataverse, V1.

Khoury, P. S., & Kostiner, J. (Eds.). (1990). *Tribes and state formation in the Middle East.* University of California Press.

Klasen, S., & Nestmann, T. (2006). Population, population density and technological change. *Journal of Population Economics, 19*(3), 611-626.

Kremer, M. (1993). Population growth and technological change: One million BC to 1990. *The Quarterly Journal of Economics, 108*(3), 681-716.

Krugman, P. (1992). *Geography and trade.* MIT press.

Liebeschuetz, J. H. W. G. (2015). Arab tribesmen and desert frontiers in late antiquity. *Journal of Late Antiquity, 8*(1), 62-96.

Lindner, R. P. (2017). *Nomads and Ottomans in medieval Anatolia.* Routledge.

Lowes, S., Nunn, N., Robinson, J. A., & Weigel, J. L. (2017). The evolution of culture and institutions: Evidence from the Kuba Kingdom. *Econometrica, 85*(4), 1065-1091.

Ma'oz, M. (2013). The emergence of modern Syria. In *Syria Under Assad (RLE Syria)* (pp. 9-35). Routledge.

Ma'oz, M. (1968). *Ottoman reform in Syria and Palestine, 1840-1861: The impact of the Tanzimat on politics and society.* Clarendon Press.

Michaels, G., & Rauch, F. (2018). Resetting the urban network: 117–2012. *The Economic Journal, 128*(608), 378-412.

McGuirk, E. F., & Nunn, N. (2025). Transhumant pastoralism, climate change, and conflict in Africa. *Review of Economic Studies, 92*(1), 404-441.

Michalopoulos, S., & Papaioannou, E. (2013). Pre-colonial ethnic institutions and contemporary African development. *Econometrica, 81*(1), 113-152.

Lewis, N. N. (1987). *Nomads and settlers in Syria and Jordan, 1800-1980.* Cambridge University Press.

North, D. C., Wallis, J. J., & Weingast, B. R. (2009). *Violence and social orders: A conceptual framework for interpreting recorded human history.* Cambridge University Press.

Nunn, N. (2020). The historical roots of economic development. *Science, 367*(6485), eaaz9986.

Rogan, E. L. (2002). *Frontiers of the state in the late Ottoman Empire: Transjordan, 1850-1921 (Vol. 12)*. Cambridge University Press.

Robinson, E. and Smith, E. (1856). *Biblical researches in Palestine, Mount Sinai and Arabia Petraea: A journal of travels in the year 1838 (Vol. 4).* J. Murray.

Rogan, E.L., 1995. Ottomans, Merchants and Tribes in the Syrian Frontier (1867-1900). in D.Panzac, ed, *Histoire Économique et Sociale de l'Empire Ottoman et de la Turquie (1326- 1960)* (Paris: Peeters, 1995) pp. 251-62.

Rousset, M. O., Geyer, B., Awad, N., & Shabo, S. (2020). The defensive system of the arid margins of the Northern Levant during the Middle Bronze Age. In *11th International Congress on the Archaeology of the Ancient Near East* (pp. 349-362). Harrassowitz Verlag.

Oto-Peralías, D. (2020). Frontiers, warfare and economic geography: The case of Spain. *Journal of Development Economics, 146,* 102511.

Özok-Gündoğan, N. (2014). Ruling the Periphery, Governing the Land: The Making of the Modern Ottoman State in Kurdistan, 1840-70*. Comparative Studies of South Asia, Africa and the Middle East, 34*(1), 160-175.

Walker, B. J. (2012). The phenomenon of the "disappearing" villages of late medieval Jordan, as reflected in archaeological and economic sources. *Bulletin d'études orientales, (Tome LX),* 161-176.

Wickham, C. (2005). *Framing the early middle ages: Europe and the Mediterranean, 400-800.* Oxford University Press.

Williams, S. L. (1981). Ottoman land policy and social change: The Syrian provinces. *Acta Orientalia Academiae Scientiarum Hungaricae, 35*(1), 89-120.

Winter, S. (2006). The Other" Nahḍah": The Bedirxans, the Millîs and the Tribal Roots of Kurdish Nationalism in Syria. *Oriente Moderno, 25*(3), 461-474.

Appendix

Table A1: Regression discontinuity effects using 25km bands

	Dependent variable:								
	Income index	Education-based income index	Electricity network	Sanitation network	Water network	Managers	Farmers	University educated	Illiteracy rate
Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Linear model									
Stateless	-0.321***	-0.333***	-3.706***	-15.950***	-17.691***	-3.984***	2.503	-0.417***	9.216***
	(0.064)	(0.040)	(1.398)	(2.210)	(2.665)	(0.464)	(1.962)	(0.087)	(1.039)
Panel B: Cubic polynomial model									
Stateless	-0.563***	-0.548***	-10.448***	-5.398	-24.111***	-0.794	11.722***	-0.707***	17.231***
	(0.143)	(0.080)	(3.232)	(5.168)	(5.401)	(0.933)	(4.406)	(0.180)	(2.408)

*p<0.1; **p<0.05; ***p<0.01

Note: Standard errors clustered by district are in parentheses (92 clusters). All regressions control for log distance to Aleppo. The basic linear model includes controls for latitude and longitude on each side of the regression discontinuity line. The cubic polynomial model includes controls for latitude and longitude on each side of the discontinuity of the form $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$ where x and y denote longitude and latitude.

Table A2: Regression discontinuity effects with controls for line segment fixed effects

	Dependent variable:								
	Income index	Education-based income index	Electricity network	Sanitation network	Water network	Managers	Farmers	University educated	Illiteracy rate
Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Linear model									
Stateless	-1.382***	-0.718***	-7.427*	-14.025*	-39.782***	-3.009**	35.223***	-1.306***	14.230***
	(0.224)	(0.123)	(4.119)	(7.704)	(9.287)	(1.485)	(6.493)	(0.236)	(3.040)
Panel B: Cubic polynomial model									
Stateless	-1.121***	-0.600***	14.324	-3.535	4.686	-0.716	26.420**	-1.146**	6.927
	(0.418)	(0.230)	(9.319)	(14.545)	(16.443)	(2.349)	(11.860)	(0.494)	(6.459)

*p<0.1; **p<0.05; ***p<0.01

Note: Standard errors clustered by district are in parentheses (92 clusters). All regressions control for log distance to Aleppo. The basic linear model includes controls for latitude and longitude on each side of the regression discontinuity line. The cubic polynomial model includes controls for latitude and longitude on each side of the discontinuity of the form $x + y + x^2 + y^2 + xy + x^3 + y^3 + x^2y + xy^2$ where x and y denote longitude and latitude.

Table A3: RD regression models with linear longitude and latitude

	Dependent variable:								
	Income index	Education-based income	Electricity network	Sanitation network	Water network	Managers	Agriculture	University educated	Illiteracy rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Stateless	-0.532***	-0.438***	-3.410**	-21.942***	-22.491***	-5.010***	7.098***	-0.530***	9.216***
	(0.061)	(0.037)	(1.326)	(2.047)	(2.529)	(0.407)	(1.853)	(0.075)	(1.039)
log(to_Aleppo)	0.064	0.263***	1.897	-4.460**	4.832**	2.511***	1.820	0.291***	0.559
	(0.051)	(0.025)	(1.153)	(1.963)	(2.105)	(0.237)	(1.658)	(0.044)	(1.005)
Stateless:long	0.608***	0.097	2.143	14.725***	19.644***	4.324***	-16.727***	-0.082	-3.884
	(0.146)	(0.089)	(2.506)	(4.333)	(6.004)	(0.960)	(4.240)	(0.133)	(2.430)
Stateless:lat	-0.781***	-0.504***	-0.152	-22.324***	-17.730***	-7.478***	6.335*	-0.323***	13.796***
	(0.135)	(0.078)	(2.133)	(3.823)	(4.688)	(1.080)	(3.551)	(0.115)	(1.853)
long:Control	-0.517***	-0.341***	-0.849	2.815	-45.193***	-2.290***	9.300***	-0.223	4.748***
	(0.097)	(0.062)	(1.837)	(3.456)	(3.145)	(0.837)	(2.633)	(0.148)	(1.304)
lat:Control	-0.357***	-0.320***	1.638	-21.362***	8.828***	-5.116***	6.159**	-0.579***	6.489***
	(0.094)	(0.062)	(1.791)	(3.303)	(3.069)	(0.888)	(2.453)	(0.156)	(1.234)
Constant	-1.186***	-1.936***	85.048***	55.066***	35.440***	-3.640***	33.210***	-0.100	23.611***
	(0.230)	(0.114)	(5.110)	(8.812)	(9.353)	(1.085)	(7.375)	(0.215)	(4.457)
Observations	1,908	1,908	1,908	1,908	1,908	1,908	1,908	1,908	1,908
Adjusted R^2	0.197	0.407	0.004	0.121	0.248	0.308	0.059	0.196	0.206
Residual Std. Error	0.999	0.620	20.554	34.832	36.227	8.374	27.628	1.242	15.350

*p<0.1; **p<0.05; ***p<0.01

Table A4: RD regression models with cubic polynomial longitude and latitude

	Dependent variable:								
	Income index	Education-based income	Electricity network	Sanitation network	Water network	Managers	Agriculture	University educated	Illiteracy rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Stateless	-0.577***	-0.431***	-7.754***	-2.006	-27.177***	-3.648***	10.098***	-0.471***	13.676***
	(0.113)	(0.066)	(2.659)	(4.073)	(4.289)	(0.797)	(3.446)	(0.154)	(1.956)
log(to_Aleppo)	-0.431***	-0.051	6.645**	-18.919***	-3.179	-2.916***	7.264*	-0.013	0.803
	(0.119)	(0.065)	(2.691)	(4.873)	(4.970)	(0.652)	(3.752)	(0.119)	(2.348)
Stateless:long	-1.283***	-0.711***	-26.378***	3.557	-41.342***	0.638	34.690***	-0.870***	19.960***
	(0.257)	(0.160)	(6.125)	(8.990)	(11.380)	(1.319)	(7.998)	(0.223)	(4.888)
Stateless:lat	-0.693**	-0.593***	18.348**	-29.362**	10.328	-11.177***	0.627	-0.398	1.206
	(0.335)	(0.198)	(7.438)	(12.329)	(13.887)	(2.094)	(9.890)	(0.329)	(5.550)
Stateless:I(long2)	4.244***	2.006***	8.333	-39.324	168.202***	14.447***	-103.296***	0.502	-38.963**
	(0.984)	(0.576)	(23.013)	(31.054)	(40.521)	(4.633)	(30.744)	(0.957)	(17.851)
Stateless:I(lat2)	1.146***	0.485**	15.010**	6.574	39.609***	4.392	-23.156**	0.340	-19.748***
	(0.370)	(0.219)	(6.915)	(10.539)	(12.555)	(2.886)	(9.378)	(0.306)	(4.566)
Stateless:I(long3)	-1.352	-0.784	44.052	92.287**	-35.050	-7.799	63.843	0.869	20.437
	(1.540)	(0.962)	(32.183)	(46.954)	(60.395)	(8.142)	(45.528)	(1.417)	(28.023)
long:Control	-1.028***	0.189	-0.364	-8.392	-71.508***	3.428	45.334***	1.264	3.839
	(0.398)	(0.281)	(7.888)	(15.269)	(15.294)	(3.121)	(12.162)	(0.781)	(7.304)
lat:Control	-1.215***	-1.043***	12.139	-34.183**	-14.549	-19.257***	-4.806	-1.906***	6.528
	(0.400)	(0.265)	(7.509)	(14.863)	(13.643)	(3.080)	(10.787)	(0.618)	(5.222)
I(long2):Control	-0.100	1.994**	2.685	54.489	-6.858	14.513	52.462*	5.810**	-3.051
	(0.980)	(0.781)	(15.680)	(35.423)	(33.707)	(10.670)	(28.749)	(2.446)	(14.450)

Table A4 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
I(lat2):Control	0.694	0.803*	-3.888	45.299***	7.173	5.527	6.523	2.383*	-4.571
	(0.469)	(0.415)	(7.754)	(17.504)	(16.662)	(5.431)	(12.305)	(1.353)	(6.102)
I(long3):Control	0.686	1.284*	23.837	45.496	56.780*	0.755	-10.192	4.842**	-11.681
	(0.921)	(0.682)	(16.348)	(34.803)	(31.143)	(8.970)	(26.977)	(2.053)	(13.388)
Control:I(lat3)	1.862***	0.703*	-7.802	45.668***	39.235***	20.397***	-20.401*	0.409	-0.834
	(0.477)	(0.359)	(8.996)	(17.267)	(14.567)	(4.747)	(11.868)	(1.071)	(5.660)
Stateless:long:lat	-0.648	0.256	-4.485	78.115**	-106.757***	-2.458	33.977	1.926*	10.024
	(1.115)	(0.693)	(20.732)	(36.527)	(40.930)	(8.245)	(29.094)	(0.998)	(15.637)
Stateless:lat:I(long2)	-2.476	-1.295	-70.232	-156.345*	-7.889	-3.646	-48.651	-2.935	1.445
	(2.890)	(1.880)	(49.050)	(84.778)	(104.594)	(19.450)	(75.564)	(2.451)	(48.397)
Stateless:long:I(lat2)	2.321	0.927	22.072	43.891	-41.618	8.830	-6.409	0.738	-0.976
	(1.451)	(0.921)	(23.895)	(42.777)	(51.037)	(11.162)	(36.929)	(1.326)	(22.467)
long:lat:Control	0.170	-1.821	-8.093	-46.805	-4.616	-6.889	-62.188*	-6.849*	8.908
	(1.310)	(1.122)	(22.552)	(48.078)	(45.988)	(14.862)	(36.884)	(3.612)	(18.924)
lat:I(long2):Control	0.017	-3.424**	-29.093	-28.718	-4.773	0.033	-34.567	-11.416**	29.946
	(2.126)	(1.736)	(38.033)	(77.756)	(66.692)	(24.159)	(56.631)	(5.527)	(27.615)
long:I(lat2):Control	-1.759	0.902	4.612	-60.545	-27.705	-20.505	46.328	5.004	-12.402
	(1.508)	(1.309)	(26.844)	(53.739)	(43.422)	(18.363)	(36.137)	(4.313)	(16.195)
Constant	0.569	-0.835***	68.959***	92.621***	68.507***	16.131***	15.063	0.922**	21.981**
	(0.454)	(0.245)	(10.468)	(18.701)	(18.851)	(2.409)	(14.329)	(0.420)	(8.884)
Observations	1,908	1,908	1,908	1,908	1,908	1,908	1,908	1,908	1,908
Adjusted R^2	0.264	0.451	0.021	0.164	0.288	0.346	0.126	0.224	0.246

*p<0.1; **p<0.05; ***p<0.01

Table A5: RD regression models with cubic polynomial longitude and latitude and controls for ethnic majority

	Dependent variable:								
	Income index	Education-based income	Electricity network	Sanitation network	Water network	Managers	Agriculture	University educated	Illiteracy rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Stateless	-0.645***	-0.523***	-8.703***	-3.232	-28.155***	-4.279***	11.210***	-0.653***	14.533***
	(0.111)	(0.062)	(2.729)	(4.122)	(4.289)	(0.735)	(3.448)	(0.145)	(1.977)
log(to_Aleppo)	-0.318***	0.062	7.177***	-17.825***	-2.463	-1.680***	5.594	0.144	-0.266
	(0.118)	(0.065)	(2.589)	(4.930)	(4.991)	(0.645)	(3.644)	(0.121)	(2.320)
Ethnicity Alawi	0.913***	0.615***	2.342	1.473	7.301**	9.941***	-15.601***	0.575***	-6.887***
	(0.109)	(0.072)	(2.200)	(3.515)	(3.267)	(1.117)	(2.508)	(0.137)	(0.974)
Ethnicity Christian	0.620***	1.320***	4.993	33.353***	18.697***	0.051	-12.192***	3.253***	-8.362***
	(0.188)	(0.142)	(3.275)	(5.305)	(4.191)	(1.721)	(4.480)	(0.389)	(1.336)
Ethnicity Druze	-0.139	0.307*	0.143	-34.472***	13.482***	3.329	15.540*	0.648*	-4.832
	(0.361)	(0.165)	(2.648)	(5.707)	(4.097)	(2.527)	(8.859)	(0.339)	(5.197)
Ethnicity Kurdish	-0.864***	-0.179**	-8.462**	-17.445***	-21.603***	-0.143	26.962***	-0.153	8.061**
	(0.121)	(0.079)	(3.509)	(6.217)	(6.610)	(0.578)	(4.434)	(0.117)	(3.703)
Ethnicity Mixed	0.557***	0.489***	6.035***	3.694	3.765	5.394***	-10.857***	0.651***	-4.311***
	(0.143)	(0.090)	(1.403)	(4.713)	(4.608)	(1.323)	(3.165)	(0.190)	(1.481)
Ethnicity Other minority	0.234	0.341***	-0.621	-1.590	0.866	4.922**	-1.375	0.670*	-1.308
	(0.157)	(0.124)	(3.433)	(5.284)	(5.279)	(1.972)	(3.655)	(0.346)	(1.792)
Ethnicity Sunni tribes	-0.177**	0.008	-5.617***	-3.789	-7.262**	0.782	5.829**	0.048	3.975***
	(0.084)	(0.048)	(2.137)	(2.857)	(3.444)	(0.488)	(2.450)	(0.095)	(1.527)
Stateless:long	-1.374***	-0.891***	-23.633***	5.607	-38.304***	-2.326*	34.216***	-1.139***	18.958***
	(0.250)	(0.153)	(6.161)	(9.443)	(11.333)	(1.272)	(7.990)	(0.229)	(4.930)

Table A5 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Stateless:lat	-0.007	-0.110	26.177***	-23.109*	20.637	-5.846***	-12.982	0.249	-6.712
	(0.346)	(0.202)	(7.438)	(12.814)	(14.537)	(2.229)	(10.125)	(0.366)	(5.686)
Stateless:I(long2)	4.089***	1.785***	12.068	-37.138	172.891***	10.542**	-103.059***	0.298	-40.040**
	(0.993)	(0.550)	(23.437)	(31.591)	(40.980)	(4.690)	(31.119)	(0.885)	(18.322)
Stateless:I(lat2)	1.252***	0.558***	18.982***	9.473	43.458***	4.270	-26.260***	0.514	-22.104***
	(0.358)	(0.206)	(7.068)	(10.736)	(12.832)	(2.771)	(9.339)	(0.314)	(4.663)
Stateless:I(long3)	-0.943	-0.303	37.241	87.194*	-41.753	1.046	61.276	1.364	21.677
	(1.529)	(0.931)	(32.728)	(47.702)	(61.267)	(8.142)	(46.004)	(1.349)	(28.378)
long:Control	-1.127***	-0.164	3.941	-11.906	-70.013***	1.432	44.910***	0.511	2.803
	(0.389)	(0.261)	(8.253)	(15.368)	(15.625)	(2.959)	(11.940)	(0.747)	(7.307)
lat:Control	-0.616	-0.514**	14.839**	-28.179*	-7.312	-13.959***	-15.054	-1.152*	0.712
	(0.382)	(0.246)	(7.537)	(14.733)	(13.724)	(2.904)	(10.422)	(0.600)	(5.119)
I(long2):Control	-1.097	0.753	-5.134	35.502	-22.775	6.729	68.851**	3.451	8.594
	(0.941)	(0.721)	(16.563)	(35.549)	(34.067)	(10.151)	(27.657)	(2.313)	(14.466)
I(lat2):Control	0.306	0.401	-10.575	38.115**	-1.785	3.209	13.678	1.603	1.862
	(0.464)	(0.380)	(8.330)	(17.439)	(16.870)	(5.202)	(12.252)	(1.275)	(6.290)
I(long3):Control	0.458	0.984	15.380	35.523	47.777	0.975	-5.278	4.104**	-5.099
	(0.909)	(0.653)	(17.369)	(34.670)	(31.475)	(8.705)	(26.442)	(2.013)	(13.650)
Control:I(lat3)	1.273***	0.127	-10.770	42.466**	30.516**	15.028***	-11.666	-0.420	5.748
	(0.467)	(0.348)	(9.403)	(17.262)	(14.813)	(4.492)	(11.741)	(1.019)	(5.471)
Stateless:long:lat	-0.723	0.276	-15.534	68.430*	-118.121***	1.651	39.843	1.732*	15.930
	(1.103)	(0.669)	(20.893)	(37.247)	(42.357)	(8.317)	(29.381)	(0.992)	(16.054)
Stateless:lat:I(long2)	-2.976	-1.858	-54.139	-143.214*	5.727	-17.365	-48.190	-3.310	-3.323
	(2.843)	(1.827)	(49.207)	(85.862)	(106.240)	(19.563)	(76.424)	(2.413)	(48.582)

Table A5 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Stateless:long:I(lat2)	1.846	0.690	6.350	32.165	-57.984	9.405	7.359	0.256	9.169
	(1.461)	(0.907)	(23.698)	(43.558)	(52.394)	(11.631)	(38.144)	(1.344)	(22.534)
long:lat:Control	0.664	-1.132	2.346	-30.733	11.331	-5.679	-71.460**	-5.265	-1.222
	(1.253)	(1.023)	(23.463)	(47.283)	(45.730)	(14.216)	(35.427)	(3.435)	(18.953)
lat:I(long2):Control	1.472	-2.130	-11.188	7.709	20.903	7.614	-67.074	-8.931*	12.267
	(2.090)	(1.627)	(39.323)	(77.138)	(67.016)	(23.127)	(55.540)	(5.262)	(27.839)
long:I(lat2):Control	-1.858	1.136	-2.225	-72.876	-29.311	-18.686	54.948	5.349	-10.814
	(1.493)	(1.228)	(27.845)	(53.133)	(43.632)	(17.604)	(36.024)	(4.103)	(16.162)
Constant	0.233	-1.230***	70.215***	91.168***	69.758***	11.062***	18.491	0.415	23.798***
	(0.448)	(0.244)	(10.468)	(18.862)	(18.999)	(2.334)	(13.895)	(0.424)	(8.931)
Observations	1,908	1,908	1,908	1,908	1,908	1,908	1,908	1,908	1,908
Adjusted R^2	0.317	0.508	0.028	0.180	0.295	0.414	0.161	0.305	0.261
Residual Std. Error	0.921	0.565	20.302	33.645	35.074	7.703	26.082	1.155	14.806

*p<0.1; **p<0.05; ***p<0.01

Table A6: RD regression models with cubic polynomial longitude and latitude and controls for geographic variables

	Dependent variable:								
	Income index	Education-based income	Electricity network	Sanitation network	Water network	Managers	Farmers	University educated	Illiteracy rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Stateless	-0.500***	-0.409***	-7.338***	-0.903	-26.805***	-3.589***	7.743**	-0.421***	13.541***
	(0.112)	(0.067)	(2.673)	(4.080)	(4.313)	(0.806)	(3.379)	(0.157)	(1.959)
log(to_Aleppo)	-0.490***	-0.055	4.725*	-23.100***	-3.729	-1.779**	9.849**	-0.136	0.165
	(0.125)	(0.068)	(2.809)	(5.009)	(5.101)	(0.728)	(3.860)	(0.131)	(2.367)
Slope	-0.018***	-0.009***	0.104	-0.688***	-0.375**	-0.078	0.475***	-0.001	0.325***
	(0.005)	(0.003)	(0.087)	(0.178)	(0.172)	(0.052)	(0.128)	(0.008)	(0.093)
Elevation	0.001*	0.0003*	0.004	-0.013	-0.003	0.001	-0.009	0.001***	0.002
	(0.0003)	(0.0002)	(0.006)	(0.009)	(0.009)	(0.003)	(0.007)	(0.0004)	(0.004)
Temperature	-0.354***	-0.052	-0.393	-1.274	-0.192	-1.707**	11.347***	0.017	0.591
	(0.078)	(0.049)	(1.889)	(2.834)	(2.680)	(0.690)	(1.917)	(0.098)	(1.115)
Precipitation	0.001	0.001	-0.073***	-0.218***	-0.035	0.061***	0.044	-0.002	-0.027*
	(0.001)	(0.001)	(0.026)	(0.045)	(0.040)	(0.014)	(0.032)	(0.002)	(0.015)
Flow accumulation	-0.00001**	0.00000	0.00003	-0.00003	0.0001	-0.00005	0.0002**	-0.00000	-0.00002
	(0.00000)	(0.00000)	(0.0001)	(0.0002)	(0.0002)	(0.00004)	(0.0001)	(0.00001)	(0.0001)
Stateless:long_d	-0.751***	-0.547***	-31.095***	-10.987	-43.756***	7.412***	22.488***	-0.990***	17.227***
	(0.289)	(0.185)	(6.651)	(10.124)	(12.152)	(1.831)	(8.531)	(0.294)	(5.114)

Table A6 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Stateless:lat_d	-0.909***	-0.680***	20.556***	-23.125*	11.294	-14.023***	4.788	-0.398	2.667
	(0.336)	(0.202)	(7.527)	(12.531)	(13.996)	(2.177)	(9.808)	(0.339)	(5.579)
Stateless:I(long_d2)	1.921*	1.335**	10.368	-12.633	173.235***	-1.267	-42.243	-0.196	-34.851*
	(1.039)	(0.615)	(23.421)	(34.136)	(42.726)	(5.466)	(31.991)	(1.063)	(18.364)
Stateless:I(lat_d2)	0.202	0.222	15.520**	14.597	41.007***	-1.586	2.193	0.072	-17.876***
	(0.388)	(0.235)	(7.655)	(11.714)	(13.527)	(3.037)	(9.824)	(0.349)	(4.923)
Stateless:I(long_d3)	0.987	-0.138	47.430	75.874	-39.012	4.029	-0.072	1.900	18.839
	(1.559)	(0.981)	(32.642)	(48.635)	(61.571)	(8.612)	(46.025)	(1.480)	(28.336)
long_d:Control	-0.427	0.408	-4.559	-23.642	-73.894***	10.151***	32.243***	1.330*	1.093
	(0.411)	(0.295)	(8.055)	(15.817)	(15.539)	(3.427)	(12.219)	(0.804)	(7.472)
lat_d:Control	-1.814***	-1.250***	12.085	-29.157**	-13.984	-22.635***	10.179	-2.237***	7.832
	(0.398)	(0.276)	(7.785)	(14.873)	(13.923)	(3.309)	(10.553)	(0.645)	(5.436)
I(long_d2):Control	-1.243	1.449	33.451*	152.109***	10.985	-17.403	63.976*	6.913***	8.188
	(1.193)	(0.891)	(19.601)	(41.577)	(39.731)	(12.167)	(33.986)	(2.638)	(17.231)
I(lat_d2):Control	-0.015	0.556	5.721	78.501***	13.698	-6.697	20.087	2.693*	-0.600
	(0.528)	(0.445)	(9.151)	(19.496)	(18.878)	(5.822)	(13.825)	(1.401)	(7.127)
I(long_d3):Control	-0.487	0.810	38.228**	92.847**	64.121*	-17.573*	11.999	5.076**	-4.182
	(0.984)	(0.707)	(17.179)	(36.612)	(33.133)	(9.303)	(29.064)	(2.085)	(14.543)
Control:I(lat_d3)	2.714***	1.009***	-11.762	36.207**	39.776***	27.821***	-41.251***	0.522	-5.653
	(0.479)	(0.368)	(9.618)	(17.655)	(15.376)	(4.983)	(11.794)	(1.084)	(6.227)

Table A6 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Stateless:long_d:lat_d	0.922	0.686	-1.337	73.452*	-107.873***	4.354	-9.536	2.667***	8.469
	(1.120)	(0.700)	(21.574)	(37.493)	(41.795)	(8.321)	(29.451)	(1.026)	(15.911)
Stateless:lat_d:I(long_d2)	-4.370	-1.777	-79.273	-158.128*	-6.621	-8.285	6.184	-4.139*	0.232
	(2.869)	(1.879)	(49.795)	(85.222)	(105.066)	(19.548)	(75.038)	(2.465)	(48.470)
Stateless:long_d:I(lat_d2)	2.882**	1.019	28.647	57.094	-39.893	6.978	-25.715	1.212	0.519
	(1.438)	(0.919)	(24.360)	(43.019)	(51.325)	(11.194)	(36.469)	(1.338)	(22.506)
long_d:lat_d:Control	1.375	-1.340	-33.811	-125.850**	-18.904	21.178	-80.308**	-7.892**	-1.483
	(1.449)	(1.193)	(25.392)	(52.107)	(50.663)	(15.684)	(40.550)	(3.731)	(21.340)
lat_d:I(long_d2):Control	3.655	-1.883	-85.986**	-210.771**	-35.007	66.972**	-95.514	-12.781**	3.120
	(2.490)	(1.903)	(43.866)	(88.714)	(78.429)	(26.396)	(67.055)	(5.791)	(33.675)
long_d:I(lat_d2):Control	-4.441**	-0.219	44.061	53.082	-11.693	-66.670***	94.641**	6.009	9.388
	(1.738)	(1.416)	(31.061)	(61.185)	(52.032)	(19.828)	(42.741)	(4.489)	(21.026)
Constant	7.009***	-0.135	90.924**	163.611***	79.725	35.155**	-204.279***	0.896	16.027
	(1.592)	(0.992)	(40.893)	(58.041)	(56.421)	(13.712)	(40.404)	(1.893)	(24.634)
Observations	1,907	1,907	1,907	1,907	1,907	1,907	1,907	1,907	1,907
Adjusted R^2	0.288	0.454	0.023	0.176	0.287	0.361	0.156	0.229	0.249
Residual Std. Error	0.940	0.595	20.357	33.730	35.263	8.044	26.161	1.216	14.932

*p<0.1; **p<0.05; ***p<0.01

Table A7: RD effects on informality in property ownership

	Dependent variable:					
	Official deed		Agricultural deed		Public notary	
	(1)	(2)	(3)	(4)	(5)	(6)
Stateless	7.376*	6.517	-4.981	-4.410	-0.435	-0.487
	(3.885)	(3.969)	(4.113)	(4.185)	(0.751)	(0.739)
Ethnicity	-	Y	-	Y	-	Y
Mean Y (Control)	37.894	37.894	20.615	20.615	1.105	1.105
SD	36.521	36.521	34.773	34.773	5.460	5.460
Observations	1,908	1,908	1,908	1,908	1,908	1,908

*p<0.1; **p<0.05; ***p<0.01

Note: The table shows the regression discontinuity effect on the share of properties owned according to official deeds, agricultural deeds, and public notary deeds, using a cubic polynomial function in longitude and latitude.

Table A8: Effects of institutional variables on development outcomes

	Dependent variable:								
	Income index	Education-based income index	Electricity network	Sanitation network	Water network	Managers	Farmers	University educated	Illiteracy rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Official deed	0.003***	0.002***	0.150***	0.329***	0.215***	0.007	-0.067***	0.005***	-0.031***
	(0.001)	(0.0005)	(0.015)	(0.026)	(0.027)	(0.007)	(0.020)	(0.001)	(0.011)
Agricultural deed	0.002**	-0.0001	0.117***	0.015	-0.029	0.002	-0.022	-0.0001	0.009
	(0.001)	(0.0004)	(0.015)	(0.020)	(0.027)	(0.006)	(0.021)	(0.001)	(0.012)
Public notary	0.002	0.009***	0.188***	0.835***	0.619***	0.027	-0.028	0.008**	-0.252***
	(0.003)	(0.002)	(0.038)	(0.172)	(0.114)	(0.018)	(0.104)	(0.004)	(0.068)
Observations	1,908	1,908	1,908	1,908	1,908	1,908	1,908	1,908	1,908

*p<0.1; **p<0.05; ***p<0.01

Note: The table shows the effects of institutional variables on development outcomes, treating the study region as one area and controlling for linear terms in longitude and latitude.

Table A9: RD effects on age structure controlling for fertility rates

	Dependent variable:					
	Males under 15	Males 15 to 64	Males 65 and over	Males/females under 15	Males/females 15 to 64	Males/females 65 and over
	(1)	(2)	(3)	(4)	(5)	(6)
Stateless	2.215***	-1.966***	-0.249**	0.029**	-0.044**	0.126*
	(0.356)	(0.332)	(0.120)	(0.011)	(0.020)	(0.067)
Mean Y (Control)	43.012	53.174	3.813	1.081	1.027	1.412
SD	7.314	6.174	2.459	0.153	0.311	0.832
Observations	1,872	1,872	1,872	1,872	1,872	1,856

*p<0.1; **p<0.05; ***p<0.01

Note: The table shows the RD effects on the proportion of each age group within the male population (columns 1 to 3), and the RD effects on the sex ratio within each age group(columns 4 to 6), controlling for linear terms in longitude and latitude.

Table A10: RD effects on age structure in the male population and the sex ratio according to age (Sunni Muslim sample)

	Dependent variable:					
	Males under 15	Males 15 to 64	Males 65 and over	Males/females under 15	Males/females 15 to 64	Males/females 65 and over
	(1)	(2)	(3)	(4)	(5)	(6)
Stateless	1.416***	-1.575***	0.158	0.025**	-0.046**	0.068
	(0.381)	(0.362)	(0.118)	(0.012)	(0.018)	(0.075)
Mean Y (Control)	45.821	50.915	3.262	1.079	1.028	1.486
SD	5.440	5.006	1.839	0.177	0.281	1.021
Observations	1,406	1,406	1,393	1,406	1,406	1,393

*p<0.1; **p<0.05; ***p<0.01

Note: The table shows the RD effects on the proportion of each age group within the male population (columns 1 to 3), and the RD effects on the sex ratio within each age group (columns 4 to 6), controlling for linear terms in longitude and latitude. The sample excludes non-Sunni Muslim groups and adds a dummy variable for areas with a tribal majority.

Figure A1: Bandwidth sensitivity of the effect of statelessness on outcomes

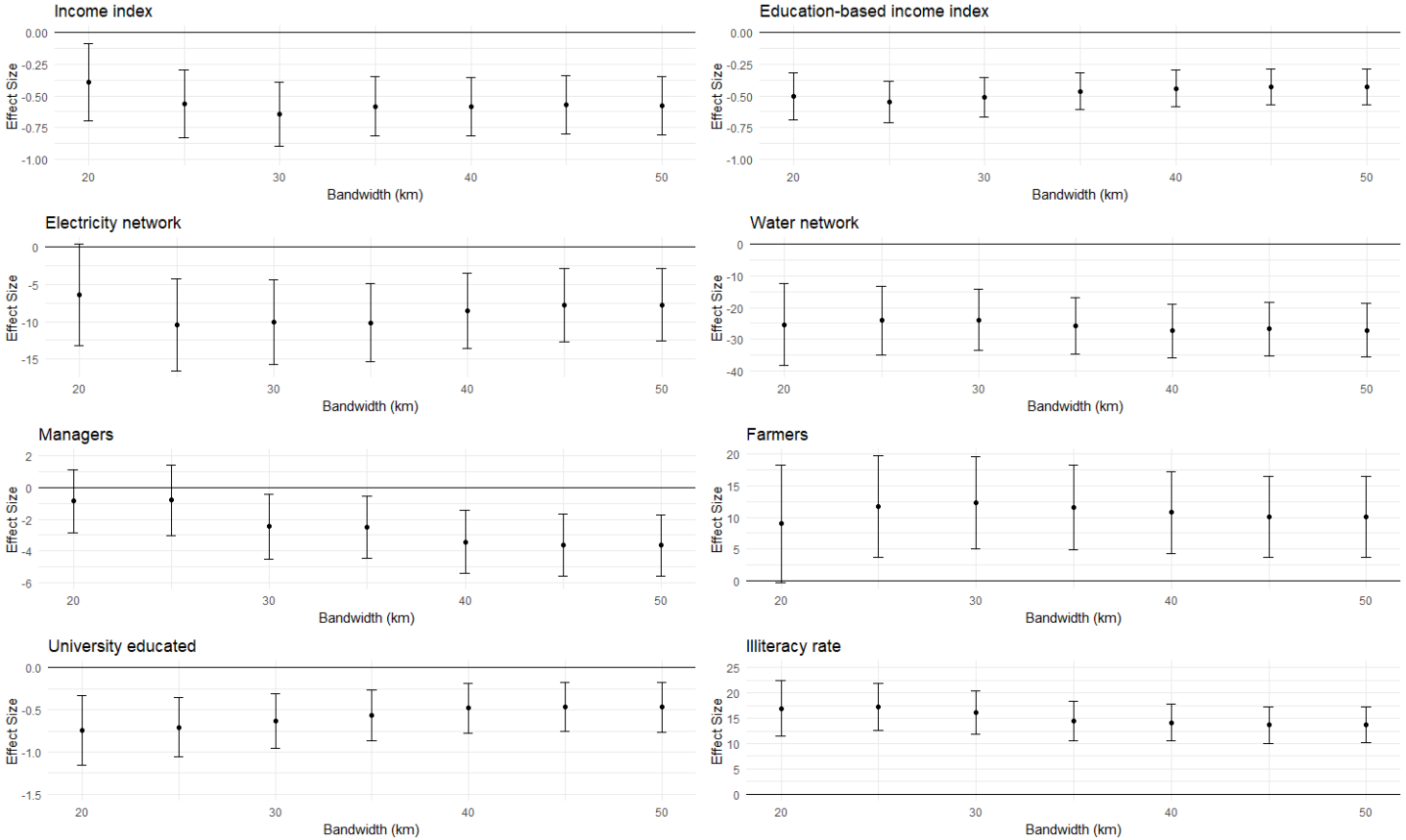
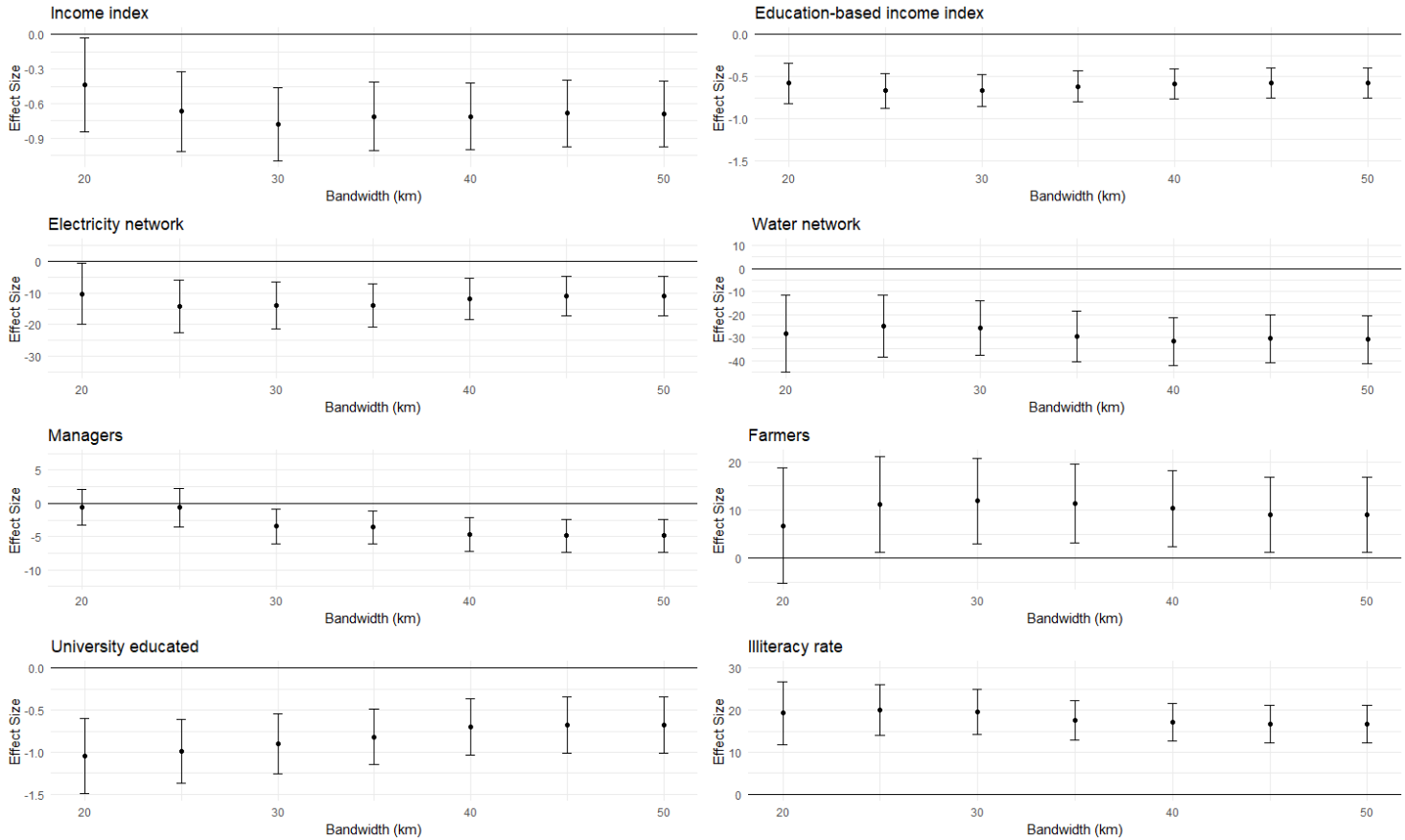


Figure A2: Bandwidth sensitivity with a 5k donut



Paper II

Journal of Demographic Economics (2025), 1–33
doi:10.1017/dem.2025.10007

CAMBRIDGE UNIVERSITY PRESS

RESEARCH PAPER

Ethnic and religious differences in female labor force participation: evidence from Syrian census data

Rami Zalfou

Department of Economic History, Centre for Economic Demography, and Centre for Advanced Middle Eastern Studies, Lund University, Lund, Sweden
Email: rami.zalfou@ekh.lu.se

(Received 11 July 2024; revised 28 July 2025; accepted 30 July 2025)

Abstract

This paper investigates the magnitude and drivers of ethnic and religious disparities in Female Labor Force Participation (FLFP) in Syria. Using census data and the Gelbach decomposition method, the analysis reveals substantial FLFP gaps across groups, reaching up to 18 percentage points. To understand the factors underlying these gaps, the analysis exploits the heterogeneity in economic development, demographic profiles, educational attainment, and geography across towns and villages. The findings indicate that differences in age distribution, income levels, education, and public sector employment shares are the primary contributors to FLFP gaps. While social norms are often considered key determinants of FLFP, their role is assessed using gender parity indices for educational enrollment and attainment but shows little explanatory power for ethnic and religious disparities in FLFP. These results highlight the importance of structural economic and demographic factors over cultural constraints in shaping group differences in FLFP in Syria.

Keywords: female labor force participation (FLFP); ethnic and religious disparities; Syrian demographics; economic development; Gelbach decomposition

1. Introduction

The conventional narrative surrounding low female labor force participation (FLFP) rates in the Middle East and North Africa (MENA) region attributes these rates to cultural and social norms (Dildar, 2015; Giuliano, 2017), particularly those related to Islam (Algan & Cahuc, 2006; Bayanpour-Tehrani & Sylwester, 2013; Kızılca, 2016; Atasoy, 2017; Akyol & Ökten, 2024). However, by focusing on religious and cultural norms, existing analyses often overlook critical socio-economic factors and regional disparities that are instrumental in shaping FLFP. This paper delves deeper into these factors by examining FLFP gaps among ethnic and religious groups in Syria and their determinants.

Syria represents an interesting case study as a country with low FLFP emblematic of the MENA region, having reached 26% in 2021 from a low of 13% in 2010 (World Bank, 2022), and with a Muslim majority and a diverse cultural landscape that includes indigenous Christian groups and various Muslim sects and ethnic minorities.

At the country level, the religion of Islam is associated with low FLFP (Ross, 2008; Bayanpour-Tehrani & Sylwester, 2013). This paper documents substantial within-country gaps in FLFP across different religions and ethnic groups in the context of a Muslim-majority country. In the population census of 2004, Syrian towns and villages with Christian majorities are found to have an average FLFP rate of 26%, as compared to Sunni Muslim majority areas where the FLFP rate is at only 10%. A similar gap is found between Islamic sects as well, with Alawi-majority areas having an FLFP rate of 28%.

It remains debated whether the correlation between Islam and low FLFP is due to the influence of social norms specific to Muslims, such as restrictions on women's freedoms and particular Islamic views on gender roles, or if it is due to structural economic factors specific to countries of the MENA (Ross, 2008; Assaad, 2014; Assaad et al., 2020). MENA economies share certain common characteristics such as large and stagnant public sector shares, weak private sectors, and a prevalence of informal and low-quality jobs that may have contributed to the low and stagnant FLFP rates. Within-country differences in these structural factors may also shape the gaps in FLFP between ethnic and religious groups. Furthermore, previous empirical work on FLFP documents the presence of a U-shaped relationship between economic development and FLFP (Goldin, 1995; Mancini, 2018; Tunalı et al., 2021). As a consequence, within-country differences in economic development may also shape FLFP gaps between ethnic and religious groups, when large regional inequalities are present within the same country.

To explore competing hypotheses, the paper makes use of census data aggregated at the local level and the Gelbach decomposition method which allows for a comparative approach that can determine the relative importance of a range of factors affecting group differences in FLFP. The census presents a snapshot of the country prior to the civil war which began in 2011 and provides a view of large inequalities across the country in terms of economic development and of diversity in measures of income, demography, and human capital. To account for differences in the nature of women's work in such a diverse context, the analysis makes use of an expanded definition of FLFP, which includes both employment in the modern sector as well as work in the family sector which remains prevalent in many agricultural areas. The analysis explores the roles of economic and demographic development, as well as the dynamics of the family sector and the public sector. To explore the role of social norms in shaping FLFP gaps across ethnic and religious groups, the analysis makes use of Gender Gap Indices (GPI's) that measure the gender gap in school enrollment among school-aged children, and the gender gaps in the educational attainment of adults at the secondary level and above, and at the tertiary levels. These indices proxy for bias against women associated with the under-investment in girls' education as compared to boys. The analysis shows that these indices exert a weak influence on FLFP, and only contribute in small part to explaining gaps in FLFP across groups, accounting for only 10% of the FLFP premiums in Alawi and Christian communities.

The findings from the decomposition analysis suggest that the main contributing factors to ethnic and religious gaps in FLFP are the age structure of the population, income, and the public sector employment share, alongside educational attainment. Regional inequality contributes to the observed gaps in FLFP, with industrial regions having the lowest FLFP, while rates are elevated in both agricultural regions and in areas where the service economy predominates. The findings support the idea that economic

development shapes the gaps in FLFP between groups in accordance with the U-feminization hypothesis. Meanwhile, elevated FLFP in certain minorities like the Alawis correlates with high public sector employment shares. The public sector in the MENA context is an important share of employment for politically important groups (Assaad, 2014), and its contribution to ethnic and religious gaps in FLFP has not been studied previously. Differences in educational attainment between groups also contribute to shaping FLFP, especially amongst the Christian communities which have a legacy of early progress in educational attainment (Saleh, 2016). On the other hand, differences in the age structure between ethnic and religious groups contribute significantly to the observed gaps in FLFP, explaining over a third of the FLFP premium held by Christian and Alawi women. The important role of the age structure, accompanied by an insignificant role for fertility rates, highlights the role of the demographic dividend in shaping FLFP.

The paper contributes to the debate on religious and ethnic identity, social and cultural norms, and gender inequality. The dynamics of the labor force, the population structure, and educational attainment are crucial to explaining the existing ethnic and religious gaps in FLFP, while observed social norms could not account for these gaps. In exploring these determinants, the study provides a more nuanced understanding of the interplay between culture, gender, and labor force dynamics.

2. Theory and previous research

Religious and ethnic identity can impact FLFP through several mechanisms. One key channel is the influence of social norms, which are shared expectations and informal rules that govern behavior within a group. These norms can shape attitudes toward women's work and influence both individual choices and structural constraints in the labor market. Religious doctrines, for instance, may prescribe specific roles for women, reinforcing expectations that prioritize domestic responsibilities over workforce participation. In particular, norms related to gender segregation or the belief that women should occupy a subordinate role in the home or public sphere may create significant barriers to FLFP. Because norms are socially constructed and transmitted across generations, they often vary across religious and ethnic groups.

Empirical research has shown that countries with Muslim majorities tend to have significantly lower FLFP compared to the global average (Tzannatos, 1999). Several studies have examined the role of religiosity in shaping these patterns (Dildar, 2015; Ugur, 2018). Differences in social norms across religious sects can also contribute to variations in FLFP. For example, Akyol and Ökten (2024) argue that more gender-equal views among Alevi Muslims in Turkey led to an FLFP premium of about twelve percentage points in this group compared to their Sunni Muslim counterparts, even after controlling for observed characteristics. Likewise, ethnic cultures may carry deeply ingrained beliefs about gender roles. In the MENA region, for example, social norms specific to Arab culture have been theorized as a major factor in explaining low FLFP rates (Korotayev et al., 2015). These findings highlight how religious and ethnic norms can act as both direct constraints on FLFP and as broader structural forces that shape societal expectations and labor market dynamics.

On the other hand, group inequalities in socio-economic status or in access to education may also influence FLFP. Different ethnic and religious groups may face different economic opportunities, particularly if these groups are geographically segregated, or when institutional barriers or discrimination affects certain groups.

Differences in economic development across areas can also shape FLFP rates in line with the U-feminization hypothesis. Other variables that influence FLFP in a general sense may also contribute to differences in FLFP across groups, such as the age distribution of the population, when differences in FLFP are present across age groups concurrently with differences in the age distributions of different religious/ethnic groups.

Early research into ethnic and religious differences in FLFP employed data from developed countries such as the US. Reimers (1985) explores the ethnic gaps in FLFP in the US and finds that these can be mostly explained by observed characteristics such as age, education level, and family size. Similarly, Lehrer (1995) studies FLFP gaps between Christian denominations in the US and finds that the differences between denominations are partly but not fully accounted for by observed characteristics. More recent studies include cross-country analyses, as well as analyses that focus on developing country contexts. Ross (2008) theorizes that low FLFP in Muslim countries is due to the presence of oil rents, which reduces FLFP through the growth of non-tradable sectors (which include construction and retail services) and through high household incomes which increase women's reservation wages. While cross-country regressions do not find a strong correlation between fuel exports and FLFP (Bayanpour-Tehrani & Sylwester, 2013), the nature of Middle Eastern economies can contribute to the observed low FLFP. Demand-side factors, such as the shrinking public sector (Assaad et al., 2020), and dualism in the labor market where public sector jobs are highly sought after while outside options are limited to low-quality informal jobs, are additional constraints on FLFP (Assaad, 2014). Gender segregation within the workplace may also shape the demand for female workers as Chatterjee and Vanneman (2022) show in the case of India. In a similar vein, İlkkaracan (2012) suggests that the lack of a demand-side challenge to breadwinner model resulted in the institutionalization of the gendered labor division in Turkey.

Gaps in economic development across regions may contribute to FLFP gaps between ethnic and religious groups, especially when these groups are geographically segregated. A substantial body of literature explores the impact of economic development on FLFP, frequently employing the feminization U hypothesis (Boserup, 1970; Goldin, 1995). The hypothesis suggests a pattern in FLFP across different stages of economic development. In agricultural economies, FLFP is generally high as women often participate in farming and related activities. However, as these economies transition to industrialization, FLFP tends to decline. This decline is attributed to the shift toward factory jobs and other industrial work, which historically have been less accessible or less culturally acceptable for women. As economies further evolve into post-industrial stages, focusing more on services and knowledge-based industries, FLFP begins to increase again. These sectors often offer more opportunities and greater flexibility for women, leading to higher FLFP. Research focused on developed countries shows the presence of long-term secular increase in female labor force participation over the course of the 20th century (Stanfors & Goldscheider, 2017; Mancini, 2018). Changing cultural norms and the family division of labor are part of this story, while structural transformation may have also played an important role (Olivetti & Petrongolo, 2016). While this theory has been used to account for differences in FLFP across countries and across time, it has seldom been used to address FLFP gaps within countries or across ethnic and religion groups.

Syria's complex social fabric and its historical progression from an agriculture-based economy to a more diverse economy[1] have resulted in differential economic

[1]The share of workers employed in the agricultural sector declined in Syria from around 70% in the middle of the 20th century (World Bank 1955) to 15% according to the population census of 2004.

opportunities across groups, possibly impacting FLFP. The different FLFP rates across groups could be shaped by differentials in labor market conditions and in economic development as suggested by the feminization U hypothesis. Economic development can shape FLFP gaps, not only through regional differences in the nature of the economy and available opportunities for women, but also through differences in education attainment across groups, and the demographic composition of different communities, as well as through changing norms that evolve in the process of economic development.

This study contributes to the literature on ethnic and religious disparities in FLFP by focusing on Syria, a country that exhibits both the broader Middle Eastern pattern of low FLFP and significant heterogeneity in labor market outcomes across religious and ethnic groups. While much of the existing research attributes low FLFP in the Middle East and North Africa (MENA) region to cultural and religious norms (Algan & Cahuc, 2006; Bayanpour-Tehrani & Sylwester, 2013; Dildar, 2015; Kızılca, 2016; Atasoy, 2017; Ugur, 2018; Akyol & Ökten, 2024), this paper shifts the focus to structural economic and demographic determinants of FLFP disparities within a single Muslim-majority country. The analysis will examine the contributions of various factors to explaining ethnic and religious gaps in FLFP, including economic and demographic structure, and cultural norms. This paper takes an exploratory approach to assessing the relative importance of these factors. Rather than assuming a dominant explanatory mechanism, the analysis tests the strength of association between these factors and FLFP gaps, allowing for an empirical evaluation of whether structural economic and demographic conditions outweigh the influence of cultural and social norms.

3. Ethnic and religious differences in Syria

Syria stands out for its extraordinary ethnic and religious diversity, often described as a "mosaic society" to emphasize the coexistence of various cultural groups (Hinnebusch, 2004). This diversity is deeply rooted in the country's geography, which has long facilitated migratory movements and reinforced internal fragmentation. Historically, the region that is now Syria was part of the Ottoman Empire until the First World War and later came under French rule before gaining independence in 1945. Since the early 20th century, Syria has undergone significant economic and social transformations. These include state-led industrialization starting in the mid-20th century, followed by economic liberalization in the late 20th century, alongside rising regional economic inequality in the decades leading up to the recent civil war (Hassine, 2015).

Syria provides a uniquely valuable case study for analyzing FLFP in relation to religious, ethnic, and economic factors. While Syria's FLFP rates are low, in line with other Middle Eastern countries, its diversity and pronounced regional economic inequalities offer a rich spectrum of opportunities and constraints for women. This combination allows for a deeper exploration of how cultural norms and economic conditions shape women's participation in the labor market. As a Middle Eastern country, Syria has a Muslim and an Arabic-speaking majority, and its low levels of FLFP make it a microcosm for examining mechanisms in the broader region. Using data from the pre-civil war period allows for comparisons across groups in the same institutional context prior to the cantonization brought on by the war which starting in 2011.

Figure 1 shows a map of Syrian cities, towns, and villages according to ethnic majority. The analysis in this paper follows the definition of ethnic identity which relates this identity to claims of shared ancestry and the cultural practices associated with those claims (Khaddour & Mazur, 2018). Ethnic identity is used here as an umbrella term for

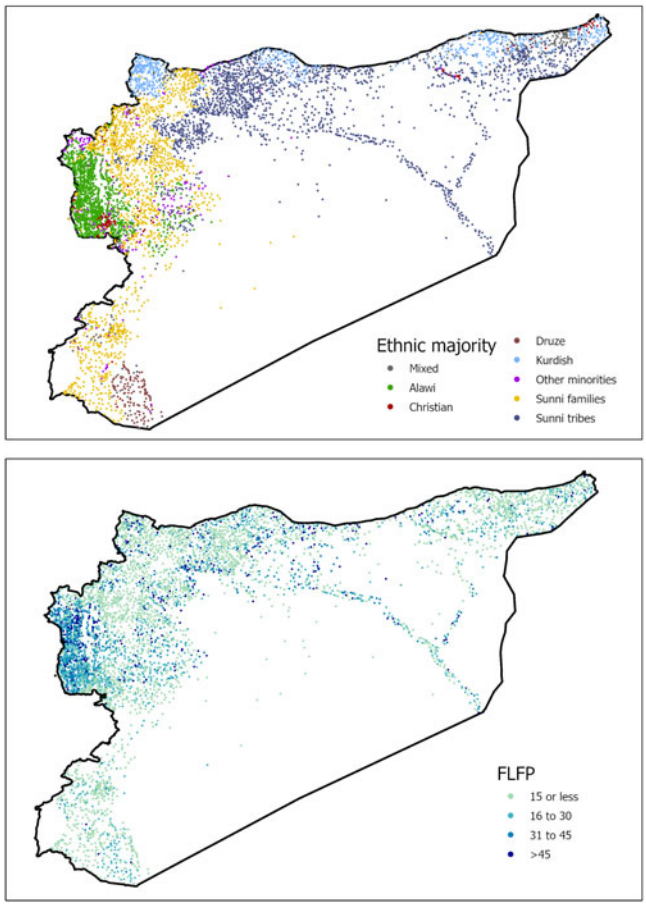


Figure 1. Ethnic composition and female labor force participation (FLFP).

religious, sectarian, and national differences. The ethnic majority group in Syria is typically conceptualized as the Arab Sunni Muslim population. However, Khaddour and Mazur (2018) suggest a more nuanced view, separating this group into tribal and non-tribal groups, employing the terms "Sunni tribes" and "Sunni families" – a terminology adopted in this paper. The Sunni tribes occupy the north-east of the country and have a distinct culture despite sharing the same language and religious affiliation with the Sunni families in the west. Apart from Sunni Muslims, major ethnic groups in Syria include Alawi Muslims, Druzes, Kurds, and Christians, as well as smaller groups such as Shia Muslims, Ismaili Muslims, and Circassians who are all aggregated under the category "Other minorities." Most of the cities have mixed populations and are categorized as such, while smaller towns and villages are mostly homogenous with a clear majority group in each.

Most Sunni Muslims in Syria follow the Hanafi school of Islamic jurisprudence, noted for its relative adaptability compared to other Islamic legal traditions. Still, religion holds a prominent place in daily life and may shape individuals' viewpoints and

preferences. Several other Muslim and non-Muslim groups are present alongside the Sunni Muslim majority. Antoun (1991) emphasizes the distinctiveness of Syria's religious sects, which typically occupy peripheral locations, cluster demographically in territorial "homelands," and are exclusive in culture and social organization. Muslim sects like Alawis and Druzes, alongside Christians, do not follow the practices that may be common amongst Sunni Muslims, like gender segregation, such as wearing the veil which is common amongst Sunni Muslim women. Cultural and social norms of this kind may play a role in determining FLFP within each group. Sect groups in Syria not only have distinct informal norms, but they also maintain separate courts that deal with matters of personal law such as marriage and divorce, and inheritance (Van Eijk, 2016), a practice with deep historical roots in the MENA region.

Other groups like the Sunni tribal communities, Kurds, and other minorities like the Ismailis, Circassians, and Armenians all tend to have distinct social and cultural norms that distinguish each of them. Even though communities like Sunni tribes, Kurds, and Circassians, largely follow a similar form of Sunni Islam, the distinct norms in these communities may shape the lives of women and may affect their propensity to enter the labor market, based on views prevalent in each community regarding women's proper role in society.

Ethnic and religious inequalities in socio-economic outcomes remain under-studied in this context. However, studies that address neighboring countries can be useful to inform our expectations as many ethnic communities cut across national borders in the region. An early study of religious gaps in fertility in Lebanon found lower fertility among Christian as compared to Muslims, as well as differences between Muslim and Christian sects, but these appeared to converge at higher levels of education (Chamie, 1977). Differentials in education levels between religious groups in the region are long-standing, with Christian school enrollment overtaking Muslim enrollment in the early 20th century (Sassmannshausen, 2019). Under the French mandate of Syria and Lebanon (1923–1946), private and missionary school were mostly attended by Christian students, while public schools served a majority of Muslim students (Abi-Rached & Diwan, 2022), which suggests differences in access to education across groups. The post-colonial period saw the establishment of modern primary schools, which coincided with an increase in Muslim education levels in Egypt, and the beginning of a process of convergence in Muslim and Christian educational levels and labor market outcomes in that country (Saleh, 2016). Still, Christians in the MENA have overall higher incomes, higher educational attainment, and lower fertility than Muslims. Several studies draw on data from Turkey, where religiosity is found to have a negative association with FLFP (Dildar, 2015; Ugur, 2018). Additionally, Alevi Muslim women in Turkey have higher participation than their Sunni Muslim counterparts (Akyol & Ökten, 2024). Some studies also assess religious differences in FLFP among Arab women in Israel (e.g., Khattab, 2002; Yonay et al., 2015) finding significant gaps between Christian, Druze, and Sunni Muslim women in that context.

Courbage and Todd (2014) address the fertility gaps among ethnic groups in Syria, focusing on differences in fertility between Alawis, Sunni Muslims, and the tribal regions of the north-east. The authors propose that these divergences are a product of entrenched family structures. Specifically, they highlight that the Alawi community has preserved traditional matrilineal practices, while the Sunni tribes remain the most patrilineal among the nation's ethnic cohorts. Balanche (2015), on the other hand, links the diminished Alawi fertility with the relative advancement of the Alawi minority which commenced in the second half of the 20th century. Balanche posits that the

acceleration in the fertility transition was due to increased education levels and the induction of Alawi women into public sector employment. These contrasting theories on Alawi fertility can also be used to address women's entry into the labor force. Elevated FLFP in Alawi regions might stem from beliefs or traditions that inform preferences within this group. Alternatively, higher FLFP could be attributed to more recent factors pertaining to the labor market structure in Alawi areas. The same argumentation can be applied to addressing the Christian advantage in FLFP. The FLFP advantage may be due to differences in deep rooted beliefs among Christians and Sunni Muslims, or they may be a product of other factors including the demographic and economic structure at the local level.

Previous research has highlighted some implications of structural change for gender dynamics in the MENA and how the resulting changes in patterns of life and work affect men and women differently. Abdelali-Martini et al. (2003) describes the feminization of agricultural work which took place in MENA countries including Syria (see also Abdelali-Martini & Dey de Pryck, 2015). The increased availability of work outside of the agricultural sector led to an exodus of men into more productive and higher paid occupations, increasing the proportion of women among agricultural workers. This points to the spread of the male breadwinner model as a result of structural change. Female labor in the agricultural sector in the MENA includes both waged work in agricultural gangs such as those investigated by Abdelali-Martini, as well as traditional forms of work within the family unit which endures in many areas.

Figure 2 shows the sectoral composition of the labor force across the geography of Syria, which clarifies the inequalities discussed above. The north-east of Syria has a predominantly agricultural economy, while the secondary sector is largest in the western parts of the country apart from the coastal region, and the latter has a mainly service-based economy coinciding with the highest rates of employment in the public sector.

Based on the figure above and the U-feminization hypothesis, we would expect large differences to emerge in FLFP rates between groups in Syria. The first panel in Figure 3 shows FLFP in Syrian towns and villages as a function of income. The line of fit shows a clear U-shaped pattern, with FLFP decreasing as a function of income at low values of income, and then increasing at higher values of income. Areas with a majority of Sunni tribes have higher FLFP that is falling with income, while Sunni families appear to have reached the lowest FLFP rates in the mid-section of the U-curve. Meanwhile, Alawis and Christians have higher FLFP owing to their high average incomes and higher economic development and associated education levels in their areas.

Ethnicity-specific patterns can be seen more clearly in Figure 3b, which breaks down the same relationship according to ethnic group. The U-curve is most found clearly for Sunni families and Sunni tribes, where the U shape resembles that in the previous figure. The dashed line in Figure 3b shows a line of fit based on a similar model that adds a control variable for the share of the working population employed in the public sector, with the line of fit estimated at zero public sector shares for all groups. Controlling for public sector employment lowers the fitted values of FLFP in all groups, but its effect is seen most clearly in Alawi areas where a stark difference between the adjusted and unadjusted fit are found. Public sector employment may have shielded this group from the decline in FLFP that is found in Sunni majority areas. Meanwhile Kurdish FLFP appears to be falling with income, and Druze and Christian FLFP rates are rising with income. The limited data on these smaller groups and their geographic clustering may limit the extent to which the full U pattern can be seen. Similar results are found when

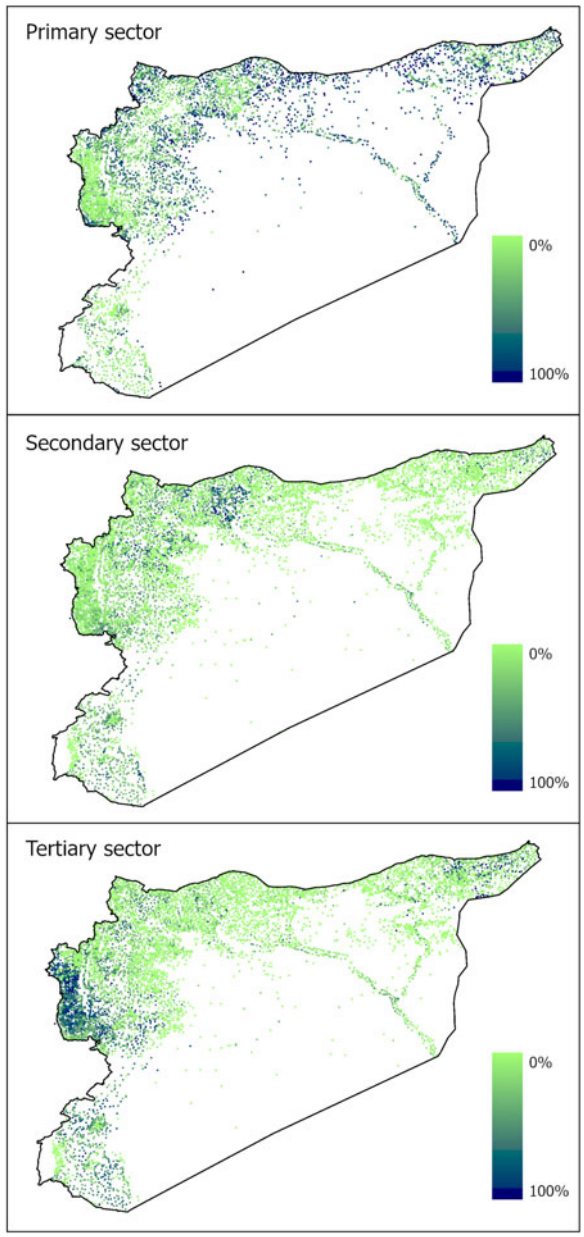


Figure 2. Labor force composition by sector at the local level in Syria.

using the alternative measure of income imputed using education data (See Appendix Figure 1). The finding that FLFP follows a U-shaped pattern as a function of income accords with the U-feminization hypothesis and justifies the use of second order polynomials in the regression equations in the next section.

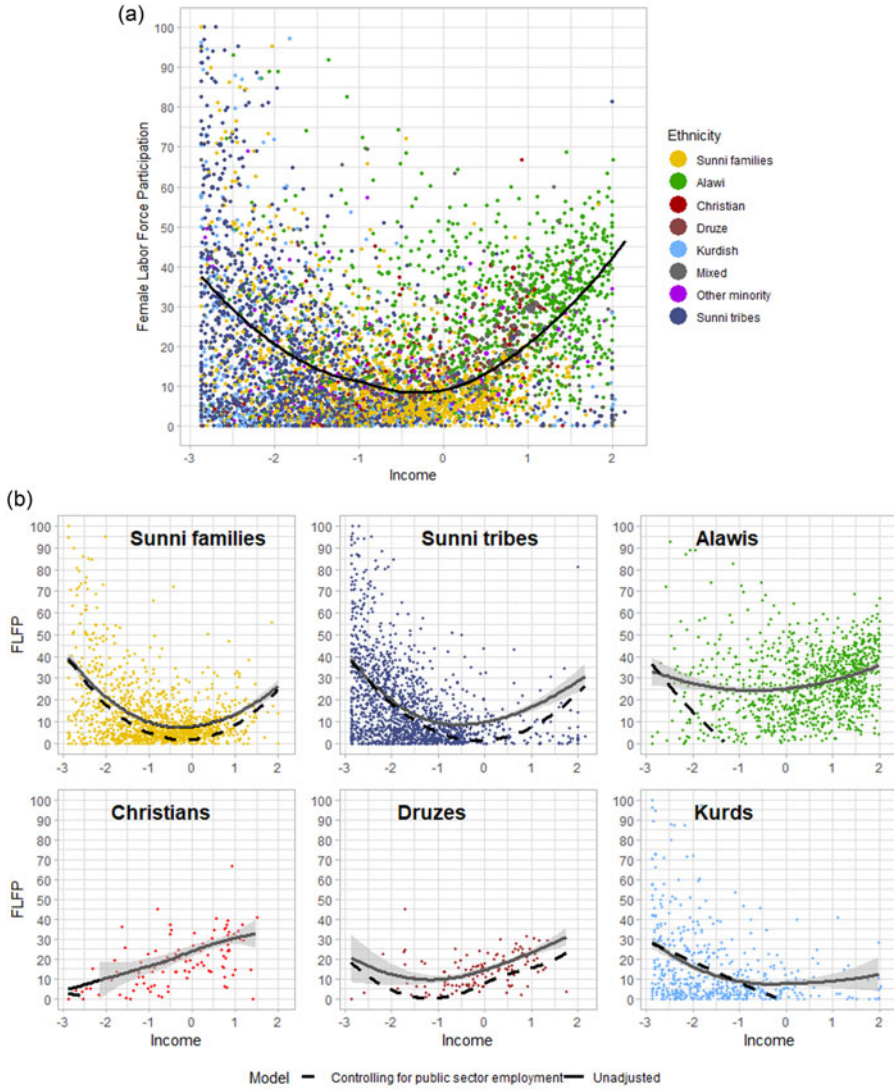


Figure 3. (a) Female labor force participation as a function of income. (b) Female labor force participation as a function of income by ethnic majority. *Note*: The income measure is imputed based on the sectoral composition of the labor force. The line of fit is estimated using a non-parametric loess function with observations weighed according to population size.

4. Data and methods

The data employed in the analysis comes from the population census of 2004, which is the last census conducted in the country to date. The data is aggregated geographically at the level of cities, towns, and villages. It includes measures of female and male labor force participation, as well as data on demographics, economic structure, and education levels. Ethnicity data comes from previous work on the social structures in Syria (Khaddour & Mazur, 2018). A city, town, or village is assigned a majority based on an estimated 80%

cutoff (i.e., 80% of the village is of the same ethnic group for the village to have an ethnic majority), with most towns and villages having clear ethnic majorities while cities are typically ethnically mixed. The data necessitates that the ethnic gaps are measured between ethnic majority areas but does not capture ethnic differences between populations within cities.

The use of aggregated data is justified by the nature of the research question, which seeks to explore gaps in FLFP across geographically distinct ethnic majority areas. Additionally, individual-level census data is unavailable in this context. It is important to note that this approach has several limitations: Aggregation omits within-area variation, making it impossible to capture individual-level differences within the same area, or those between ethnic minorities and majorities in mixed urban areas. Another limitation relates to the data on ethnic composition, which was collected based on expert interviews, rather than measured directly. Meaning that ethnic composition shares are likely to be highly imprecise. For this reason, a conservative 80% cutoff is used for assigning the ethnic majority for any city, town, or village, with areas at or below the cutoff treated as "mixed" areas.

To evaluate the relationship between FLFP and economic development, two alternative measures of income are imputed based on the sectoral composition of the labor force or alternatively based on the educational attainment of the labor force. Real wage data is used for the calculations which comes from Syrian Labour Force Survey of 2007. The imputation formula takes the following form for sector-based income:

$$Income_i = \sum_{j=1}^{7} Sector\ wage_j \ .\ \ Sector\ share_{ij}$$

Where i refers to the city, town, or village, j to the sector. *Sector wage* refers to the national-level average wage of sector j, and *Sector share* is the share of workers in the city/town/village employed in that sector. The set of sectors used to construct the sectoral income measure are agriculture, industry, construction, hotels and restaurants, transportation, finance and real estate, and other services. While wage imputation is standardly used in the study of income inequality (such as income scores, see Sobek 1995), its limitations should also be understood. In this case, the measure captures the sectoral composition of the labor force, with low values of sectoral income corresponding to areas where the labor force is predominantly employed in agriculture, while increasing values of the measure track with the transition of the labor force into secondary and tertiary sector occupations and eventually into service-sector and white-collar dominant economies. The sectoral income measure is standardized according to its z-score value, with the national average weighted by population size taking a value of 0 with a standard deviation of 1. The alternative specification of the income measure uses educational attainment instead of income using the following formula:

$$Income_i = \sum_{j=1}^{7} Average\ wage_j \ .\ \ Educational\ attainment\ share_{ij}$$

Where the educational attainment share refers to the share of the adult population in attainment categories: literate, elementary school, middle school, high school, middle academy, and university educated, with the illiterate group treated as the reference group. Average wage refers to the national-level average wage conditional on the educational attainment category.

To understand the contribution of social norms to explaining ethnic gaps in FLFP, gender parity indices (GPIs) in education are used to proxy for bias against girls and women. This bias may take the form of lower investment in girls' and women's education as compared to their male counterparts, or as prescriptive norms that limit their freedoms and participation in activities outside the home by limiting their education. Three GPIs are calculated based on the secondary school enrollment ratio between girls and boys, the schooling attainment ratio between adult women and adult men at the secondary level and above, and at the tertiary level. The GPIs measure at 100 in places with equality between the genders, and measure below 100 if male outcomes are higher than female outcomes, and above 100 when female outcomes overtake male outcomes. The enrollment GPI is calculated based on the enrollment of girls and boys aged 15–17 during the stage of upper secondary schooling. In the Syrian context, schooling is mandatory until the age of 16, which allows for variation in the enrollment rate that is determined by parental decisions after the age of 16. The GPI for secondary education and above (termed secondary+) captures the gap in educational attainment at the secondary level and above. And another GPI is used for educational attainment at the tertiary level.

Educational attainment of both women and men is determined by parental investment during late teens and early 20s. Students in Syria typically rely on family support, and in particular women who choose to pursue education beyond the minimum required by law. Pursuing education at above the secondary level also requires postponing marriage, and parental decisions typically play an important role in the timing of marriage. The GPIs are used in the analysis to proxy for the kinds of social norms that inhibit women's participation in wider society, including prescriptive norms that curb mixing between women and men, or those that favor the seclusion of women or their early marriage. While secondary level students in Syria are often segregated on gender, there is no such segregation above the secondary level. Furthermore, secondary schooling establishments for boys and girls are typically located in close proximity and attending school typically offers opportunity for students of different genders to interact. In effect, low values of the GPIs correspond to places with high differential in educational investment in favor of males, which proxy for both favoritism towards males and limitations on women's freedom to pursue activities outside the home on par with their male counterparts. Since educational enrollment in Syria is free at all levels, costs of education are not expected to shape the decision to attend school at the secondary or tertiary levels. Though the opportunity cost of education may play a role, particularly in agricultural regions where the labor of young men and women is used on the farm. This possibility is accounted for by including secondary polynomial terms for all GPI measures in the regression analysis.

The Gelbach decomposition method (Gelbach, 2016) is used to evaluate the role of various factors that determine the ethnic gaps in FLFP. This method allows for estimating the contribution of each individual factor to explaining gaps between groups, without introducing bias due to the order with which the different variables are included. The Gelbach decomposition starts with an unadjusted ordinary least-squares model of the form:

$$FLFP_i = \alpha + \sum_{d=1}^{7} \mu_{d,\ unadjusted}\ D_i + \varepsilon_i \tag{1}$$

Where i refers to the city, town, or village. *FLFP* is the female labor force participation ratio, which measures the ratio of the female population in working age (18–65) who are

engaged in the labor force as proportion of the total female population in the same age group. *D* includes a set of seven ethnic majority dummy variables, which separate the cities, towns, and villages into different ethnic majority groups with Sunni Families as the reference category.

The adjusted model takes a similar form and adds a set of controls:

$$FLFP_i = \alpha + \sum_{d=1}^{7} \mu_{d,\ adjusted}\ D_i + \sum \gamma\ X_i + \varepsilon_i \tag{2}$$

The set of controls X_i include male labor force participation (MLFP), demographic controls, the imputed income variables, the shares of workers employed in the public sector and in the family sector, and gender parity indices. The demographic controls include the marriage fertility rate, and the proportion of the population aged 15–64. Second-order polynomial terms are used for the demographic controls, income measures, and the gender parity indices to account for the non-linearity due to the U-feminization pattern. An additional specification adds a broader set of controls that include education composition of the female population, a set of infrastructure variables that measure the proportion of homes connected to the electricity network, the sanitation network, and drinking water network, and a set of controls that capture information on informality in home ownership, as well as a broad set of geographic controls that include longitude and latitude and information on agricultural suitability. The observations are weighed by population size, which allows for generalizable analysis that is not skewed by the most numerous villages with the smallest populations.

The expanded set of variables is meant to examine mechanisms and additional competing hypotheses. Education, infrastructure, and housing informality capture various mechanisms for the effect of economic development on FLFP. The inclusion of geographic variables is motivated by the idea that social norms around FLFP may be shaped by environmental factors, particularly those related to agricultural suitability. Theories of gendered division of labor suggest that historical reliance on different forms of agriculture – such as plow-based versus shifting cultivation – has long-term effects on gender roles, including attitudes toward women's work outside the household (Alesina et al., 2013). By including controls for agricultural suitability, the analysis accounts for the possibility that historical economic conditions contributed to the development of group-specific labor market norms that persist today.

The FLFP measure is based on an expanded definition which includes both employment in the "modern" labor force, as well as work performed in both the informal sector and in the family sector. The family sector counts activities that are performed within the family unit – which mostly refers to agricultural work but also includes any production activities performed within the household, without necessarily receiving a wage in exchange for this activity. This definition of FLFP does not include any service activities that are not intended for marketization, such as household chores or care activities. This definition accords with the available data and is standardly used in census data collection in developing countries (Gaddis & Klasen, 2014).

Having run models (1) and (2), the explained ethnic gap in FLFP that is due to the added explanatory variables can be expressed as $\mu_{d,\ adjusted}$ - $\mu_{d,\ unadjusted}$." To estimate the contribution of each variable to this gap according to the Gelbach method, a set of auxiliary regressions were run for each added covariate and each ethnic group, with each variable used in equation (2) as the dependent variable. For each auxiliary regression, an indicator variable is included for the ethnic majority, alongside controls for longitude

and latitude and a constant term. For example, an auxiliary regression for MLFP for Alawis takes the following form (with Sunni families as the base category, and restricting the sample to only Sunni Families and Alawis):

$$MLFP_i = \alpha + \tau_{MLFP,\ Alawi} + \varepsilon_i$$

Then the contribution of *MLFP* to explaining the Alawi-Sunni gap in FLFP is calculated using the following formula:

$$\delta_{MLFP,\ Alawi} = \gamma_{MLFP}\tau_{MLFP,Alawi}$$

Where γ_{MLFP} is the coefficient for *MLFP* obtained from equation (2). The same process is repeated for every ethnic group and every variable introduced in equation (2).

In addition to the Gelbach decomposition analysis, the contribution of each variable to explaining ethnic FLFP gaps is tested using models that include each variable as the sole explanatory variable (besides the ethnicity dummy variables and latitude and longitude). The findings from this analysis show similar results to the Gelbach decomposition and are therefore included in the appendix (see Appendix Tables A.1–A.7).

It is important to note that the Gelbach decomposition method explains changes in coefficients due to correlation, but does not imply causal mechanisms. The method has the same limitations as those of OLS, like sensitivity to omitted variable bias. The results presented in the forthcoming section should be read with this limitation in mind. While the available data allows for exploring the roles of a rich set of variables, it is not possible to rule out the presence of other factors outside of the included variables. For example, social norms that are unrelated to the included GPIs and geographic variables, such as those related to the male breadwinner status.

5. Results

5.1. Descriptive statistics and group differences

Descriptive statistics are presented in in Table 1. Average FLFP varies widely by group, from 10% for Sunni families to 28% for Alawis and 26% for Christians. The former group is the largest numerically in terms of population and is thus chosen as the reference group in the analysis. In contrast to FLFP, male labor force participation is far more limited in range and only varies between 77% and 81% at the group level. Figure 1 shows a map of FLFP, highlighting a concentration of high FLFP in the coastal areas in the west of Syria where Alawi and Christian majorities predominate. Meanwhile, the inland regions of the country show varied levels of FLFP across geography, with no clear geographic trends or patterns apart from the coastal-inland division, and a mix of both high and low FLFP.

The income *z*-scores shown in Table 1 differ by group, highlighting large regional inequalities across the country. Alawis, Christians, Druze, and Mixed (urban) areas have the highest average incomes, while Kurdish and tribal areas have lower incomes in comparison at around one standard deviation below the mean. While the two income measures are not always equal, they are consistent in terms of sign for all ethnic groups and in terms of magnitude (with the exception of Christian areas where education-based income outpaces sector-based income). As the largest ethnic group, Sunni families are also the nearest to the national average, which is also convenient for the analysis with this group used as the base category. Large regional inequalities are also present in public sector employment, which is highest in Alawi areas at around 60%, but only reaches 27%

Table 1. Descriptive statistics

		Sunni Families	Alawis	Christians	Druzes	Kurds	Mixed	Other Minorities	Sunni tribes
	N	1315	1048	106	128	536	327	151	1569
	FLFP	10.27	28.34	25.75	17.22	12.47	15.36	15.61	17.06
	MLFP	80.73	78.28	80.33	76.59	81.99	80.67	79.17	77.08
Demographics	Aged 15–64	55.0	63.5	63.7	64.0	56.9	59.8	58.7	51.7
	Age dependency ratio	82.6	58.3	57.7	56.7	77.0	67.8	71.1	94.8
	Marriage fertility rate	6.4	5.3	4.8	4.5	5.9	5.1	5.8	8.0
	Female headed families	7.5	7.7	9.8	10.7	6.9	8.9	7.8	7.9
Income and employment	Income (education-based)	−0.20	0.60	1.08	0.52	−1.11	0.54	0.10	−1.09
	Income (sectors-based)	−0.12	0.48	0.26	0.18	−1.05	0.56	0.14	−1.17
	Public sector share	27.2	59.5	42.3	40.1	16.0	33.0	40.2	19.8
	Family sector share	10.3	13.9	9.6	8.6	19.7	3.3	6.8	23.9
Gender parity indices	GPI (enrollment 15–17)	99.0	99.9	99.1	100.0	94.8	99.2	99.1	86.1
	GPI (secondary education)	59.2	82.7	87.5	78.9	48.9	85.8	73.8	28.4
	GPI (tertiary education)	31.7	47.8	44.3	39.2	15.8	52.7	39.0	13.6
Education	Female years of education	4.8	6.2	6.7	6.3	2.9	5.9	5.3	2.5
	Male years of education	6.1	7.3	7.6	7.5	4.2	6.6	6.4	4.2
Infrastructure	Drinking water network	90.4	84.9	89.4	95.3	57.8	95.1	88.3	68.1
	Sanitation network	73.2	54.5	80.3	34.6	46.6	92.0	79.2	34.8
	Electricity network	97.2	97.2	97.1	97.8	95.4	97.6	97.0	94.1
	Crowding rate	1.8	1.4	1.2	1.3	2.1	1.5	1.7	2.0

(*Continued*)

Table 1. (*Continued*)

		Sunni Families	Alawis	Christians	Druzes	Kurds	Mixed	Other Minorities	Sunni tribes
Private property institutions	Official deeds	60.4	56.0	61.2	76.7	23.9	56.3	43.3	35.6
	Agricultural deeds	11.6	13.1	9.8	3.7	16.8	12.6	9.7	17.4
	Public notary	8.2	1.2	2.1	1.3	16.7	15.0	21.2	3.2
	Owned dwellings	87.6	90.6	82.8	92.6	87.6	83.1	85.4	85.3
Geography	Slope	2.3	5.9	4.4	2.7	3.5	1.7	2.0	1.1
	Elevation	560.8	364.2	433.1	1054.8	440.0	480.5	524.0	297.1
	Temperature	16.9	16.3	16.6	15.8	17.7	17.1	17.1	19.6
	Precipitation	148	317	201	120	207	178	143	114
	Flow accumulation	280	69	292	56	62	6	431	674
	Wheat suitability	9046	7864	8165	7742	8743	9692	9479	8310
	Cotton suitability	8517	7100	7456	7057	7919	8965	8819	8305

Note: The columns represent means weighted by population size in each ethnic majority group of cities, towns, and villages.

for Sunni families. Public sector employment is also elevated among Christinas, Druze, and other minorities, reaching around 40%, and is lowest for Kurds and Sunni tribes at 16% and 20%, respectively. In the latter groups, rates of work in the family sector are relatively high at 20% and 24%, respectively, due to the higher rates of agricultural work in the north-east of Syria. Work in the family sector is lowest in mixed areas at only 3%, and measures at 10% for Sunni families.

The infrastructure variables show some disparities between groups, particularly for the sanitation network, while the gaps in the development of the electricity network and the drinking water network are overall smaller and concentrated between the north-east region (Kurds and Sunni tribes) and the rest of the country. Similar patterns can be seen in the variables capturing the formality of the private property institutions. The data records the share of housing deeds according to the type of deed, where different deed types correspond to different levels of formality. Three types are recorded in the data – formal deeds, agricultural deeds, and public notary contracts. Formal deeds correspond to the share of formally owned homes. Agricultural deeds correspond to the share of homes built on formerly agricultural land without a government permit. And public notary contracts correspond to the share of homes where owners possess a notarized contract verifying their purchase of the property from the previous owner, but without holding an official deed to the property.

In terms of geography, there is high agricultural suitability across most regions. The coastal Alawi region receives more rain than is typical for Syria but suffers from relatively lower suitability due to high ruggedness (slope). Druze areas also appear to have lower agricultural suitability as compared to other regions.

The Gender Parity Indices (GPIs) in Table 1 show small gaps in rates of school enrollment for boys and girls aged 15–17, and large gaps in GPIs for educational attainment in the adult population (aged 18 and above) at both the secondary levels and over (secondary+), and at the tertiary levels. Enrollment rates (aged 15–17) are at 99 for all groups with the exception of Kurds and Sunni tribes where they measure at 95 and 86, respectively. These lower rates are likely a function of lower levels of economic development and lower overall educational levels in the north-east of Syria. This can be seen in the low average years of education of adult males which measures at around 4 years in both groups. However, GPIs in the adult population tell a different story, with large gaps in secondary and tertiary educational attainment between men and women. Secondary education GPIs are lowest amongst Sunni tribes at 28, but they are also relatively low for Sunni families and Kurds at 49 and 59, respectively. The measure is higher amongst Alawis, Christians, and mixed populations where it reaches above 80 in all cases, and among Druzes and other minorities at 79 and 74, respectively. Similar patterns are found for tertiary education GPIs, with very large gaps amongst tribes and Kurds. Overall, the highest value is at 53 in mixed (urban) populations where the proportion of women with a tertiary education reaches almost half of that of men.

The gaps in economic development and educational attainment across the country are also associated with differences in demographics across groups. Fertility rates differ markedly, with the highest fertility amongst Sunni tribes, and the lowest amongst the Druzes, Christians, and mixed populations. The age structure and the age dependency ratio differ in turn, with very high dependency ratios and low proportions of the population in the working age amongst Sunni tribes and Sunni families, whereas Alawis, Christians, and Druzes enjoy low dependency ratios and relatively high proportions of the population in working age. The differences in age structure are stark even by global standards. As a comparison, shares of the population in at ages 15–64 in Africa and the

European Union measured at 56% and 65%, respectively in 2021 (UN, World Population Prospects, 2022).

5.2. Gelbach decomposition of covariates

Table 2 shows the Gelbach decomposition of covariates. For each ethnic group, the coefficient on the ethnicity dummy variable is shown for both the unadjusted and adjusted models. Below that, the Gelbach decomposition includes each variable name, the contribution of that variable calculated as described in the methods section (along with the standard error), and the contribution as a percentage of the explained gap (i.e., the difference between the ethnicity dummy variables in the unadjusted and adjusted models). The percentage of contributions are summed for each group of variables, for ease of interpretation, and to address the multicollinearity of closely related variables.

As the ethnicity coefficients show, the majority of FLFP gaps are accounted for by the included variables. The Alawi FLFP premium is reduced from 18.5 to 4.5 percentage points, while the Christian premium is reduced from 16.8 to 3.2 percentage points. The residual gaps in the adjusted model remain statistically significant in some cases, but they are at much reduced levels and do not measure more than 5 percentage points in any ethnic groups.

Turning to the decomposition of covariates, the most important set of variables appear to be the income and employment variables, accounting for 62% of the Alawi FLFP premium and 58% of the Christian premium, and significant shares of premiums in other groups. The public sector share of workers is also an important contributor, especially in the case of the Alawi premium (49%). In the cases of Kurds and Sunni tribes, the negative coefficients indicates that these groups enjoy higher FLFP despite the very low public sector shares in their areas. Instead, the family sector share in these two groups appears to explain a share of their FLFP premia, due to a higher share of workers in agriculture, in line with the U-feminization hypothesis.

The set of demographic variables contribute to explaining around a quarter of the Alawi FLFP premium, and a third of the Christian FLFP premium. Most of this is due to the contribution of the age structure of the population and the age dependency ratio, though the two variables are highly correlated which makes it difficult to separate the two contributions. In contrast, the marriage fertility rate contributes a very small share of the gap (summing the first and second order terms yields 4% for Alawis and 5% for Christians).

The contributions of the gender parity indices are small overall, measuring at 6% for Alawis and 8% for Christians, and are even less significant for Druzes, but they do appear to be larger for mixed (urban) areas at 25% and for Sunni tribes at –37%. The GPI contribution is due almost entirely to the gender gap in secondary educational attainment for the adult population, whereas school-aged and tertiary education GPIs are not contributors. The contributions are positive for the secondary attainment GPI, meaning that a smaller educational gap between women and men is associated with higher FLFP. With the only exception being in Kurds and Sunni tribes where the contributions are negative, which reflect the lower gender parity indices in these group. Here again, the negative relationship between the GPIs and FLFP in these groups is in line with the U-feminization hypothesis. The same results are found when including each GPI separately in both linear and polynomial forms without additional controls

Table 2. Gelbach decomposition of covariates

	Alawis		Christians		Druzes		Kurds		Mixed		Other minorities		Sunni tribes	
Unadjusted model	18.129***		15.497***		6.937***		2.180**		5.081***		5.301***		6.773***	
	(0.571)		(1.490)		(1.158)		(0.907)		(0.331)		(0.812)		(0.417)	
Adjusted model	4.921***		2.950***		−0.174		−1.633**		1.116***		1.042*		3.957***	
	(0.553)		(1.120)		(0.900)		(0.746)		(0.329)		(0.603)		(0.439)	
Gelbach decomposition of covariates:														
Spatial trends		**[4%]**		**[2%]**		**[−14%]**		**[20%]**		**[8%]**		**[2%]**		**[−1%]**
Latitude	0.438***	[3%]	0.443***	[4%]	−1.006***	[−14%]	1.569***	[41%]	0.489***	[12%]	0.190***	[4%]	0.900***	[32%]
	(0.086)		(0.136)		(0.191)		(0.273)		(0.086)		(0.070)		(0.155)	
Longitude	0.109***	[1%]	−0.142**	[−1%]	−0.011	[−0%]	−0.792***	[−21%]	−0.172***	[−4%]	−0.113***	[−3%]	−0.918***	[−33%]
	(0.042)		(0.072)		(0.043)		(0.261)		(0.057)		(0.048)		(0.301)	
Demographics		**[23%]**		**[31%]**		**[47%]**		**[7%]**		**[53%]**		**[25%]**		**[−72%]**
Aged 15–64	103.560***	[784%]	104.991***	[837%]	108.974***	[1532%]	22.502***	[590%]	57.600***	[1453%]	44.572***	[1047%]	−41.378***	[−1469%]
	(62.206)		(63.410)		(65.631)		(14.190)		(34.602)		(27.031)		(24.911)	
Aged 15–64^2	−45.999***	[−348%]	−46.345***	[−369%]	−48.023***	[−675%]	−9.710***	[−255%]	−24.411***	[−616%]	−18.863***	[−443%]	16.500***	[586%]
	(33.989)		(34.361)		(35.543)		(7.409)		(18.040)		(14.029)		(12.215)	
Age dependency ratio	−72.168***	[−546%]	−73.875***	[−589%]	−76.887***	[−1081%]	−16.589***	[−435%]	−43.865***	[−1106%]	−34.080***	[−800%]	35.843***	[1273%]
	(37.897)		(39.131)		(40.543)		(9.338)		(23.031)		(18.125)		(18.857)	
Age dependency ratio^2	18.749***	[142%]	19.349***	[154%]	20.195***	[284%]	4.588***	[120%]	12.562***	[317%]	9.824***	[231%]	−11.906***	[−423%]
	(9.665)		(10.101)		(10.473)		(2.586)		(6.471)		(5.140)		(6.142)	

(*Continued*)

Table 2. (*Continued*)

	Alawis		Christians		Druzes		Kurds		Mixed		Other minorities		Sunni tribes	
Female headed families	0.094	[1%]	1.327***	[11%]	1.818***	[26%]	−0.325*	[−9%]	0.803***	[20%]	0.180	[4%]	0.229***	[8%]
	(0.117)		(0.349)		(0.334)		(0.189)		(0.124)		(0.167)		(0.090)	
Female headed families^2	−0.066	[−0%]	−0.682***	[−5%]	−0.872***	[−12%]	0.038	[1%]	−0.239***	[−6%]	−0.049	[−1%]	−0.324***	[−12%]
	(0.069)		(0.229)		(0.231)		(0.108)		(0.064)		(0.097)		(0.085)	
Fertility rate	−1.424***	[−11%]	−2.063***	[−16%]	−2.419***	[−34%]	−0.647***	[−17%]	−1.629***	[−41%]	−0.752***	[−18%]	2.027***	[72%]
	(0.255)		(0.480)		(0.461)		(0.244)		(0.257)		(0.230)		(0.320)	
Fertility rate^2	0.895***	[7%]	1.342***	[11%]	1.610***	[23%]	0.095	[2%]	1.298***	[33%]	0.660***	[15%]	−2.073***	[−74%]
	(0.195)		(0.400)		(0.370)		(0.205)		(0.225)		(0.212)		(0.351)	
Income and employment		**[62%]**		**[58%]**		**[42%]**		**[91%]**		**[14%]**		**[50%]**		**[176%]**
Income (education)	3.101***	[23%]	4.937***	[39%]	2.777***	[39%]	−3.550***	[−93%]	2.829***	[71%]	1.141***	[27%]	−3.477***	[−123%]
	(0.286)		(0.586)		(0.414)		(0.382)		(0.228)		(0.267)		(0.282)	
Income (education)^2	0.309***	[2%]	1.137***	[9%]	0.147*	[2%]	0.605***	[16%]	0.283***	[7%]	0.266***	[6%]	0.805***	[29%]
	(0.074)		(0.249)		(0.094)		(0.137)		(0.061)		(0.081)		(0.160)	
Income (sectors)	−2.826***	[−21%]	−1.759***	[−14%]	−1.433***	[−20%]	4.320***	[113%]	−3.174***	[−80%]	−1.228***	[−29%]	4.883***	[173%]
	(0.293)		(0.569)		(0.443)		(0.457)		(0.257)		(0.315)		(0.380)	
Income (sectors)^2	1.636***	[12%]	0.013	[0%]	0.031	[0%]	3.779***	[99%]	−0.300***	[−8%]	−0.128	[−3%]	4.766***	[169%]
	(0.210)		(0.507)		(0.394)		(0.359)		(0.114)		(0.276)		(0.271)	
Public sector	6.497***	[49%]	3.037***	[24%]	2.612***	[37%]	−2.269***	[−60%]	1.147***	[29%]	2.606***	[61%]	−1.492***	[−53%]
	(0.447)		(0.533)		(0.420)		(0.335)		(0.132)		(0.316)		(0.167)	

(*Continued*)

Table 2. (*Continued*)

	Alawis		Christians		Druzes		Kurds		Mixed		Other minorities		Sunni tribes	
Family sector	0.111***	[1%]	−0.022	[−0%]	−0.048	[−1%]	0.288***	[8%]	−0.214***	[−5%]	−0.110***	[−3%]	0.418***	[15%]
	(0.042)		(0.069)		(0.055)		(0.094)		(0.065)		(0.049)		(0.124)	
MLFP	−0.632***	[−5%]	−0.104	[−1%]	−1.066***	[−15%]	0.308*	[8%]	−0.017	[−0%]	−0.415**	[−10%]	−0.959***	[−34%]
	(0.123)		(0.304)		(0.245)		(0.186)		(0.068)		(0.168)		(0.103)	
Gender parity indices		**[6%]**		**[8%]**		**[10%]**		**[−10%]**		**[25%]**		**[13%]**		**[−37%]**
GPI (enrollment)	−0.010*	[−0%]	−0.001	[−0%]	−0.010	[−0%]	0.042***	[1%]	−0.002	[−0%]	−0.002	[−0%]	0.132***	[5%]
	(0.022)		(0.014)		(0.024)		(0.091)		(0.005)		(0.008)		(0.284)	
GPI (enrollment)^2	−0.003	[−0%]	0.000	[−0%]	−0.003	[−0%]	0.009***	[0%]	0.000	[−0%]	−0.002	[−0%]	0.032***	[1%]
	(0.015)		(0.006)		(0.015)		(0.046)		(0.002)		(0.009)		(0.163)	
GPI (secondary)	1.170***	[9%]	1.400***	[11%]	0.984***	[14%]	−0.538***	[−14%]	1.331***	[34%]	0.730***	[17%]	−1.544***	[−55%]
	(0.281)		(0.379)		(0.274)		(0.173)		(0.311)		(0.200)		(0.361)	
GPI (secondary)^2	−0.348***	[−3%]	−0.448***	[−4%]	−0.266***	[−4%]	0.099**	[3%]	−0.368***	[−9%]	−0.174***	[−4%]	0.347***	[12%]
	(0.177)		(0.236)		(0.144)		(0.066)		(0.186)		(0.095)		(0.175)	
GPI (tertiary)	0.001***	[0%]	0.001***	[0%]	0.001***	[0%]	−0.002***	[−0%]	0.002***	[0%]	0.001***	[0%]	−0.002***	[−0%]
	(0.148)		(0.117)		(0.071)		(0.168)		(0.198)		(0.068)		(0.176)	
GPI (tertiary)^2	0.017***	[0%]	0.014***	[0%]	0.004	[0%]	−0.008***	[−0%]	0.012***	[0%]	0.003	[0%]	−0.008***	[−0%]
	(0.089)		(0.073)		(0.020)		(0.040)		(0.063)		(0.018)		(0.041)	
Total	**13.208**	**[100%]**	**12.547**	**[100%]**	**7.111**	**[100%]**	**3.813**	**[100%]**	**3.965**	**[100%]**	**4.259**	**[100%]**	**2.816**	**[100%]**

Note: The table shows the coefficients on ethnic group dummies in the unadjusted and the adjusted models. The unadjusted model is a regression of FLFP on ethnic group majority dummies, and the adjusted model adds the control variables shown in the Gelbach decomposition. The Gelbach decomposition shows the contribution of each control variable to explaining the difference in ethnic coefficients between the unadjusted and adjusted models. The percentages in brackets measure the contribution of each control variable as a share of the total gap between the unadjusted and adjusted model ethnicity coefficients. Grouped contributions are shown in bold, where the contributions of each set of variables are grouped together for ease of interpretation. See appendix table A.8 (column 1) for the adjusted model full regression output.

(See Appendix Table A.2), which suggests that the small size of GPI contributions are not due to the inclusion of other control variables.

5.3. Gelbach decomposition with additional controls

Table 3 repeats the same decomposition analysis with the addition of variables capturing women's education, infrastructure development, private property institutions, and geography. The expanded model does not differ significantly from the model outlined above in terms of explanatory power, but it allows us to examine some of the mechanisms and additional competing hypotheses.

The contributions of the demographic variable appear robust and similar to that found in the previous model. On the other hand, the contribution of the income and employment variables is reduced, and the education variables capture an important share of those contributions. Income and employment explain 40% of the Alawi FLFP premium, and only 23% of that of Christians. The education variables capture 29% of the Alawi premium, and 52% of that of Christians. For Kurds and Sunni tribes, the contribution of education is negative, reflecting the lower average educational attainment in these areas compared to the reference group. The contribution of GPI is very small in the expanded model for both Alawis and Christians, with most of the effect captured by the education controls.

The contributions of infrastructure variables are small and negative overall, exceeding 10% only in the case of Kurds. The negative sign indicates that, on average, groups with better infrastructure tend to have lower FLFP, consistent with the broader U-shaped relationship between economic development and FLFP. However, the modest size of these contributions suggests that infrastructure is not a major mechanism in explaining FLFP differences across groups, particularly compared to the role of educational attainment. The limited effect may also reflect the relatively small variation in infrastructure development, such as the near-universal access to electricity, across the groups studied.

The next set of variables capture the role of private property institutions. The contributions are very small in the case of Alawis and Christians at 3% and 1%, respectively. The effects are slightly larger in the case of Kurds, Sunni tribes, and other minorities reaching up to 9%. In these groups, the effects align with high female participation in agricultural activities, which may be linked to lower land ownership rates in rural areas.

The last set of variables capture the role of geography. The contributions vary in size, measuring at only 1% for Christians, somewhat higher at 11% for Alawis, and reaching up to –67% and –29% for Kurds and Sunni tribes, respectively. In the case of Alawis, the largest contribution comes from the elevation variables at 8%, with higher elevation associated with higher FLFP. For Kurds and Sunni tribes, the contribution of geography appears to be driven by latitude and longitude. The North and East of Syria are the areas inhabited by these groups, which may explain the geographic correlation with FLFP. Wheat suitability and cotton suitability are not major contributors to the observed gaps, with cotton suitability contributing at most 8% in the case of Kurds.

6. Discussion and conclusions

The analysis presented here reveals the presence of large gaps in FLFP between ethnic and religious groups in Syria. Syrian women in Alawi and Christian communities join

Table 3. Gelbach decomposition of covariates with additional controls

	Alawis		Christians		Druzes		Kurds		Mixed		Other minorities		Sunni tribes	
Unadjusted model	18.129***		15.497***		6.937***		2.180**		5.081***		5.301***		6.773***	
	(0.571)		(1.490)		(1.158)		(0.907)		(0.331)		(0.812)		(0.417)	
Adjusted model	4.908***		2.053*		0.389		−0.527		1.830***		0.688		3.277***	
	(0.620)		(1.138)		(1.012)		(0.775)		(0.382)		(0.610)		(0.520)	
Gelbach decomposition of covariates:														
Demographic variables		**[28%]**		**[31%]**		**[68%]**		**[10%]**		**[70%]**		**[35%]**		**[−39%]**
Aged 15–64	19.057***	[144%]	19.320***	[144%]	20.053***	[306%]	4.141***	[153%]	10.599***	[326%]	8.202***	[178%]	−7.614***	[−218%]
	(64.317)		(65.217)		(67.685)		(13.998)		(35.773)		(27.690)		(25.700)	
Aged 15–64^2	−3.133***	[−24%]	−3.156***	[−23%]	−3.271***	[−50%]	−0.661***	[−24%]	−1.663***	[−51%]	−1.285***	[−28%]	1.124***	[32%]
	(35.214)		(35.480)		(36.764)		(7.435)		(18.688)		(14.441)		(12.631)	
Age dependency ratio	−16.970***	[−128%]	−17.371***	[−129%]	−18.080***	[−276%]	−3.901***	[−144%]	−10.315***	[−317%]	−8.014***	[−174%]	8.428***	[242%]
	(38.878)		(39.815)		(41.429)		(8.972)		(23.630)		(18.372)		(19.311)	
Age dependency ratio^2	5.022***	[38%]	5.183***	[39%]	5.409***	[83%]	1.229***	[45%]	3.365***	[103%]	2.632***	[57%]	−3.189***	[−92%]
	(9.813)		(10.136)		(10.574)		(2.418)		(6.574)		(5.147)		(6.232)	
Female headed families	0.071	[1%]	0.997***	[7%]	1.366***	[21%]	−0.245*	[−9%]	0.604***	[19%]	0.136	[3%]	0.172***	[5%]
	(0.088)		(0.288)		(0.299)		(0.145)		(0.118)		(0.127)		(0.071)	
Female headed families^2	−0.047	[−0%]	−0.487***	[−4%]	−0.622***	[−9%]	0.027	[1%]	−0.170***	[−5%]	−0.035	[−1%]	−0.231***	[−7%]
	(0.050)		(0.192)		(0.209)		(0.077)		(0.058)		(0.070)		(0.077)	
Fertility rate	−0.800***	[−6%]	−1.158***	[−9%]	−1.358***	[−21%]	−0.363***	[−13%]	−0.914***	[−28%]	−0.422***	[−9%]	1.138***	[33%]
	(0.227)		(0.371)		(0.396)		(0.159)		(0.248)		(0.159)		(0.309)	
Fertility rate^2	0.520***	[4%]	0.780***	[6%]	0.936***	[14%]	0.055	[2%]	0.755***	[23%]	0.383***	[8%]	−1.205***	[−35%]

(*Continued*)

Table 3. (*Continued*)

	Alawis		Christians		Druzes		Kurds		Mixed		Other minorities		Sunni tribes	
	(0.163)		(0.292)		(0.302)		(0.119)		(0.215)		(0.151)		(0.340)	
Income and employment		**[40%]**		**[23%]**		**[15%]**		**[177%]**		**[−30%]**		**[18%]**		**[166%]**
Income (education)	1.427***	[11%]	2.272***	[17%]	1.278***	[20%]	−1.634***	[−60%]	1.302***	[40%]	0.525***	[11%]	−1.600***	[−46%]
	(0.687)		(1.107)		(0.634)		(0.792)		(0.624)		(0.277)		(0.767)	
Income (education)^2	0.024***	[0%]	0.088***	[1%]	0.011*	[0%]	0.047***	[2%]	0.022***	[1%]	0.020***	[0%]	0.062***	[2%]
	(0.086)		(0.316)		(0.041)		(0.168)		(0.079)		(0.074)		(0.224)	
Income (sectors)	−2.313***	[−18%]	−1.440***	[−11%]	−1.173***	[−18%]	3.536***	[131%]	−2.598***	[−80%]	−1.005***	[−22%]	3.997***	[115%]
	(0.278)		(0.474)		(0.370)		(0.431)		(0.263)		(0.265)		(0.395)	
Income (sectors)^2	1.704***	[13%]	0.013	[0%]	0.033	[0%]	3.935***	[145%]	−0.312***	[−10%]	−0.133	[−3%]	4.962***	[142%]
	(0.217)		(0.528)		(0.410)		(0.370)		(0.118)		(0.288)		(0.274)	
Public sector	4.914***	[37%]	2.297***	[17%]	1.976***	[30%]	−1.716***	[−63%]	0.867***	[27%]	1.971***	[43%]	−1.128***	[−32%]
	(0.487)		(0.435)		(0.347)		(0.281)		(0.117)		(0.277)		(0.150)	
Family sector	0.125***	[1%]	−0.025	[−0%]	−0.054	[−1%]	0.325***	[12%]	−0.241***	[−7%]	−0.124***	[−3%]	0.471***	[14%]
	(0.044)		(0.077)		(0.062)		(0.096)		(0.064)		(0.053)		(0.123)	
MLFP	−0.634***	[−5%]	−0.104	[−1%]	−1.069***	[−16%]	0.309*	[11%]	−0.017	[−1%]	−0.416**	[−9%]	−0.962***	[−28%]
	(0.123)		(0.305)		(0.246)		(0.187)		(0.068)		(0.168)		(0.104)	
Gender parity indices		**[−2%]**		**[−2%]**		**[−4%]**		**[1%]**		**[−8%]**		**[−4%]**		**[9%]**
GPI (enrollment)	0.019*	[0%]	0.002	[0%]	0.019	[0%]	−0.083***	[−3%]	0.004	[0%]	0.003	[0%]	−0.259***	[−7%]
	(0.025)		(0.028)		(0.031)		(0.100)		(0.007)		(0.015)		(0.309)	
GPI (enrollment)^2	−0.015	[−0%]	−0.002	[−0%]	−0.014	[−0%]	0.047***	[2%]	−0.002	[−0%]	−0.009	[−0%]	0.167***	[5%]
	(0.019)		(0.028)		(0.026)		(0.050)		(0.007)		(0.018)		(0.165)	

(*Continued*)

Table 3. (*Continued*)

	Alawis		Christians		Druzes		Kurds		Mixed		Other minorities		Sunni tribes	
GPI (secondary)	−0.539***	[−4%]	−0.645***	[−5%]	−0.453***	[−7%]	0.248***	[9%]	−0.613***	[−19%]	−0.336***	[−7%]	0.711***	[20%]
	(0.323)		(0.395)		(0.279)		(0.158)		(0.366)		(0.206)		(0.425)	
GPI (secondary)^2	0.165***	[1%]	0.213***	[2%]	0.126***	[2%]	−0.047**	[−2%]	0.175***	[5%]	0.082***	[2%]	−0.165***	[−5%]
	(0.178)		(0.232)		(0.139)		(0.054)		(0.189)		(0.091)		(0.178)	
GPI (tertiary)	0.127***	[1%]	0.100***	[1%]	0.061***	[1%]	−0.144***	[−5%]	0.170***	[5%]	0.059***	[1%]	−0.151***	[−4%]
	(0.166)		(0.134)		(0.082)		(0.189)		(0.222)		(0.078)		(0.198)	
GPI (tertiary)^2	0.004***	[0%]	0.003***	[0%]	0.001	[0%]	−0.002***	[−0%]	0.003***	[0%]	0.001	[0%]	−0.002***	[−0%]
	(0.089)		(0.073)		(0.020)		(0.040)		(0.063)		(0.017)		(0.041)	
Education		**[29%]**		**[52%]**		**[49%]**		**[−11%]**		**[67%]**		**[41%]**		**[9%]**
Tertiary	−0.279***	[−2%]	−0.818***	[−6%]	−0.196**	[−3%]	0.430***	[16%]	−1.025***	[−32%]	−0.187***	[−4%]	0.398***	[11%]
	(0.171)		(0.500)		(0.144)		(0.265)		(0.612)		(0.126)		(0.239)	
Middle Academy	2.459***	[19%]	5.967***	[44%]	2.321***	[35%]	−1.174***	[−43%]	1.840***	[57%]	1.389***	[30%]	−0.760***	[−22%]
	(0.372)		(0.909)		(0.415)		(0.258)		(0.272)		(0.265)		(0.141)	
Upper secondary	2.232***	[17%]	1.898***	[14%]	1.756***	[27%]	−1.031***	[−38%]	1.402***	[43%]	0.724***	[16%]	−1.319***	[−38%]
	(0.656)		(0.594)		(0.538)		(0.328)		(0.411)		(0.242)		(0.389)	
Lower secondary	−1.273***	[−10%]	−0.863***	[−6%]	−1.054***	[−16%]	0.998***	[37%]	−0.380***	[−12%]	−0.286***	[−6%]	1.422***	[41%]
	(0.344)		(0.271)		(0.302)		(0.280)		(0.106)		(0.109)		(0.382)	
Elementary	0.316***	[2%]	0.474***	[4%]	−0.137***	[−2%]	0.293***	[11%]	0.316***	[10%]	0.163***	[4%]	0.488***	[14%]
	(0.143)		(0.219)		(0.075)		(0.135)		(0.141)		(0.079)		(0.218)	
Literate	0.392***	[3%]	0.324***	[2%]	0.511***	[8%]	0.185***	[7%]	0.039***	[1%]	0.095***	[2%]	0.092***	[3%]
	(0.177)		(0.159)		(0.234)		(0.092)		(0.023)		(0.055)		(0.045)	

(*Continued*)

Table 3. (*Continued*)

	Alawis		Christians		Druzes		Kurds		Mixed		Other minorities		Sunni tribes	
Infrastructure		**[−3%]**		**[−4%]**		**[−6%]**		**[−12%]**		**[−5%]**		**[−3%]**		**[−7%]**
Drinking water network	−0.098***	[−1%]	−0.019	[−0%]	0.086*	[1%]	−0.577***	[−21%]	0.083***	[3%]	−0.038	[−1%]	−0.394***	[−11%]
	(0.040)		(0.061)		(0.055)		(0.195)		(0.031)		(0.035)		(0.131)	
Sanitation network	−0.023***	[−0%]	0.009	[0%]	−0.048***	[−1%]	−0.033***	[−1%]	0.023***	[1%]	0.007**	[0%]	−0.048***	[−1%]
	(0.091)		(0.034)		(0.188)		(0.128)		(0.091)		(0.029)		(0.186)	
Electricity network	−0.001	[−0%]	−0.002	[−0%]	0.010	[0%]	−0.031**	[−1%]	0.007	[0%]	−0.004	[−0%]	−0.055***	[−2%]
	(0.008)		(0.021)		(0.019)		(0.029)		(0.008)		(0.012)		(0.047)	
Crowding rate	−0.323***	[−2%]	−0.500***	[−4%]	−0.446***	[−7%]	0.313***	[12%]	−0.287***	[−9%]	−0.082***	[−2%]	0.250***	[7%]
	(0.145)		(0.228)		(0.202)		(0.142)		(0.128)		(0.046)		(0.112)	
Property institutions		**[−3%]**		**[−1%]**		**[−4%]**		**[1%]**		**[10%]**		**[9%]**		**[−9%]**
Official deeds	−0.032***	[−0%]	0.006	[0%]	0.116***	[2%]	−0.261***	[−10%]	−0.029***	[−1%]	−0.122***	[−3%]	−0.177***	[−5%]
	(0.026)		(0.029)		(0.088)		(0.193)		(0.023)		(0.091)		(0.131)	
Agricultural deeds	−0.005	[−0%]	0.006	[0%]	0.025***	[0%]	−0.016***	[−1%]	−0.003	[−0%]	0.006	[0%]	−0.019***	[−1%]
	(0.010)		(0.014)		(0.050)		(0.032)		(0.007)		(0.013)		(0.037)	
Public notary	−0.255***	[−2%]	−0.221***	[−2%]	−0.253***	[−4%]	0.316***	[12%]	0.249***	[8%]	0.477***	[10%]	−0.182***	[−5%]
	(0.074)		(0.091)		(0.087)		(0.095)		(0.069)		(0.135)		(0.053)	
Owned dwellings	−0.078***	[−1%]	0.124**	[1%]	−0.130***	[−2%]	0.001	[0%]	0.115***	[4%]	0.057*	[1%]	0.058***	[2%]
	(0.033)		(0.068)		(0.060)		(0.034)		(0.038)		(0.036)		(0.024)	
Geography		**[11%]**		**[1%]**		**[−18%]**		**[−67%]**		**[−4%]**		**[4%]**		**[−29%]**
Latitude	−0.525***	[−4%]	−0.531***	[−4%]	1.205***	[18%]	−1.880***	[−69%]	−0.586***	[−18%]	−0.227***	[−5%]	−1.079***	[−31%]
	(0.250)		(0.282)		(0.570)		(0.878)		(0.274)		(0.129)		(0.503)	

(*Continued*)

Table 3. (*Continued*)

	Alawis		Christians		Druzes		Kurds		Mixed		Other minorities		Sunni tribes	
Longitude	−0.021***	[−0%]	0.027**	[0%]	0.002	[0%]	0.149***	[6%]	0.032***	[1%]	0.021***	[0%]	0.173***	[5%]
	(0.137)		(0.178)		(0.016)		(0.991)		(0.215)		(0.141)		(1.149)	
Latitude^2	−0.105***	[−1%]	−0.155**	[−1%]	0.023	[0%]	−0.563***	[−21%]	−0.105***	[−3%]	0.013	[0%]	−0.448***	[−13%]
	(0.089)		(0.142)		(0.053)		(0.461)		(0.087)		(0.036)		(0.366)	
Longitude^2	0.211***	[2%]	0.172***	[1%]	−0.407***	[−6%]	0.112***	[4%]	0.086***	[3%]	0.076***	[2%]	0.199***	[6%]
	(0.272)		(0.224)		(0.525)		(0.145)		(0.111)		(0.099)		(0.257)	
Slope	0.068***	[1%]	0.039***	[0%]	0.007	[0%]	0.023***	[1%]	−0.013***	[−0%]	−0.007*	[−0%]	−0.025***	[−1%]
	(0.273)		(0.154)		(0.030)		(0.092)		(0.051)		(0.029)		(0.098)	
Slope^2	0.034***	[0%]	0.027***	[0%]	0.000	[0%]	0.038***	[1%]	−0.006***	[−0%]	0.000	[0%]	−0.008***	[−0%]
	(0.107)		(0.088)		(0.006)		(0.120)		(0.021)		(0.005)		(0.027)	
Elevation	1.016***	[8%]	0.659***	[5%]	−2.530***	[−39%]	0.625***	[23%]	0.415***	[13%]	0.190**	[4%]	1.361***	[39%]
	(0.471)		(0.349)		(1.170)		(0.306)		(0.195)		(0.129)		(0.627)	
Elevation^2	−0.028***	[−0%]	−0.025***	[−0%]	0.127***	[2%]	−0.026***	[−1%]	−0.015***	[−0%]	−0.009**	[−0%]	−0.043***	[−1%]
	(0.264)		(0.233)		(1.187)		(0.246)		(0.139)		(0.081)		(0.405)	
Temperature	−0.465***	[−4%]	−0.183	[−1%]	−0.808***	[−12%]	0.567***	[21%]	0.145***	[4%]	0.113	[2%]	1.923***	[55%]
	(0.247)		(0.262)		(0.440)		(0.316)		(0.090)		(0.145)		(0.945)	
Temperature^2	0.209**	[2%]	−0.056	[−0%]	1.235***	[19%]	−0.907***	[−34%]	−0.393***	[−12%]	−0.293**	[−6%]	−3.256***	[−93%]
	(0.142)		(0.257)		(0.639)		(0.472)		(0.202)		(0.200)		(1.602)	
Precipitation	−2.662***	[−20%]	−0.825***	[−6%]	0.447***	[7%]	−0.920***	[−34%]	−0.465***	[−14%]	0.086	[2%]	0.549***	[16%]
	(1.397)		(0.485)		(0.290)		(0.500)		(0.249)		(0.129)		(0.294)	
Precipitation^2	3.543***	[27%]	0.864***	[6%]	−0.398*	[−6%]	0.704***	[26%]	0.685***	[21%]	0.089	[2%]	−0.438***	[−13%]

(*Continued*)

Table 3. (*Continued*)

	Alawis		Christians		Druzes		Kurds		Mixed		Other minorities		Sunni tribes	
	(1.167)		(0.410)		(0.265)		(0.293)		(0.234)		(0.164)		(0.166)	
Flow accumulation	−0.004	[−0%]	0.000	[0%]	−0.004	[−0%]	−0.004	[−0%]	−0.005	[−0%]	0.003	[0%]	0.008	[0%]
	(0.010)		(0.020)		(0.017)		(0.014)		(0.010)		(0.012)		(0.014)	
Flow accumulation^2	0.000	[0%]	0.000	[−0%]	0.000	[0%]	0.000	[0%]	0.000	[0%]	−0.001	[−0%]	−0.008	[−0%]
	(0.010)		(0.026)		(0.021)		(0.016)		(0.006)		(0.014)		(0.012)	
Wheat suitability	−0.096***	[−1%]	−0.071***	[−1%]	−0.107***	[−2%]	−0.024**	[−1%]	0.052***	[2%]	0.035***	[1%]	−0.060***	[−2%]
	(0.762)		(0.566)		(0.845)		(0.194)		(0.415)		(0.280)		(0.474)	
Wheat suitability^2	−0.091***	[−1%]	−0.061***	[−0%]	−0.087***	[−1%]	−0.027***	[−1%]	0.053***	[2%]	0.037***	[1%]	−0.060***	[−2%]
	(0.754)		(0.508)		(0.720)		(0.222)		(0.441)		(0.308)		(0.497)	
Cotton suitability	−0.905***	[−7%]	−0.676***	[−5%]	−0.937***	[−14%]	−0.382***	[−14%]	0.287***	[9%]	0.195*	[4%]	−0.132***	[−4%]
	(0.555)		(0.450)		(0.587)		(0.257)		(0.179)		(0.155)		(0.095)	
Cotton suitability^2	1.270***	[10%]	0.976***	[7%]	1.042***	[16%]	0.706***	[26%]	−0.285***	[−9%]	−0.146	[−3%]	0.348***	[10%]
	(0.549)		(0.477)		(0.480)		(0.333)		(0.132)		(0.141)		(0.162)	
Total	**13.209**	**[100%]**	**13.444**	**[100%]**	**6.548**	**[100%]**	**2.707**	**[100%]**	**3.251**	**[100%]**	**4.613**	**[100%]**	**3.485**	**[100%]**

Note: The table shows the coefficients on ethnic group dummies in the unadjusted and the adjusted models. The unadjusted model is a regression of FLFP on ethnic group majority dummies, and the adjusted model adds the control variables shown in the Gelbach decomposition. The Gelbach decomposition shows the contribution of each control variable to explaining the difference in ethnic coefficients between the unadjusted and adjusted models. The percentages in brackets measure the contribution of each control variable as a share of the total gap between the unadjusted and adjusted model ethnicity coefficients. Grouped contributions are shown in bold, where the contributions of each set of variables are grouped together for ease of interpretation. See appendix table A.8 (column 1) for the adjusted model full regression output.

the labor force at more than twice the rate of their counterparts in the neighboring Sunni Muslim communities. Gaps in FLFP between Muslim sects have only been explored in few countries in the MENA like Turkey (Akyol & Ökten, 2024), and the analysis here is the first to measure these gaps in the Syrian context. The presence of Christian-majority areas in Syria also allows for the measurement of Muslim-Christian gaps in FLFP, and not just gaps between Muslim sects. Whereas most previous research on Christian-Muslim gaps in FLFP was undertaken at the country level. Surprisingly, gaps between Muslim sects appear to be at least as large as those between Christians and Muslims, with Alawi Muslim women joining the labor force at slightly higher rates than Christians. And while the Christian premium in FLFP in Syria is significant, Christian FLFP in this context is still consistent with the overall rates found in the MENA region, below the global average and below those of Christian-majority regions of the world like Europe, the Americas, and sub-Saharan Africa.

The decomposition analysis shows that economic and demographic structure matter for the ethnic and religious gaps in FLFP. The analysis does not find a large contribution for social norms as proxied by gender parity in school enrollment and educational attainment, even though women's education levels are important determinants of FLFP gaps between groups. The analysis contributes to the literature on women's work in the MENA region by elucidating the factors behind ethnic and religious gaps in FLFP. Recent research has emphasized social and cultural norms as main determinants for divergences in FLFP within and across societies, but the analysis here shows that broader structural factors appear to dominate. The small effects sizes and contributions of GPIs suggest that parental bias against girls is not a major factor behind gaps in FLFP between Muslim sects and communities, or even between Muslims and Christians.

Differences in the age structure across communities appear to play an important and underappreciated role for FLFP gaps, even after controlling for other important factors such as income and education levels. This effect may be due to the reduced dependency ratio resulting from a higher proportion of the population in the working age, leading to lower caregiving burdens which typically fall on women. The effect may also be due to differences in life stages, with a lower proportion of women in childbearing ages amongst those in the working age overall, in groups with relatively high FLFP like Alawis and Christians. Overall, the effect suggests a possible role for a demographic dividend for FLFP, with a more rapid fertility decline in Alawi and Christians communities leading to increased FLFP rates, even though current fertility rates do not play a significant role. Previous literature found that differences in age structure explain substantial parts of the ethnic FLFP gaps in the US (Reimers, 1985). Likewise, large differences in age structure exist between ethnic groups in Syria, with resultant differences in FLFP. The evidence is also in line with recent research suggesting the presence of a demographic dividend for FLFP in Sub-Saharan Africa (Backhaus & Loichinger, 2022).

One explanation for the link between demographic structure and female labor force participation (FLFP) is the influence of social norms. Gender norms shape demographic patterns by affecting family size and the timing of marriage; societies that value early marriage and large families tend to have more dependents, limiting women's work opportunities. In turn, demographic structure can reinforce expectations that women provide unpaid care when many children or elderly are present. Yet if fertility and caregiving norms were the main drivers, we would expect fertility rates to strongly predict FLFP, which is not the case. This suggests other factors are also at play, such as differences in age composition itself, which may help explain why fertility rates do not contribute significantly to explaining FLFP gaps.

Income appears to matter for FLFP gaps, but the size of its contribution depends on the ethnic group, as does the contribution of public sector employment share. The finding here supports earlier research on the Alawi community in Syria, and the role of public sector employment in the entry of Alawi women into the labor force (Balanche, 2015). Public sector employment intertwines with economic development, with the poorest areas in the north-east of Syria having relatively low public sector shares, despite FLFP that is higher than average. This is in part due to women's participation in the agricultural economy, especially amongst Kurds and Sunni tribes, where the contribution of the public sector share is low and negative suggesting that women's participation in these areas occurs despite the lack of public sector employment opportunities.

The gender gaps in educational attainment and school enrollment play a small role in explaining FLFP gaps between groups, despite the large overall contribution of educational attainment for FLFP gaps. The effect of the educational composition mirrors that of the U-feminization pattern, with the lowest FLFP levels associated with lower secondary education (grade 9 which is typically completed at age 14). Meanwhile, the shares of the lowest education categories are associated with higher FLFP, corresponding to areas with women's participation in the agricultural economy. The U-shaped pattern also holds between FLFP and years of education (See Appendix Table A.3). Educational attainment can proxy for economic development, capturing the human capital component of the development process (Acemoglu et al., 2014). Declining FLFP with increased levels of education may be driven in part by structural change, but the U-shaped relationship between FLFP and education is also present after controlling for income and demographics.

Overall, the structural correlates of FLFP in Syria mirror previous work on FLFP. Goldin (1990) emphasizes the role of increased demand of clerical workers for employment gains of women in the US during the first half of the 20th century. Similar patterns appear to be behind the FLFP premiums of Alawi and Christian women in Syria, with relatively high shares of educated workers employed in the mostly white-collar public sector in those groups. This finding also accords with previous macro level evidence on the role of public sector employment in expanding FLFP, both globally and in the MENA region (Klasen, 2019; Assaad et al., 2020).

By focusing on areas of majority for each group, the analysis inadvertently emphasizes their differences, meaning that the results of the analysis are conservative in the sense that FLFP differences are likely to be less divergent within cities. For example, Sunni Muslims and Christians may have more similar outcomes within Damascus due to acculturation or due to having more similar characteristics. Meanwhile, Sunni Muslims and Christians living in separate towns or villages have little or no interaction between them, and more divergent underlying factors. Future research should explore these settings further to understand whether different dynamics predominate in urban areas.

The civil war that began in 2011 redrew the population map of Syria, but the findings in this paper remain relevant for Syria and the MENA region more broadly. According to World Bank estimates, FLFP in Syria doubled from 13% in 2010 to 26% in 2021 (World Bank, 2022). This increase may have been driven by the fall in incomes, with a shortage of male breadwinners due to conflict and associated deaths and migrations. The war likely changed the calculus regarding the costs and benefits of wage-earning activity for women, and therefore helped push many towards increased participation. Though we should expect FLFP to continue to be shaped by the structural factors explored above,

we may see further declines in FLFP in areas where the agricultural economy is in transition, particularly in the north-east of Syria. There is little reason to expect convergence in FLFP rates between groups, as long as regional inequalities remain unaddressed. The de-facto fragmentation of the country following the civil war may result in further divergence between groups, adding political and institutional divergence to the litany of factors involved in shaping FLFP, but it may also lead inadvertently to more equitable distribution of resources which shape FLFP (education, investments, government jobs), especially in historically marginalized regions of the country.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/dem.2025.10007

Acknowledgements. I am grateful to Martin Dribe and Anna Tegunimataka for their guidance and support, and to participants at the Nordic Conference in Development Economics at Gothenburg University, the 12th Annual Workshop on "Growth, History and Development" at the University of Southern Denmark, and seminar participants at the Centre for Advanced Middle East Studies and the Department of Economic History at Lund University.

Funding statement. This research was supported by the Centre for Advanced Middle Eastern Studies (CMES) at Lund University.

Competing interests. None.

References

Abdelali-Martini, M., & Dey de Pryck, J. (2015). Does the feminisation of agricultural labour empower women? Insights from female labour contractors and workers in Northwest Syria. *Journal of International Development*, 27(7), 898–916.

Abdelali-Martini, M., Goldey, P., Jones, G., & Bailey, E. (2003). Towards a feminization of agricultural labour in northwest Syria. *The Journal of Peasant Studies*, 30(2), 71–94.

Abi-Rached, J. M., & Diwan, I. (2022). *The economic legacy of the French mandate in Lebanon*. Cairo, Egypt: Economic Research Forum (ERF).

Acemoglu, D., Gallego, F. A., & Robinson, J. A. (2014). Institutions, human capital, and development. *Annual Review of Economics*, 6(1), 875–912.

Akyol, P., & Ökten, Ç. (2024). The role of religion in female labour supply: evidence from two Muslim denominations. *Journal of Demographic Economics*, 90(1), 116–153.

Alesina, A., Giuliano, P., & Nunn, N. (2013). On the origins of gender roles: women and the plough. *The Quarterly Journal of Economics*, 128(2), 469–530.

Algan, Y., & Cahuc, P. (2006). Job protection: The macho hypothesis. *Oxford Review of Economic Policy*, 22(3), 390–410.

Antoun, R. T. (1991). Ethnicity, clientship, and class: Their changing meaning. In *Syria: Society, culture and polity* (pp. 1–12). Albany, NY: State University of New York Press.

Assaad, R. (2014). Making sense of Arab labour markets: the enduring legacy of dualism. *IZA Journal of Labour & Development*, 3(1), 1–25.

Assaad, R., Hendy, R., Lassassi, M., & Yassin, S. (2020). Explaining the MENA paradox: rising educational attainment, yet stagnant female labour force participation. *Demographic Research*, 43, 817.

Atasoy, B. S. (2017). Female labour force participation in Turkey: the role of traditionalism. *The European Journal of Development Research*, 29(4), 675–706.

Backhaus, A., & Loichinger, E. (2022). Female labour force participation in Sub-Saharan Africa: A cohort analysis. *Population and Development Review*, 48(2), 379–411.

Balanche, F. (2015). Go to Damascus, my son. In *The Alawis of Syria: War, faith and politics in the Levant* (pp. 79–106). Oxford, UK: Oxford University Press.

Bayanpour-Tehrani, G., & Sylwester, K. (2013). Female labour force participation and religion: a cross-country analysis. *Bulletin of Economic Research*, 65(2), 107–133.
Boserup, E. (1970). *Woman's role in economic development*. London, UK: George Allen & Unwin.
Chamie, J. (1977). Religious differentials in fertility: Lebanon, 1971. *Population Studies*, 31(2), 365–382.
Chatterjee, E., & Vanneman, R. D. (2022). Women's low employment rates in India: cultural and structural explanations. *Population and Development Review*, 48(2), 445–474.
Courbage, Y., & Todd, E. (2014). *A convergence of civilizations: The transformation of Muslim societies around the world*. New York, NY: Columbia University Press.
Dildar, Y. (2015). Patriarchal norms, religion, and female labour supply: evidence from Turkey. *World Development*, 76, 40–61.
Gaddis, I., & Klasen, S. (2014). Economic development, structural change, and women's labour force participation. *Journal of Population Economics*, 27(3), 639–681.
Gelbach, J. B. (2016). When do covariates matter? And which ones, and how much? *Journal of Labour Economics*, 34(2), 509–543.
Giuliano, P. (2017). *Gender: An historical perspective*. Bonn: Institute of Labor Economics (IZA).
Goldin, C. (1990). *Understanding the gender gap: An economic history of American women*.
Goldin, C. (1995). The U-shaped female labour force function in economic development and economic history. In *Investment in women's human capital* (pp. 61–90). Chicago, IL: University of Chicago Press.
Hassine, N. B. (2015). Economic inequality in the Arab region. *World Development*, 66, 532–556.
Hinnebusch, R. (2004). *Syria: Revolution from above*. London, UK: Routledge.
İlkkaracan, İ. (2012). Why so few women in the labour market in Turkey? *Feminist Economics*, 18(1), 1–37.
Khaddour, K., & Mazur, K. (2018). Syria town database. Harvard Dataverse, V1. https://doi.org/10.7910/DVN/YQQ07L
Khattab, N. (2002). Ethnicity and female labour market participation: a new look at the Palestinian enclave in Israel. *Work, Employment and Society*, 16(1), 91–110.
Kızılca, K. F. (2016). Breaking with dogma: unorthodox consumption patterns and women's labour market outcomes in Turkey. *Feminist Economics*, 22(4), 1–30.
Klasen, S. (2019). What explains uneven female labour force participation levels and trends in developing countries? *The World Bank Research Observer*, 34(2), 161–197.
Korotayev, A. V., Issaev, L. M., & Shishkina, A. R. (2015). Female labour force participation rate, Islam, and Arab culture in cross-cultural perspective. *Cross-Cultural Research*, 49(1), 3–19
Lehrer, E. L. (1995). The effects of religion on the labor supply of married women. *Social Science Research*, 24(3), 281–301.
Mancini, G. (2018). Women's labour force participation in Italy, 1861–2011. *Rivista di Storia Economica*, 34(1), 3–68.
Olivetti, C., & Petrongolo, B. (2016). The evolution of gender gaps in industrialized countries. *Annual Review of Economics*, 8, 405–434.
Reimers, C. W. (1985). Cultural differences in labour force participation among married women. *The American Economic Review*, 75(2), 251–255.
Ross, M. L. (2008). Oil, Islam, and women. *American Political Science Review*, 102(1), 107–123.
Saleh, M. (2016). Public mass modern education, religion, and human capital in twentieth-century Egypt. *The Journal of Economic History*, 76(3), 697–735.
Sassmannshausen, C. (2019). Educated with distinction: educational decisions and girls' schooling in late Ottoman Syria. *Journal of the Economic and Social History of the Orient*, 62(1), 222–256.
Sobek, M. (1995). The comparability of occupations and the generation of income scores. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 28(1), 47–51.
Stanfors, M., & Goldscheider, F. (2017). The forest and the trees: industrialization, demographic change, and the ongoing gender revolution in Sweden and the United States, 1870–2010. *Demographic Research*, 36, 173–226.
Tunalı, İ., Kırdar, M. G., & Dayıoğlu, M. (2021). Down and up the "U" – a synthetic cohort (panel) analysis of female labour force participation in Turkey, 1988–2013. *World Development*, 146, 105609.
Tzannatos, Z. (1999). Women and labour market changes in the global economy: growth helps, inequalities hurt and public policy matters. *World Development*, 27(3), 551–569.
Ugur, Z. B. (2018). The wearing of the headscarf & labour market outcomes for women in Turkey. *KADEM Kadın Araştırmaları Dergisi*, 4, 51–75.

UN, World Population Prospects (2022). Processed by Our World in Data. "Working age (ages 15–64)" [dataset]. UN, World Population Prospects (2022) [original data]. https://ourworldindata.org/grapher/population-young-working-old?country = Africa+%28UN%29~Europe+%28UN%29

Van Eijk, E. (2016). *Family law in Syria: Patriarchy, pluralism and personal status laws*. Oxford, UK: Bloomsbury Publishing.

World Bank. (1955). *The economic development of Syria: Report of a mission organized by the International Bank for Reconstruction and Development at the request of the Government of Syria*. Baltimore, MD: The Johns Hopkins Press.

World Bank. (2022). *Syria economic monitor – Spring 2022: Lost generation of Syrians*. Washington, DC: World Bank.

Yonay, Y. P., Yaish, M., & Kraus, V. (2015). Religious heterogeneity and cultural diffusion: the impact of Christian neighbors on Muslim and Druze women's participation in the labour force in Israel. *Sociology*, 49(4), 660–678.

Cite this article: Zalfou, R. (2025). Ethnic and religious differences in female labor force participation: evidence from Syrian census data. *Journal of Demographic Economics* 1–33. https://doi.org/10.1017/dem.2025.10007

Appendix

Table A1: Regression results – Income

	Dependent variable:								
	Female Labour Force Participation								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Income variable:	-	Education -based				Sector-based			
Ethnicity									
Sunni families	ref	ref	ref	ref	ref	ref	ref	ref	ref
Alawi	18.294***	13.997***	14.028***	11.616***	10.892***	13.196***	9.422***	9.242***	8.676***
	(0.837)	(0.695)	(0.772)	(0.747)	(0.741)	(0.971)	(1.188)	(1.187)	(1.207)
Christian	16.472***	6.163***	6.154***	5.621***	4.834***	14.981***	13.554***	13.492***	12.987***
	(2.381)	(1.402)	(1.399)	(1.339)	(1.334)	(2.250)	(2.311)	(2.318)	(2.242)
Druze	5.604***	4.762***	4.768***	4.136***	4.114***	5.295***	3.874***	3.862***	3.809***
	(1.593)	(0.825)	(0.834)	(0.783)	(0.790)	(1.157)	(1.049)	(1.047)	(1.071)
Kurdish	6.515***	1.327	1.318	1.637*	1.794**	1.736	2.330*	2.420**	2.648**
	(1.736)	(1.151)	(1.161)	(0.967)	(0.890)	(1.252)	(1.207)	(1.209)	(1.228)
Mixed	6.222***	1.418*	1.406*	2.074***	1.702**	4.298***	6.193***	6.169***	6.146***
	(1.680)	(0.845)	(0.825)	(0.659)	(0.698)	(1.261)	(0.826)	(0.826)	(0.892)
Other minorities	5.903**	3.153***	3.164***	3.309***	3.040**	5.194**	3.627*	3.692*	3.439*
	(2.450)	(1.186)	(1.215)	(1.161)	(1.184)	(2.108)	(2.025)	(2.017)	(1.966)
Sunni tribes	10.548***	5.322***	5.320***	5.010***	5.918***	4.893***	4.616***	4.652***	5.424***
	(1.328)	(0.927)	(0.929)	(0.815)	(0.764)	(1.131)	(1.009)	(1.008)	(0.997)
Income	-	2.874***	2.899***	3.800***	4.023***	3.161***	-1.335*	-0.932	-1.236
		(0.395)	(0.397)	(0.372)	(0.361)	(0.615)	(0.699)	(0.717)	(0.766)
Income^2	-	3.203***	3.202***	2.590***	2.546***	4.473***	3.499***	3.438***	3.288***
		(0.336)	(0.336)	(0.296)	(0.306)	(0.310)	(0.215)	(0.216)	(0.216)
Public sector	-	-	-0.002	0.037***	0.083***	-	0.218***	0.213***	0.264***
			(0.014)	(0.013)	(0.013)		(0.037)	(0.037)	(0.039)
Family sector	-	-	-	0.180***	0.183***	-	-	0.034**	0.035***
				(0.014)	(0.014)			(0.014)	(0.014)
MLFP	-	-	-	-	0.265***	-	-	-	0.219***
					(0.023)				(0.023)
Lat & Long	Y	Y	Y	Y	Y	Y	Y	Y	Y
Constant	8.222***	9.353***	9.407***	6.937***	-15.602***	6.167***	-0.074	-0.203	-19.260***
	(0.676)	(0.400)	(0.630)	(0.587)	(1.977)	(0.642)	(1.447)	(1.447)	(2.273)
Observations	5,175	5,175	5,175	5,175	5,175	5,175	5,175	5,175	5,175
Adjusted R^2	0.206	0.369	0.369	0.414	0.442	0.405	0.440	0.441	0.460
F Statistic	149.920***	275.613***	252.600***	282.115***	293.784***	321.454***	339.698***	314.989***	315.902***

Table A2: Regression results – Gender Parity Indices

	Dependent variable:						
	Female Labour Force Participation						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
GPI variable:	-	Enrolment (15 to 17)	Enrolment (15 to 17)	Secondary education	Secondary education	Tertiary education	Tertiary education
Ethnicity							
Sunni families	ref	ref	ref	ref	ref	ref	ref
Alawi	18.292***	18.380***	18.386***	17.252***	17.417***	17.504***	17.497***
	(0.837)	(0.838)	(0.838)	(1.080)	(1.086)	(1.047)	(1.055)
Christian	16.482***	16.562***	16.573***	15.209***	15.203***	15.849***	15.841***
	(2.382)	(2.389)	(2.390)	(2.305)	(2.262)	(2.446)	(2.441)
Druze	5.606***	5.653***	5.651***	5.316***	5.337***	5.667***	5.664***
	(1.593)	(1.592)	(1.591)	(1.522)	(1.497)	(1.540)	(1.538)
Kurdish	6.511***	6.432***	6.372***	6.087***	5.881***	6.255***	6.269***
	(1.737)	(1.716)	(1.708)	(1.838)	(1.720)	(1.913)	(1.914)
Mixed	6.222***	6.325***	6.345***	4.997***	5.223***	4.855***	4.767***
	(1.680)	(1.675)	(1.672)	(1.461)	(1.375)	(1.437)	(1.369)
Other minorities	5.904**	5.968**	5.945**	5.257**	5.458**	5.423**	5.386**
	(2.450)	(2.474)	(2.476)	(2.265)	(2.232)	(2.330)	(2.325)
Sunni tribes	10.548***	9.582***	9.364***	11.184***	10.661***	10.877***	10.920***
	(1.329)	(1.399)	(1.406)	(1.462)	(1.417)	(1.445)	(1.453)
GPI	-	-0.101***	-0.187***	0.037**	-0.014	0.057***	0.065**
		(0.029)	(0.043)	(0.015)	(0.030)	(0.019)	(0.026)
GPI^2	-	-	0.0004***	-	0.0004**	-	-0.0001
			(0.0001)		(0.0002)		(0.0001)
Lat & Long	Y	Y	Y	Y	Y	Y	Y
Constant	8.222***	18.089***	22.547***	6.393***	7.677***	6.721***	6.596***
	(0.676)	(2.826)	(3.419)	(0.951)	(0.983)	(0.834)	(0.817)
Observations	5,127	5,127	5,127	5,144	5,144	4,045	4,045
Adjusted R^2	0.206	0.213	0.215	0.213	0.218	0.241	0.241
F Statistic	148.705***	139.539***	128.597***	140.011***	131.297***	129.071***	117.431***

*p<0.1; **p<0.05; ***p<0.01

Table A3: Regression results – Educational attainment

	Dependent variable:					
	Female Labour Force Participation					
	(1)	(2)	(3)	(4)	(5)	(6)
Ethnicity						
Sunni families	ref	ref	ref	ref	ref	ref
Alawi	18.294***	15.190***	13.678***	13.507***	11.883***	11.200***
	(0.837)	(1.004)	(0.666)	(0.674)	(0.778)	(0.776)
Christian	16.472***	12.107***	7.284***	7.132***	4.542***	4.651***
	(2.381)	(2.004)	(1.249)	(1.263)	(1.055)	(1.068)
Druze	5.604***	4.404***	3.077***	3.024***	2.192**	1.479
	(1.593)	(1.292)	(0.862)	(0.896)	(1.060)	(1.037)
Kurdish	6.515***	6.612***	1.978*	1.592	2.650**	1.435
	(1.736)	(1.645)	(1.052)	(1.123)	(1.183)	(1.242)
Mixed	6.222***	3.180**	1.513*	1.389*	0.303	-0.106
	(1.680)	(1.541)	(0.821)	(0.842)	(0.985)	(0.936)
Other minorities	5.903**	4.547***	3.403***	3.282***	2.541***	2.242***
	(2.450)	(1.627)	(0.866)	(0.887)	(0.778)	(0.691)
Sunni tribes	10.548***	12.374***	5.183***	5.157***	5.591***	5.041***
	(1.328)	(1.507)	(0.941)	(0.930)	(1.054)	(1.040)
Female education years	-	1.890***	-7.744***	-6.373***	-	-
		(0.422)	(0.731)	(1.028)		
Female education years^2	-	-	0.997***	0.887***	-	-
			(0.074)	(0.102)		
Male education years	-	-	-	-2.181**	-	-
				(1.076)		
Male education years^2	-	-	-	0.154*	-	-
				(0.090)		
Female education shares	-	-	-	-	Y	Y
Male education shares	-	-	-	-	-	Y
Lat & Long	Y	Y	Y	Y	Y	Y
Constant	76.025***	14.506	7.974	9.672	65.992***	68.605***
	(19.142)	(12.442)	(9.405)	(9.712)	(13.901)	(13.242)
Observations	5,175	5,175	5,175	5,175	5,175	5,175
Adjusted R^2	0.206	0.250	0.378	0.379	0.381	0.395
F Statistic	149.920***	173.854***	286.539***	244.020***	200.010***	147.728***

*p<0.1; **p<0.05; ***p<0.01

Table A4: Regression results – Demographic variables

	Dependent variable:								
	Female Labour Force Participation								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ethnicity									
Sunni families	ref	ref	ref	ref	ref	ref	ref	ref	ref
Alawi	18.294***	18.158***	18.026***	17.949***	17.967***	11.403***	9.467***	13.039***	9.323***
	(0.837)	(0.901)	(0.923)	(0.822)	(0.816)	(1.284)	(1.842)	(1.187)	(1.403)
Christian	16.472***	16.245***	16.073***	15.343***	15.259***	9.160***	7.371***	10.848***	7.041***
	(2.381)	(2.445)	(2.484)	(2.418)	(2.387)	(2.376)	(2.661)	(2.363)	(2.401)
Druze	5.604***	5.399***	5.240***	4.988***	4.906***	-0.322	-2.045	1.097	-2.132
	(1.593)	(1.525)	(1.507)	(1.668)	(1.681)	(1.761)	(2.091)	(1.687)	(1.841)
Kurdish	6.515***	6.319***	6.119***	5.957***	5.929***	2.533*	2.068	3.340**	1.704
	(1.736)	(1.835)	(1.856)	(1.717)	(1.734)	(1.419)	(1.426)	(1.485)	(1.345)
Mixed	6.222***	6.033***	5.982***	5.422***	5.109***	1.843	1.864	2.612*	1.280
	(1.680)	(1.601)	(1.594)	(1.591)	(1.582)	(1.535)	(1.348)	(1.564)	(1.382)
Other minorities	5.903**	5.812**	5.810**	5.690***	5.597***	2.727	2.778**	3.271*	2.390*
	(2.450)	(2.439)	(2.432)	(2.188)	(2.151)	(1.702)	(1.406)	(1.894)	(1.387)
Sunni tribes	10.548***	10.616***	10.487***	10.046***	10.122***	11.188***	9.688***	11.469***	9.797***
	(1.328)	(1.326)	(1.347)	(1.230)	(1.246)	(1.260)	(1.195)	(1.352)	(1.090)
Marriage fertility rate	-	-0.122	-0.519	-	-	-	-	-	-
		(0.134)	(0.352)						
Marriage fertility rate^2	-	-	0.027	-	-	-	-	-	-
			(0.023)						
Female headed families	-	-	-	0.387***	0.799***	-	-	-	-
				(0.104)	(0.179)				
Female headed families^2	-	-	-	-	-0.018***	-	-	-	-
					(0.005)				
Aged 15 to 64	-	-	-	-	-	0.778***	-4.533**	-	-
						(0.119)	(2.076)		
Aged 15 to 64^2	-	-	-	-	-	-	0.047**	-	-
							(0.019)		
Age dependency ratio	-	-	-	-	-	-	-	-0.205***	-1.121***
								(0.034)	(0.224)
Age dependency ratio	-	-	-	-	-	-	-	-	0.005***
									(0.001)
Lat & Long	Y	Y	Y	Y	Y	Y	Y	Y	Y
Constant	8.222***	9.063***	10.427***	5.708***	3.922***	-33.366***	117.211**	26.166***	64.498***
	(0.676)	(1.384)	(1.786)	(1.082)	(1.311)	(6.664)	(56.862)	(2.770)	(9.755)
Observations	5,175	5,175	5,175	5,175	5,175	5,175	5,175	5,175	5,175
Adjusted R^2	0.206	0.206	0.206	0.219	0.223	0.274	0.301	0.255	0.301
F Statistic	149.920***	135.261***	123.297***	146.482***	135.971***	195.944***	203.539***	178.528***	203.499***

*p<0.1; **p<0.05; ***p<0.01

Table A5: Regression results – Demographic variables

	Dependent variable:						
	Female Labour Force Participation						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ethnicity							
Sunni families	ref	ref	ref	ref	ref	ref	ref
Alawi	18.294***	18.230***	18.195***	17.954***	17.853***	18.290***	18.364***
	(0.837)	(0.819)	(0.806)	(0.814)	(0.787)	(0.835)	(0.835)
Christian	16.472***	16.486***	16.439***	16.628***	16.644***	16.462***	16.474***
	(2.381)	(2.426)	(2.463)	(2.427)	(2.448)	(2.388)	(2.388)
Druze	5.604***	5.592***	5.621***	4.864***	4.840***	5.633***	5.765***
	(1.593)	(1.591)	(1.586)	(1.846)	(1.864)	(1.589)	(1.589)
Kurdish	6.515***	5.825***	5.761***	6.149***	6.195***	6.436***	6.235***
	(1.736)	(1.607)	(1.624)	(1.620)	(1.649)	(1.724)	(1.724)
Mixed	6.222***	6.379***	6.413***	6.597***	6.673***	6.228***	6.193***
	(1.680)	(1.605)	(1.582)	(1.582)	(1.520)	(1.675)	(1.675)
Other minorities	5.903**	5.862**	5.812**	6.028**	6.000**	5.894**	5.915**
	(2.450)	(2.451)	(2.439)	(2.438)	(2.427)	(2.449)	(2.449)
Sunni tribes	10.548***	10.006***	9.955***	9.953***	9.907***	10.434***	10.227***
	(1.328)	(1.416)	(1.441)	(1.403)	(1.397)	(1.344)	(1.344)
Drinking water	-	-0.024*	0.027	-	-	-	-
		(0.013)	(0.050)				
Drinking water^2	-	-	-0.0005	-	-	-	-
			(0.001)				
Sanitation network	-	-	-	-0.018*	0.010	-	-
				(0.010)	(0.046)		
Sanitation network^2	-	-	-	-	-0.0003	-	-
					(0.0005)		
Electricity network	-	-	-	-	-	-0.030	0.151***
						(0.025)	(0.025)
Electricity network^2	-	-	-	-	-		-0.002
Lat & Long	Y	Y	Y	Y	Y	Y	Y
Constant	76.025***	79.686***	78.974***	79.177***	79.006***	78.588***	74.127***
	(19.142)	(19.658)	(19.490)	(19.723)	(19.449)	(19.602)	(19.602)
Observations	5,175	5,175	5,175	5,175	5,175	5,175	5,175
Adjusted R^2	0.206	0.208	0.208	0.208	0.208	0.206	0.207
F Statistic	149.920***	136.938***	124.841***	136.841***	124.580***	135.297***	123.900***

*p<0.1; **p<0.05; ***p<0.01

Table A6: Regression results – Private property institutions

	Dependent variable:								
	Female Labour Force Participation								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ethnicity									
Sunni families	ref	ref	ref	ref	ref	ref	ref	ref	ref
Alawi	18.294***	18.297***	17.949***	18.287***	18.334***	17.786***	17.402***	18.390***	18.631***
	(0.837)	(0.820)	(0.770)	(0.826)	(0.827)	(0.806)	(0.801)	(0.837)	(0.815)
Christian	16.472***	16.469***	16.205***	16.476***	16.454***	15.964***	15.775***	16.261***	16.308***
	(2.381)	(2.383)	(2.435)	(2.387)	(2.408)	(2.371)	(2.364)	(2.463)	(2.483)
Druze	5.604***	5.595***	5.256***	5.635***	5.919***	4.973***	4.514***	5.857***	6.127***
	(1.593)	(1.536)	(1.636)	(1.542)	(1.488)	(1.633)	(1.663)	(1.544)	(1.533)
Kurdish	6.515***	6.533***	6.244***	6.493***	6.612***	7.061***	6.353***	6.444***	6.472***
	(1.736)	(1.639)	(1.571)	(1.690)	(1.732)	(1.651)	(1.469)	(1.709)	(1.728)
Mixed	6.222***	6.223***	6.783***	6.218***	5.998***	6.756***	8.495***	6.022***	5.553***
	(1.680)	(1.688)	(1.519)	(1.688)	(1.722)	(1.605)	(1.483)	(1.675)	(1.618)
Other minorities	5.903**	5.914**	6.084**	5.908**	5.921**	6.943***	7.501***	5.807**	5.621**
	(2.450)	(2.452)	(2.431)	(2.450)	(2.448)	(2.266)	(2.414)	(2.475)	(2.515)
Sunni tribes	10.548***	10.559***	10.464***	10.527***	10.558***	9.912***	9.796***	10.425***	10.432***
	(1.328)	(1.347)	(1.322)	(1.323)	(1.317)	(1.236)	(1.190)	(1.314)	(1.326)
Official deeds	-	0.001	-0.074	-	-	-	-	-	-
		(0.012)	(0.048)						
Official deeds^2	-	-	0.001*	-	-	-	-	-	-
			(0.0004)						
Agricultural deeds	-	-	-	0.003	0.055	-	-	-	-
				(0.015)	(0.052)				
Agricultural deeds^2	-	-	-	-	-0.001	-	-	-	-
					(0.001)				
Public notary	-	-	-	-	-	-0.082**	-0.331***	-	-
						(0.038)	(0.128)		
Public notary^2	-	-	-	-	-	-	0.005**	-	-
							(0.002)		
Owned dwellings	-	-	-	-	-	-	-	-0.040***	0.091
								(0.013)	(0.063)
Owned dwellings^2	-	-	-	-	-	-	-	-	-0.001*
									(0.001)
Lat & Long	Y	Y	Y	Y	Y	Y	Y	Y	Y
Constant	76.025***	75.837***	79.619***	75.899***	73.084***	74.331***	74.285***	78.464***	72.948***
	(19.142)	(20.954)	(21.644)	(18.767)	(17.251)	(18.417)	(16.766)	(19.322)	(18.887)
Observations	5,175	5,175	5,175	5,175	5,175	5,175	5,175	5,175	5,175
Adjusted R^2	0.206	0.206	0.208	0.206	0.206	0.214	0.225	0.208	0.210
F Statistic	149.920***	134.904***	124.569***	134.926***	123.362***	141.931***	137.486***	137.131***	126.148***

*p<0.1; **p<0.05; ***p<0.01

Table A7: Regression results – Geography variables

	Dependent variable:						
	Female labour force participation						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ethnicity							
Sunni families	ref	ref	ref	ref	ref	ref	ref
Alawi	18.067***	13.670***	16.214***	17.393***	14.272***	17.659***	17.695***
	(0.593)	(1.071)	(1.408)	(0.637)	(0.866)	(0.762)	(0.655)
Christian	15.475***	14.184***	14.927***	14.951***	15.461***	16.298***	16.003***
	(2.434)	(2.251)	(2.296)	(2.434)	(2.598)	(2.509)	(2.471)
Druze	6.942***	10.719***	5.644***	7.393***	6.608***	5.380***	5.403***
	(1.146)	(1.528)	(2.162)	(1.256)	(1.160)	(1.696)	(1.687)
Kurdish	2.193**	8.492***	1.894*	1.803	3.475***	6.589***	6.053***
	(1.063)	(1.468)	(1.149)	(1.135)	(1.187)	(1.669)	(1.537)
Mixed	5.084**	6.090***	4.697*	4.682**	4.102***	6.704***	6.245***
	(2.361)	(1.052)	(2.406)	(2.254)	(1.511)	(1.508)	(1.595)
Other minorities	5.335**	6.140**	5.331**	5.050**	4.431**	6.259**	5.853**
	(2.198)	(2.468)	(2.112)	(2.166)	(2.258)	(2.435)	(2.517)
Sunni tribes	6.787***	12.526***	4.933***	5.130***	6.119***	10.455***	10.270***
	(0.834)	(1.334)	(1.131)	(1.701)	(0.850)	(1.364)	(1.381)
Latitude	-	93.343***	-	-	-	-0.930	-0.911
		(20.267)				(0.714)	(0.681)
Latitude^2	-	-1.338***	-	-	-	-	-
		(0.290)					
Longitude	-	-107.117***	-	-	-	-1.059**	-0.946*
		(24.178)				(0.511)	(0.512)
Longitude^2	-	1.371***	-	-	-	-	-
		(0.311)					
Slope	-	-	-0.019	-	-	-	-
			(0.320)				
Slope^2	-	-	0.002	-	-	-	-
			(0.009)				
Elevation	-	-	-0.024***	-	-	-	-
			(0.008)				
Elevation^2	-	-	0.00002***	-	-	-	-
			(0.00001)				
Temperature	-	-	-	-1.651***		-	-
				(0.602)			
Temperature^2	-	-	-	0.062**		-	-
				(0.031)			
Precipitation	-	-	-	-	-0.082***	-	-
					(0.020)		

Table A7 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Precipitation^2	-	-	-	-	0.0002***	-	-
					(0.00004)		
Flow accumulation	-	-	-	-	0.00004	-	-
					(0.00005)		
Flow accumulation^2	-	-	-	-	-0.000	-	-
					(0.000)		
Wheat suitability	-	-	-	-	-	-	-
						(0.001)	
Wheat suitability^2	-	-	-	-	-	-0.00000	-
						(0.00000)	
Cotton suitability	-	-	-	-	-	-	0.001
							(0.001)
Cotton suitability^2	-	-	-	-	-	-	-0.00000
							(0.00000)
Constant	10.273***	467.663	17.639***	20.503***	17.512***	76.957***	75.403***
	(0.292)	(405.617)	(2.568)	(3.248)	(2.029)	(18.723)	(17.692)
Observations	5,180	5,175	5,180	5,180	5,180	5,175	5,175
Adjusted R^2	0.183	0.268	0.220	0.197	0.283	0.209	0.209
F Statistic	167.194***	173.167***	133.609***	142.158***	186.808***	125.549***	125.613***

*p<0.1; **p<0.05; ***p<0.01

Table A8: Full regression outputs for the adjusted models in decomposition analyses

	Dependent variable:	
	Female labour force participation	
	(1)	(2)
Ethnicity		
Sunni families	Ref.	Ref.
Alawi	4.921***	4.908***
	(0.553)	(0.620)
Christian	2.950***	2.053*
	(1.120)	(1.138)
Druze	-0.174	0.389
	(0.900)	(1.012)
Kurdish	-1.633**	-0.527
	(0.746)	(0.775)
Mixed	1.116***	1.830***
	(0.329)	(0.382)
Other minorities	1.042*	0.688
	(0.603)	(0.610)
Sunni tribes	3.957***	3.277***
	(0.439)	(0.520)
Demographic variables		
Aged 15 to 64	12.192*	2.243
	(7.316)	(7.572)
Aged 15 to 64^2	-0.045	-0.003
	(0.034)	(0.035)
Age dependency ratio	2.959*	0.696
	(1.551)	(1.594)
Age dependency ratio^2	-0.005*	-0.001
	(0.003)	(0.003)
Female headed families	0.573***	0.431***
	(0.074)	(0.076)
Female headed families^2	-0.013***	-0.010***
	(0.003)	(0.003)
Fertility rate	1.267***	0.711***
	(0.189)	(0.190)
Fertility rate^2	-0.074***	-0.043***
	(0.012)	(0.012)

Table A8 (continued)

	(1)	(2)
Income and employment		
Income (education)	3.876***	1.783**
	(0.279)	(0.853)
Income (education)^2	0.665***	0.051
	(0.129)	(0.185)
Income (sectors)	-4.655***	-3.810***
	(0.331)	(0.357)
Income (sectors)^2	2.515***	2.619***
	(0.122)	(0.122)
Public sector	0.201***	0.152***
	(0.013)	(0.014)
Family sector	0.031***	0.035***
	(0.009)	(0.009)
MLFP	0.257***	0.258***
	(0.016)	(0.016)
Gender parity indices		
GPI (enrolment)	-0.010	0.020
	(0.022)	(0.024)
GPI (enrolment)^2	-0.00002	-0.0001
	(0.0001)	(0.0001)
GPI (secondary)	0.050***	-0.023*
	(0.012)	(0.014)
GPI (secondary)^2	-0.0001**	0.0001
	(0.0001)	(0.0001)
GPI (tertiary)	0.0001	0.008
	(0.009)	(0.010)
GPI (tertiary)^2	0.00001	0.00000
	(0.00004)	(0.00004)
Education		
Tertiary		-0.400*
		(0.238)
Middle Academy		0.695***
		(0.099)
Upper secondary		0.318***
		(0.093)
Lower secondary		-0.180***
		(0.048)
Elementary		-0.043**
		(0.019)
Literate		-0.034**
		(0.015)

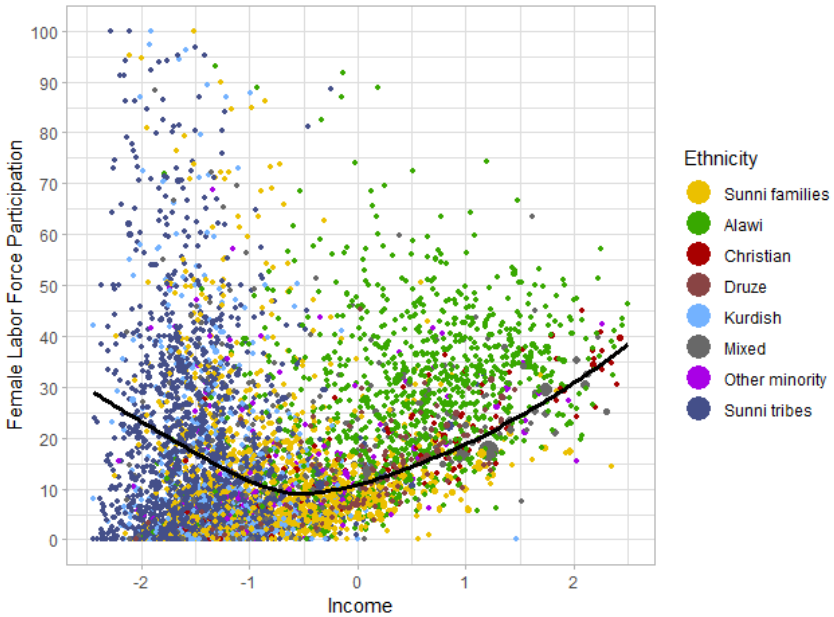
Table A8 (continued)

	(1)	(2)
Infrastructure		
Drinking water network		0.018***
		(0.006)
Sanitation network		0.001
		(0.005)
Electricity network		0.018
		(0.015)
Crowding rate		0.958**
		(0.426)
Property institutions		
Official deeds		0.007
		(0.005)
Agricultural deeds		-0.003
		(0.006)
Public notary		0.037***
		(0.010)
Owned dwellings		-0.026***
		(0.008)
Geography		
Longitude		0.063
		(0.419)
Latitude		-0.814**
		(0.379)
Longitude^2		-0.137
		(0.112)
Latitude^2		-0.087
		(0.113)
Slope		0.019
		(0.076)
Slope^2		0.001
		(0.002)
Elevation		-0.005**
		(0.002)
Elevation^2		0.00000
		(0.00000)
Temperature		0.718**
		(0.352)

Table A8 (continued)

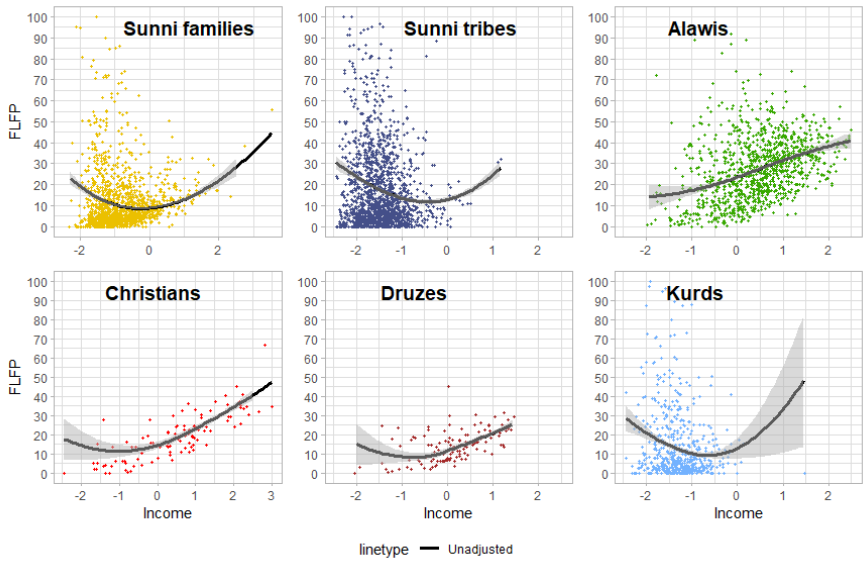
	(1)	(2)
Temperature^2		-0.033**
		(0.016)
Precipitation		-0.016*
		(0.008)
Precipitation^2		0.00004***
		(0.00001)
Flow accumulation		0.00002
		(0.00003)
Flow accumulation^2		-0.000
		(0.000)
Wheat suitability		0.0001
		(0.001)
Wheat suitability^2		0.000
		(0.00000)
Cotton suitability		0.001
		(0.0004)
Cotton suitability^2		-0.00000**
		(0.00000)
Constant	-766.388*	-187.532
	(409.233)	(422.690)
Observations	5,036	5,036
R^2	0.586	0.611
Adjusted R^2	0.584	0.606
Residual Std. Error	404.303	393.235
F Statistic	236.595***	130.311***
Note:	*p<0.1; **p<0.05; ***p<0.01	

Figure A1: Female labour force participation as a function of income (education-based)



Note: The income measure is imputed based on the educational composition of the labour force. The line of fit is estimated using a non-parametric loess function with a span of 0.75. Observations are weighed according to population size.

Figure A2: Female labour force participation as a function of income (education-based) by ethnic group



Note: The income measure is imputed based on the educational composition of the labour force. The line of fit is estimated using a non-parametric loess function with a span of 0.75. Observations are weighed according to population size.

Paper III

European Review of Economic History, **29**, 244–271

doi:10.1093/ereh/heae020 Advance Access Publication Date: February 10, 2025

Premium or Penalty? Occupations and Earnings of Ottoman Immigrants and Their Offspring in the United States, 1900–1940

Rami Zalfou[1,2,3,*], Martin Dribe[1,2]
[1]*Department of Economic History, Lund University, Scheelevägen 15, 223 63 Lund, Sweden*
[2]*Centre for Economic Demography, Lund University, Scheelevägen 15, 223 63 Lund, Sweden*
[3]*Centre for Advanced Middle Eastern Studies, Lund University, Finngatan 16, 223 62 Lund, Sweden*

Corresponding author: Centre for Economic Demography, Department of Economic History, Lund University, Scheelevägen 15, 223 63 Lund, Sweden.
Email: Rami.Zalfou@ekh.lu.se

We study the economic integration of immigrants from Ottoman Syria and Turkey and their offspring in the United States using full count census data from 1900 to 1940. Immigrants initially achieved occupational premiums due to their selection into high-reward industries, but 1940 earnings data reveals significant and growing disadvantages over time, partly due to lower educational attainment. In contrast, the second generation achieved substantial upward mobility, closing both the education and earnings gaps with native Whites. This contrasts with the experience of Northern European immigrants who matched natives more closely in terms of occupations and earnings.

1. Introduction

The economic assimilation of immigrants and their offspring in the United States has long been debated in economic history (see Abramitzky and Boustan 2017). However, existing research largely centers on two groups: pre-WWII European immigrants (e.g., Abramitzky et al. 2014; Hatton 1997; Minns 2000) and postwar immigrants from developing countries (e.g., Algan et al. 2010; Grogger and Gordon 2011). This paper fills an important gap by examining the economic integration of an overlooked group: immigrants from the Middle East during the first four decades of the 20th century and their offspring. Using comprehensive census data and the longitudinal IPUMS-MLP panel (Helgertz et al. 2020; Ruggles et al. 2021), we document their unique assimilation patterns and compare them to those of other immigrant groups, shedding light on both their challenges and successes.

Middle Eastern migration to the USA, particularly from the Ottoman Empire, has been understudied in economic history. While small in absolute terms compared to European migration flows, Ottoman migration to the Americas had a significant demographic impact in source countries, particularly in regions such as Mount Lebanon, where an estimated quarter of the population emigrated (Naff 1985). In the USA, Ottoman immigrants established vibrant communities, such as Little Syria in New York City, and engaged in mobile peddling networks that connected remote communities (Naff 1985; Karpat 1985a, 1985b).

This paper provides new evidence on the assimilation of Ottoman immigrants, comparing their economic outcomes to those of US natives and European immigrant groups. We show that Ottoman immigrant men initially held significant occupational premiums due to their selection into high-reward occupation and industries. However, their real earnings in 1940 lagged significantly behind those of US-born counterparts, with no evidence of convergence over time. In contrast, second-generation Ottoman immigrants achieved substantial upward mobility, closing most of the earnings gap with natives and even achieving earnings premiums in some cases. We compare the trajectories of Ottoman immigrants with those of Irish and Italian immigrants and find both differences and similarities. The "Old European" immigrants from North and Western Europe appear to have matched or exceeded the earnings of US natives, as in the case of Irish immigrants. However, the Ottoman experience resembles that of Southern European immigrants as in the case of Italian immigrants.

To understand the underlying factors behind this assimilation pattern, we present results on the role of human capital. Our results suggest that observed educational attainment explains substantial parts of the penalty in actual earnings among immigrants, and that the closing of the educational attainment gap with natives in the second generation plays a crucial role in their earnings assimilation and achieving earnings parity with natives, even after accounting for the effect geography. The findings here mirror earlier findings on the role of schooling in the upward mobility of second-generation Irish immigrants (Connor 2020), and contrasts with research that underplays the role of education in immigrant assimilation (Abramitzky et al. 2021; Ward 2020).

The paper proceeds as follows: We first provide a brief background on the history of Ottoman emigration. We proceed by discussing the theory and previous research on immigrant assimilation. Then, we describe the data and methods, followed by our empirical analysis and discussion of the results.

2. Ottoman migration to the USA

The regions of present-day Turkey, Syria, and Lebanon were part of the Ottoman Empire until its end in World War I. We use the term "Ottoman" to refer to immigrants in the USA who initially came from these Ottoman regions and continued to emigrate after the fall of the empire and the establishment of new nations, recognizing the continuity in migrant flows before and after the war. We categorize these immigrants into two groups: those from Turkey, which correspond roughly to those migrating from the regions of the present-day Turkish republic, and those from Syria-Lebanon, who came from the regions of these two present-day nations. The reason for combining the last two is twofold: they were initially called Turks in the USA but later came to be known as Syrians. This is partly because the term Syria referred at the time to a region that included both countries today, while the term Lebanon was seldom used to identify a nation before World War I. The data sources we use do not distinguish clearly between Lebanon and Syria in the period under study, and often codes individuals from Lebanon as Syrian, with no way of separating the two in the census data.

Ottoman law afforded a measure of conditional toleration and allowed the largest religious groups of the empire a significant degree of autonomy in running their affairs and freedom in undertaking economic activities alongside their Muslim counterparts. While sectarian conflicts took place in Lebanon and Syria during the middle of the 19th century, the period that saw the largest growth in migration volumes was a stable period that is referred to as the long peace (Akarli 1993). In contrast, the number of immigrants from Ottoman Turkey

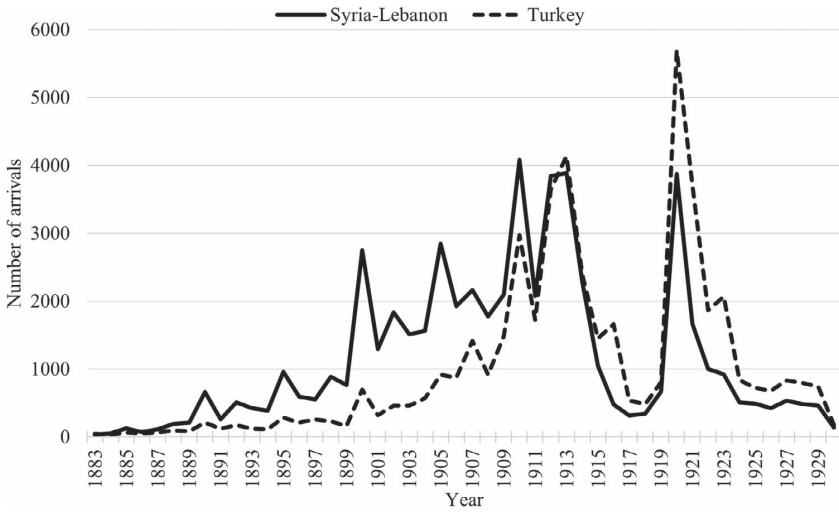


Figure 1. The number of Ottoman immigrants in the 1930 US census by arrival year and region of origin

reached peaks before and after World War I, and this migration took place in the context of forced displacement and large-scale population movements.

The 19th century saw the beginning of the Syrian-Lebanese migration to the Americas, particularly the USA, Canada, Brazil, and Argentina, alongside migration to West Africa and Australia. The number of migrants grew slowly in the second part of the 19th century and hastened in the early 20th century as economic conditions worsened and World War I took its toll. The increasing trend in emigration also coincided with the end of the prohibition on emigration from the Ottoman Empire in 1896. Karpat (1985a, 1985b) estimates that 500,000 to 800,000 Greeks, Armenians, and Arabs emigrated from the Ottoman Empire between 1854 and 1908 (before emigration rates reached their peak). The Ottoman Empire had a population of nearly 21 million in the Ottoman census of 1906, including the populations of Bulgaria, the Balkans, and Libya. According to Saliba (1981), Lebanon with a population of 442,000 in 1913 saw more than 100,000 emigrate by 1914.

Economic pressures likely played an important role, particularly amongst immigrants from Syria and Lebanon, with the region becoming increasingly integrated into world markets in the 19th century and with increased commercialization of agriculture. The economy of the region also suffered from the epidemics and crop failures which threatened the cities with famines (Fawaz 1983). The move from production for local consumption toward export also meant that some regions suffered from vulnerability to fluctuations in global markets, particularly evident in the famine that affected Mount Lebanon during World War I as its economy was blockaded by Allied forces and came under pressure from the Ottoman war efforts (Brand 2015). The economic pressures associated with the famine likely played a major role as a push factor in migration. The integration with European markets also led to decline in manufacturing and crafts which were not able to compete with the developments in European and Asian industries such as the introduction of artificial silk.

Table 1. *Place of birth of Ottoman immigrants in the Social Security NUMIDENT sample.*

Syria-Lebanon (%)		Turkey (%)	
Beirut	18	Istanbul	22
Aleppo	15	Izmir	6
Damascus	9	Harput	3
Zahlé	3	Aintab	3
Tripoli	2	Adana	3
Homs	2	Marash	2
		Armenia	2
Other cities and regions	50	Other cities and regions	58
Total sample	14,247	Total sample	15,747

Source: Social Security Administration Numerical Identification (NUMIDENT) data.

Figure 1 shows the number of yearly arrivals in the USA that are recorded in the population census of 1930. This only accounts for immigrants who came before the census was taken, and who remained in the USA and were alive in 1930. While the numbers become less representative the further back in time we go, they show the trends in the migration flow across time. There was some migration from the Ottoman Empire as early as the mid-19th century, but the volumes increased after 1890. Migration from Syria-Lebanon grew until World War I, while migration from Turkey only began to match that from Syria-Lebanon after 1910. After World War I, the migration flow from Turkey outgrew that of Syria-Lebanon on account of the upheavals and the population movements out of the country.

Previous research on the Syria-Lebanese migration to the USA is based on a variety of archival sources, such as official government records from the Ottoman, French, and US archives, as well as personal records and interviews with the immigrants (Abraham and Abraham 1983; Naff 1985). Hashimoto (1992) studies the reports of the Society of Nations and of the French Mandate which suggest that Lebanese migration was characterized by a reversible flow, with the proportion of return migrants estimated to be around 50% of all Lebanese emigrants in the period 1926 to 1933. However, this proportion is likely to differ for different destination countries. A report by a French mandate official claim that of 9188 Syrian migrants registered at their entrance to the United States between 1908 and 1909. 8725 (95%) were reported to have returned to their homeland in the following years (Hashimoto 1992).

Most Lebanese/Syrian emigrants were of peasant origin and lacked formal education (Issawi 1988). The age of mass migration coincided with increased global trade links, high demand for labor in America, and the relaxation of travel restrictions from the Ottoman Empire alongside the presence of missionary activity in Lebanon and Syria (Lindner 2009); all of which may have contributed to increasing emigration rates. The emigration from Ottoman Turkey to the USA has received less attention. The diverse nature of this group of immigrants meant that the studies that do focus on emigration from Ottoman Turkey to the USA are histories of specific communities, such as Armenians, Greeks, or Ottoman Jews, but never treat the emigration as a single phenomenon (e.g., Fittante 2017; Naar 2015).

The historical evidence suggests that economic incentives were at play as both push and pull factors. High demand for labor and new economic opportunities afforded by migration to the Americas affected those in the Ottoman Empire as it had in other nations in Europe and elsewhere. The large wage differentials between the industrializing countries and the stagnating Ottoman economies along with high potential payoffs to migration produced

strong incentives to emigrate. Economic decline in large parts of the empire meant that many were pushed to preserve and improve on their standard of living against unfavorable conditions in their home country (Khater 2001). Furthermore, the emigration of Greeks and Armenians was influenced by persecution at home, both during the Hamidian era (1876–1909) and following the Armenian genocide (1915–1917)—but not the emigration of Syrians and Lebanese before World War I, who experienced a long period of peace after the events of 1860, away from the conflicts and upheavals in Ottoman Turkey.

We use the Social Security Administration Numerical Identification (NUMIDENT) data made available by the National Archives[1] to identify the place of birth of the immigrants as shown in table 1.[2] Arrivals from the three major cities of each of Lebanon and Syria make up around 50% of the sample, while the rest of the sample came from various other towns and locales. This suggests a high degree of immigrant selection with a bias toward urban origins. As a comparison, the proportion of the total population that is urban in Syria-Lebanon and Turkey before World War I was much lower than 50% and may have been as low as 14% in 1920 and only began to increase significantly after 1950 (United Nations 1969).

When looking at the column for Turkey, we find that the two major cities of Istanbul and Izmir account for almost a third of all migrants. After the two largest cities, the southeast of Anatolia appears most dominant, which is where many Armenian refugees came from after the Hamidian massacres and later the Armenian genocide. Some immigrants gave Armenia as their place of birth, which probably refers to "Western Armenia" or the six Armenian provinces of eastern Anatolia (as opposed to the Armenian provinces of the Russian empire which today makeup the Republic of Armenia). We also find that a significant portion of immigrants came from various other cities and regions, with the major cities accounting for less than half of the immigrants in the sample.

Another way to proxy for the origin of the immigrants is to look at the mother tongue variable that is available in the census data. Table 2 shows the breakdown of the mother tongue for the immigrants in the 1930 full population census. Almost all the immigrants from Syria-Lebanon spoke Arabic, while the immigrants from Turkey were far more diverse. Armenian speakers account for just over a third of immigrants from Turkey, with Greek speakers accounting for 25% and Turkish speakers for only 18%. We find that Spanish speakers account for 11% of this group—this is likely due to the emigration of Judeo-Spanish (Ladino) speakers, a community that came to the Ottoman Empire after its expulsion from Spain in 1492.

[1] See Federal Security Agency, Social Security Board, and Social Security Administration, "Numerical Identification Files (NUMIDENT), 1936–2007."

[2] While the NUMIDENT data is limited in its coverage, it includes the place of birth at the sub-national level which was not recorded in the census. This means that we can identify the sending areas in each sending country. Selection into the sample is determined by a complex process, but it can be summarized as including some individuals who made Social Security applications or whose death was reported between 1937 and 2007 to the Social Security Administration. Selection into the data is not random, but there is no reason to suggest that it is not representative of the overall population we are interested in. The data does not include the date of arrival of immigrants but includes their place of birth and year of birth. We limit the sample further to those who were born before 1930 and exclude those with no birthplace information. While some in the sample may have arrived after the period 1900 to 1940 and it is possible that the importance of different sending regions changed through time, breaking these proportions down by birth decade does not show much difference across birth cohorts, suggesting that the sample obtained from the NUMIDENT file is representative of the overall migrant population from the Ottoman empire in the US during the period of interest.

Table 2. *Mother tongue of Ottoman immigrants in the 1930 census.*

Mother tongue (%)	Syria-Lebanon	Turkey
Arabic	94.8	2.0
Armenian	1.3	35.7
Greek	0.3	24.8
Hebrew or Yiddish	0.3	2.5
Other	2.5	6.0
Spanish	0.7	11.2
Turkish	0.2	17.8
Total sample	56,194	48,499

Source: authors' calculations based on the US census of 1930.

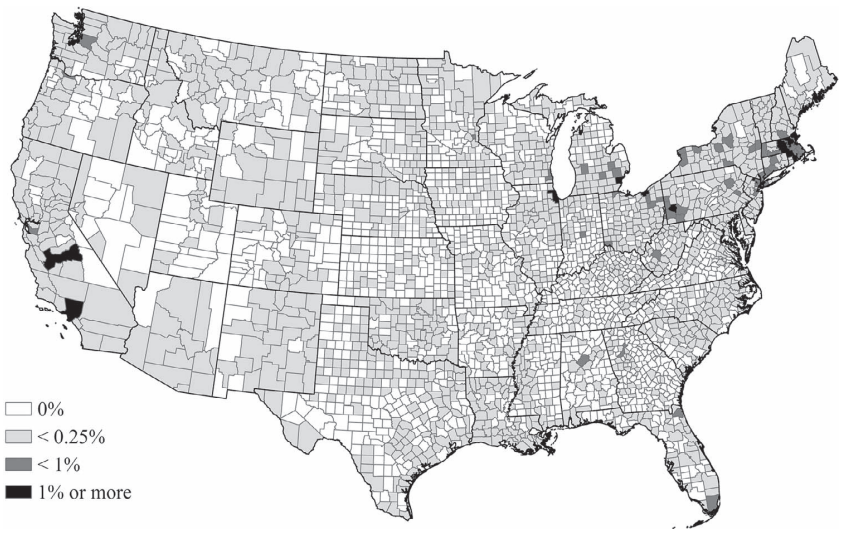


Figure 2. Map of the distribution of Ottoman immigrants and their offspring in the USA in 1940. The color categories represent the share in each county of the overall population of Ottoman immigrants and their offspring in 1940. The total population includes anyone born in Turkey, Syria, and Lebanon and anyone with a parent born in those countries, with a total of 220,986 individuals.

The settlement pattern of Ottoman immigrants in the USA is shown in figure 2, which shows the share in each US county of the total population of Ottoman immigrants and their offspring in the USA. A few counties in the Northeast account for more than 1% of the Ottoman population, with the Northeast states together accounting for around 56% of the total. California accounts for a further 7%, with the remainder distributed across the rest of the USA. The concentration in the Northeast states is typical for immigrants in the USA in the early 20th century. The highest county concentrations are in Brooklyn, New York with 8.2% of the total Ottoman population, and Wayne, Michigan with 6.6%.

The literature on Syrian-Lebanese immigrants in the USA emphasizes commerce and peddling as common activities for newly arrived immigrants, especially in the earliest immigration waves. Assimilation in the host country's labor market was ostensibly achieved through small businesses activity, such as grocery and retail stores, alongside small-scale manufacturing including clothes manufacturing and foodstuff production (Naff 1985). Employment in existing firms may not have been as common, with exceptions as in the Detroit area where many Ottoman immigrants worked in the automobile industry, alongside the more common commercial activities. While many Ottoman immigrants were of Christian background like their American and European counterparts, there was still a cultural gulf between the newcomers and the US population. The norms of the host country required adjustment from the newcomers, and they likely faced discrimination due to their unusual foreign origin and as citizens of the Ottoman Empire. Though the American press often portrayed Syrian and other Ottoman immigrants in in a positive light (Baycar 2020), and Americans largely sympathized with the plight of Armenians and Greeks fleeing persecution in Ottoman Anatolia.

Adjustment was also required by the US population and institutions. For example, Syrian-Lebanese immigrants undertook efforts to gain the rights held by other immigrant groups, including the right to US citizenship, which required that they be recognized as white persons from a legal standpoint (Gualtieri 2001). Like other immigrant groups, Ottoman immigrants established institutions that catered to their interests, including churches and mosques, printing presses in their native language, and clubs and societies, alongside political organizations that lobbied on their behalf. Despite gaining the right to become US citizens, the immigration act of 1924 affected the Ottoman immigrants by blocking further migration from that part of the world at any significant scale. Restrictions on the immigrants extended even to forced deportations, which included the deportation of Ottoman Greek refugees who were sent to Greece after the disappearance of their original communities in Anatolia. Though the number of deportations was tiny compared to the number of immigrants (Gratien and Pope-Obeda 2020).

3. Theory and previous research

We define economic integration as the participation of immigrants in the labor market of the host country and its rewards in accordance with that of their native counterparts. Economic integration entails that immigrants and their offspring earn wages that are similar to those of comparable natives. The human capital model is a starting point for understanding economic integration. Immigrants typically arrive in the host country lacking country-specific skills and knowledge, which may include language, occupation-specific knowledge that differs across countries, and social norms that are part of the host country workplace culture. Therefore, we expect immigrants to begin earning wages that are lower than their native counterparts, but they would experience faster wage growth as they accumulate country-specific skills. The accumulation of these skills leads to wage convergence, which can be measured as a shrinking of the wage gap between immigrants and natives with time spent in the host country. The investment in human capital may also be associated with forgone earnings in the period after arrival, with the aim of achieving a higher rate of future wage growth, which fits with the idea of wage convergence (Borjas 2014, 2000).

Early research on the economic integration of immigrants in the US labor market demonstrated that newly arrived immigrants earned lower wages than natives, but their earnings tended to grow rapidly and even surpass the earnings of natives after a number of

years (Chiswick 1978). A large historical literature developed which used a unified theoretical framework and empirical methods and applied them to different data sets to study the income convergence of various immigrant cohorts, and largely confirmed the earlier findings (reviewed in Abramitzky and Boustan 2017). However, methodological developments in the measurement of income gaps across time led to a challenge of the accepted wisdom. Borjas (1985) used linked census data to follow immigrant arrival cohorts across time and found that a significant part of the income convergence observed in cross-sectional data was driven by the declining skill level in later migrant cohorts as compared to earlier cohorts, which significantly inflated the estimated rate of convergence. Borjas also found a tendency toward return migration for the lowest skilled migrants within a cohort, where those with the lowest wages tended to return to their country of origin while those positively selected on skill remained in the host country. This also had the effect of inflating the observed convergence in wages between migrants and natives (see also Borjas 2014). While Borjas' seminal paper focused on the post-World War II immigrant cohorts, the economic history literature on 19th and early 20th century immigration drew on the findings and debates in contemporary economics literature. Minns (2000) deals with cohort quality by tracking arrival cohorts across the 1900 and 1910 censuses in lieu of individually connected data, however this method could not account for bias due to return migration. Abramitzky et al. (2014) use individually linked census panels, which allow them to control for both cohort quality and return migration and they find no evidence of convergence in occupational scores for immigrants from different European sending countries. Instead, they find positive earnings gaps for immigrants from richer countries in Europe and negative gaps for immigrants from poorer countries.

According to Abramitzky et al. (2014), the average long-term immigrant in the panel data held similarly paying occupations to the average native on arrival in the USA in the period 1900 to 1920. However, the average migrant masks a great deal of heterogeneity across countries of origin. Economic assimilation for any group of immigrants is determined by human capital, which may be further determined by conditions in the home country alongside the migrant selection patterns. Abramitzky et al. (2014) report occupational score premiums between $2000 and $3000 (in 2010 US$) for immigrants from England, Scotland, and Russia, and occupational score penalties of up to $4000 for immigrants from Norway and Portugal, with limited changes in the values of premiums and penalties even after 30 years spent in the USA. Both positive and negative gaps persist over time, which is evidence against the theory of wage convergence. Even though there is some convergence in immigrant-native earnings across generations, the gaps persist also in the second generation, whether they are premiums or penalties (Abramitzky et al. 2014; Borjas 1993). Furthermore, the analysis of linked father-son pairs has shown that the second generation is more upwardly mobile than the native-born with native parents (Abramitzky et al. 2021). The second-generation premium in mobility can be explained by the location choice of immigrant families, which are more likely to settle in areas with better job opportunities.

Premiums that are specific to the country of origin may be explained by the level of economic development in the source country vis-à-vis the USA, by cultural closeness, and by the migrant selectivity which all affect the human capital endowment of the migrant population. Alongside differences in skills across origin countries, there may also be differences in average skills across cohorts from the same country of origin, which is determined by the changing nature of selection driven by changing push and pull factors over time. For example, there may be stronger positive selection in the early periods of chain migration, where the first "pioneers" of a migration chain are more motivated and more willing to bear costs and risks. Complementarities between the skills of immigrants and the skill demanded in the

host country also matter, and structural changes within the economy of the host country can determine the degree of wage assimilation of immigrants. Despite the importance of skills in theory, Ward (2020) estimates a low return to English fluency on occupational income. A low proportion of immigrants in the 1900s was fluent in English on arrival, but most became fluent within their first five years in the USA. It may be that English fluency does not explain a substantial proportion of skill variation across immigrant groups, which could explain its low importance for occupational trajectories.

Some research attempts to analyze the discrimination in the hiring and remuneration of immigrant workers, and recent work has shown that structural barriers faced by immigrants are an important factor in the labor market outcomes of immigrants (Escamilla-Guerrero et al. 2021; Ferrara and Fishback 2020). Another line of research shows that immigrants mitigate discrimination by de-emphasizing their foreign roots by using less foreign-sounding names (Goldstein and Stecklov 2016). The historical literature also highlights the efforts undertaken by non-European immigrants to gain certain rights such as the right to US citizenship, including immigrants from Syria-Lebanon and those from Turkey (Gualtieri 2009). The Ottoman immigrants likely faced some of the same structural barriers that affected their counterparts from other countries such as Mexico (Escamilla-Guerrero et al. 2021). American antipathy toward the Ottoman Empire following World War I and the Armenian genocide may have played a role in the discrimination against this immigrant group (Naff 1985), while cultural and linguistic distance between the immigrants and their host communities may also have had an effect.

Since Ottoman immigrants are drawn from a region that is comparatively less developed than the US or the European sending countries, we expect to find penalties in occupational scores and in earnings when comparing Ottoman immigrants to natives and to European immigrants. Historical evidence points to low levels of urbanization and overall lower levels of education as compared to the industrialized nations, which suggests that the immigrants may assimilate into low paid occupations in the USA due to their low skill levels in the context of the US labor market. However, educational levels varied across communities, and there is historical evidence to suggest that Christians were more highly educated as compared to the Muslim majority in Ottoman Syria (Sassmannshausen 2019). Immigrants from Ottoman Syria and Lebanon were more likely to come from urban areas and were drawn from Christian communities, and both facts suggest a degree of positive selection on educational level as compared to the general population. Migrants from Ottoman Turkey may exhibit different patterns—while positive selection also applies, the migration from Turkey to the USA was part of a larger refugee movement that took place in the context of World War I, and positive selection was probably weaker. In both cases, the positive selection is expected to lead to a higher degree of occupational and wage assimilation.

4. Data and methods

We use data from the United States Decennial Census of the Population. The census includes information on every individual living in the USA at the time of its collection (Ruggles et al. 2021). The censuses include a rich array of variables, such as the individual's gender, age, nationality, year of arrival, occupation, and income. We also make use of the linked census data (Helgertz et al. 2020) which allow us to follow individuals across time if they appear in more than one census. We focus on the period 1900 to 1940. The basic variables such as gender, age, and nationality are included in every census. However, some variables of interest,

such as income and mother tongue only appear in some censuses (see Online Appendix table A1).

We are interested in measuring gaps in income between immigrants and natives. Most of the literature on economic assimilation in the 19th and early 20th centuries use occupational income scores in lieu of actual earnings data. The US Census Bureau only began to collect income as a part of the population census in 1940. However, data on occupation go back as far as the 19th century, which allows the imputation of earnings based on individual occupation, taking the information on median earnings conditional on occupation from the 1950 census (Sobek 1995). While occupational income scores do not reflect actual earnings, they are useful as a measure of occupational standing at a granular level. Recently, however, it has been demonstrated that it is quite problematic to use occupational scores when measuring earnings gaps or income mobility (Inwood et al. 2019; Saavedra and Twinam 2020; Ward 2023). Therefore, we employ three different measures of earnings: Occupational scores (Sobek 1995), LIDO scores (Saavedra and Twinam 2020), and actual earnings. The occupational score is imputed based on the individual's occupation and median earnings in the 1950 census by occupation, while the LIDO scores improve on the occupational scores by imputing wages based on a richer set of individual characteristics (occupation, industry, age, race, and state of residence) and their median corresponding earnings in 1950. Finally, we make use of the self-reported earnings data available in the census of 1940 which reports the pre-tax wage income of employed individuals.

We measure the gaps in occupational score and in actual earnings between the Ottoman immigrants and the native white population who are born in the USA and with both parents born in the USA. We also measure the gaps between the second generation and the natives, with the second generation defined as those born in the USA with one or both parents born in the Ottoman Empire. To compare the trajectories of Ottoman immigrants with those of other immigrant groups, we include immigrants from a selection of European countries with large immigrant communities in the USA in the early 20th century (Sweden, England, Ireland, and Italy). We limit our analytical sample to men aged 18 to 65 years, and exclude a subset of individuals with missing information. We convert the three outcomes into their equivalents in 2010 US dollars (see Abramitzky et al. 2014).

To compare earnings gaps across time and for demographically dissimilar groups we employ regression analysis. In our main analysis, we use the three earnings measures as dependent variables in the same regression specification, and we adjust for the age of individuals, number of years spent in the USA, state of residence. We only include individuals that appear in the linked panel which allows us to measure the occupational gap across time while accounting for return migration, which has been shown to inflate convergence trends in cross-sectional data (see Abramitzky et al. 2014). In the model, we control for the country of birth, and alternatively for mother tongue, with an interaction with time spent in the USA (or second-generation status) as our explanatory variable of interest.

We focus on the labor market outcomes of men. While immigrant men from the Ottoman Empire had similar employment rates as US natives, most immigrant women were coded as not in the labor force, with only 15% coded as employed in 1930. Furthermore, our analysis of the data shows that immigrant men outnumbered women two-to-one amongst early cohorts, and this differential was more pronounced if we were to account for return migrants.

We examine the occupational scores of Ottoman migrants compared to natives using a regression model. We only include individuals who are identified in the linked census panel, which ensures that we control for return migration (see Abramitzky et al. 2014). The base

model takes the following form:

$$\text{INCOME}_{it} = \alpha + \sum \gamma_{jm}\, Origin_j\, Time_m + \sum \delta_t\, Year_t + \sum \delta_s\, State_s + \beta_1\, age_{\text{it}} + \beta_2\, age_{\text{it}}^2 + \beta_3\, age_{\text{it}}^3 + \varepsilon_{it}$$

where i denotes the individual, j is the country of origin with US as the reference category, t is the census year with 1900 as the reference category, and m indicates time spent in the USA. $Origin_j$ includes a set of dummy variables indicating each origin country, and $Time_m$ includes a set of dummy variables that indicate time spent in the USA (with an additional dummy variable indicating second-generation status). Together the $Origin_j$ and $Time_m$ dummies separate the foreign-born into categories according to both their country of origin and the number of years spent in the USA (0 to 5 years, 6 to 10 years, 11 to 20 years, 21 years or more, or second generation). We follow Abramitzky et al. (2014) in modeling the outcomes in levels using 2010 US dollars, and we provide additional results using logged outcomes in the Online Appendix.

In the model using actual earnings in 1940 we limit the sample to those who appear in both the 1930 and 1940 censuses. This allows us to contrast results from three outcome variables using both the full sample running from 1900 to 1940 and the restricted 1940 cross-sectional sample, and we highlight the findings suggested by each. Additionally, we estimate all the models with control variables for occupation and industry categories to investigate the sources of the gaps in each outcome variable. These controls capture the broad occupation and industry categories that are shown in table 3. The purpose of these additional controls is to investigate whether gaps in income are driven by selection into occupations and industries with high average rewards, or whether they also hold within occupation and industry groups.

In addition to the above models, we estimate an individual fixed-effects regression model to examine changes in income in absolute terms by time spent in the USA. The model takes a similar form, with the exclusion of native individuals and the use of newly arrived immigrants as the reference category to track changes within the occupational scores of immigrants across time. Finally, we estimate separate models for each occupational group in order to investigate heterogeneities in earnings.

Our samples include all natives who were aged between 18 and 65 years and were identified in at least two censuses, alongside the Ottoman immigrant men in the same age bracket as well as their European counterparts from selected countries. We exclude individuals with missing information. To construct the panel data, we make use of the crosswalk data published by IPUMS (Helgertz et al. 2020) to link individuals in each pair of censuses, including the panels 1900–1910, 1910–1920, 1920–1930, and 1930–1940.

5. Descriptive statistics

Table 3 shows the descriptive statistics for the full sample and the 1940 sample, broken down by immigrant status and country of origin. Ottoman immigrants held significantly higher occupational scores as compared to natives and as compared to European immigrants. All immigrant groups held LIDO score premiums as compared to natives, with similar patterns found in both the full sample and the 1940 sample. However, the earnings figures based on

Table 3. *Descriptive statistics for the analytical sample.*

	Syria-Lebanon		Turkey		Sweden		England		Ireland		Italy		USA
Full sample (1900 to 1940)	Immi-grants	Gen 2	Immi-grants	Gen 2	Immi-grants	Gen 2	Immi-grants	Gen 2	Immi-grants	Gen 2	Immi-grants	Gen 2	
Occupational score	27,937	26,147	25,524	26,100	21,346	21,118	24,425	23,519	23,751	24,433	22,585	23,874	20,591
LIDO score	23,735	22,761	23,781	24,810	21,615	21,617	23,717	23,135	23,137	23,624	21,696	22,517	19,557
Occupation (%)													
Clerical and Kindred	1	7	2	7	2	7	5	7	5	9	1	9	5
Craftsmen	9	12	18	14	33	19	29	21	23	22	21	20	15
Farm Laborers	1	2	1	3	3	10	2	4	1	3	2	2	8
Farmer (Owner, Tenant, Manager)	3	2	4	8	26	25	9	21	5	13	3	3	35
Laborers	11	9	8	5	10	7	7	6	21	8	27	13	7
Managers, Officials, and Proprietors	42	24	26	19	7	9	11	12	8	11	12	10	9
Operatives	17	18	18	14	13	11	23	15	22	17	22	25	10
Professional, Technical	1	3	4	10	2	4	5	5	2	4	1	4	4
Sales workers	12	17	8	15	2	6	5	7	4	6	3	8	5
Service Workers	4	4	11	4	3	2	4	3	10	6	8	6	2
Industry (%)													
Agriculture	5	4	6	12	30	37	12	26	9	17	7	8	44
Finance, Insurance, and Real Estate	2	3	2	5	1	3	3	3	3	3	1	3	2
Manufacturing	27	26	31	25	26	20	35	24	27	25	30	29	16
Mining and Construction	5	6	5	6	18	10	18	13	13	12	21	14	10
Other Services	8	13	16	20	7	9	12	11	15	14	13	15	9
Transportation and Utilities	3	7	3	5	8	9	9	11	22	16	10	11	10
Wholesale and Retail Trade	50	41	38	28	9	12	11	12	11	12	18	20	10
Mean age	42	31	41	33	45	34	44	40	45	40	42	30	39

(Continued)

Table 3. *Continued*

	Syria-Lebanon		Turkey		Sweden		England		Ireland		Italy		USA
Full sample (1900 to 1940)	Immi-grants	Gen 2	Immi-grants	Gen 2	Immi-grants	Gen 2	Immi-grants	Gen 2	Immi-grants	Gen 2	Immi-grants	Gen 2	
Years in the USA (%) 0 to 5	2		3		2		4		4		3		
6 to 10	9		12		7		8		7		11		
11 to 20	31		39		26		26		26		33		
21 or more	58		46		64		62		63		53		
Year of observation (%) 1900	0.3	0.0	2.0	0.8	14.0	4.9	14.7	13.5	17.8	16.3	2.6	1.5	14.3
1910	4	1	6	3	25	16	24	24	28	27	11	7	27
1920	25	9	15	13	27	35	24	32	24	30	24	24	32
1930	42	47	43	46	22	36	23	28	19	23	36	43	24
1940	29	44	34	38	12	8	14	3	11	3	26	24	3
Observations	29,468	6764	17,863	1774	527, 007	349, 300	693, 083	872, 329	670, 834	1,617, 507	1,264, 983	177, 152	24,074, 792
1940 sample													
Occupational score	27,471	26,913	25,436	26,862	23,047	21,410	25,222	24,348	24,255	24,560	22,660	23,669	21,409
LIDO score	23,921	23,503	23,809	25,196	23,231	21,989	25,224	23,933	24,362	24,095	22,095	22,374	20,443
Earnings	9749	12,307	12,724	15,681	17,157	13,487	21,244	17,636	21,706	18,199	12,954	12,925	12,322
Non-wage income over $50 (%)	49	35	43	35	37	34	27	25	21	16	32	17	33
Education (%)													
N/A or no schooling	16	2	13	1	1	0	1	0	1	0	17	1	1
Nursery school to grade 4	18	3	16	2	4	1	4	2	6	1	31	2	5
Grade 5, 6, 7, or 8	44	37	44	28	77	49	57	43	69	42	41	55	44
Grade 9. 10, 11, or 12	15	44	17	44	14	39	28	41	17	44	6	32	36
College (any)	4	13	7	23	4	9	9	12	4	11	2	8	12
Observations	8465	3800	6253	862	64,220	31,306	96,519	29,580	73,788	53,297	340,249	56,593	698, 848

the 1940 sample show earnings penalties for Syrian immigrants as compared to natives, but earnings appear to converge to native earnings in the second generation. In contrast, northern European immigrants held earnings premiums which declined in the second generation. Comparing the full male population aged 18 to 65 in 1940 to our 1940 sample shows that earnings were similar across the two samples for all immigrant groups—confirming that the sample is representative of the larger immigrant population.

Looking at occupations and industries, we find that the largest group of Ottoman immigrants worked in wholesale and retail trade sectors (50% of Syria-Lebanon immigrants and 38% of Turkey immigrants—mainly in food and merchandise stores and in restaurants). This contrasts with the native population and with most European immigrants where only around 10% of men of worked in this industry. We also find that the proportion of Ottoman workers in trade declines by around 10 percentage points in the second generation, showing a tendency toward diversification. The second largest group of Ottoman workers is in manufacturing at 29%, not far from the proportion employed in manufacturing among the European immigrant groups but around 10 percentage points higher than in the native population.

A large proportion of immigrants worked as managers or proprietors, and this proportion is significantly higher than that of natives (9%) and is highest for immigrants from Syria-Lebanon (42%). Very few of the immigrants worked in farming as compared to natives, which is typical of many immigrant groups that congregated in urban areas.

The educational distribution of each origin group is shown for the 1940 sample. Immigrants from the Ottoman Empire show lower educational attainment than US natives and Northern European immigrants. However, their educational distribution is similar to that of Italian immigrants. We find a clear tendency toward higher educational attainment in the second generation in all immigrant groups, with the Ottoman second generation overtaking the average educational attainment of natives.

Figure 3 shows the age-earnings profiles for Ottoman immigrants by arrival cohort based on LIDO scores. Overall, the three cohort groups show similar trajectories, with only a slight difference found between early and later cohorts from Turkey. The premiums in LIDO and occupational scores are stable across age (see Appendix fig. A1) in both the full 1900–1940 sample and the 1940 sample. On the other hand, the earnings penalty is consistent after age 40 (see Appendix fig. A2). While immigrants at earlier ages appear to match or out-earn their native counterparts, most of our 1940 sample is in the older age group with 95% of immigrants from Syria-Lebanon aged 35 to 60 years and with a mean age of 49 years.

6. Estimation results

6.1 Gaps in occupational scores between immigrants and natives

Figure 4A displays the regression estimates of the model measuring gaps in occupational scores using the native population as the reference category, based on the panel spanning the censuses from 1900 to 1940. Return migration is adjusted for through the panel design (see Online Appendix, table A2 for full regression estimates). We additionally estimate the same regressions using logged outcomes (Online Appendix, table A3) which show similar results.

Ottoman immigrants have higher occupational scores than natives already upon arrival. The premiums are especially stark for immigrants from Syria-Lebanon measuring at $4700 upon arrival and increasing to $5900 after 20 years in the USA. The size of the premium

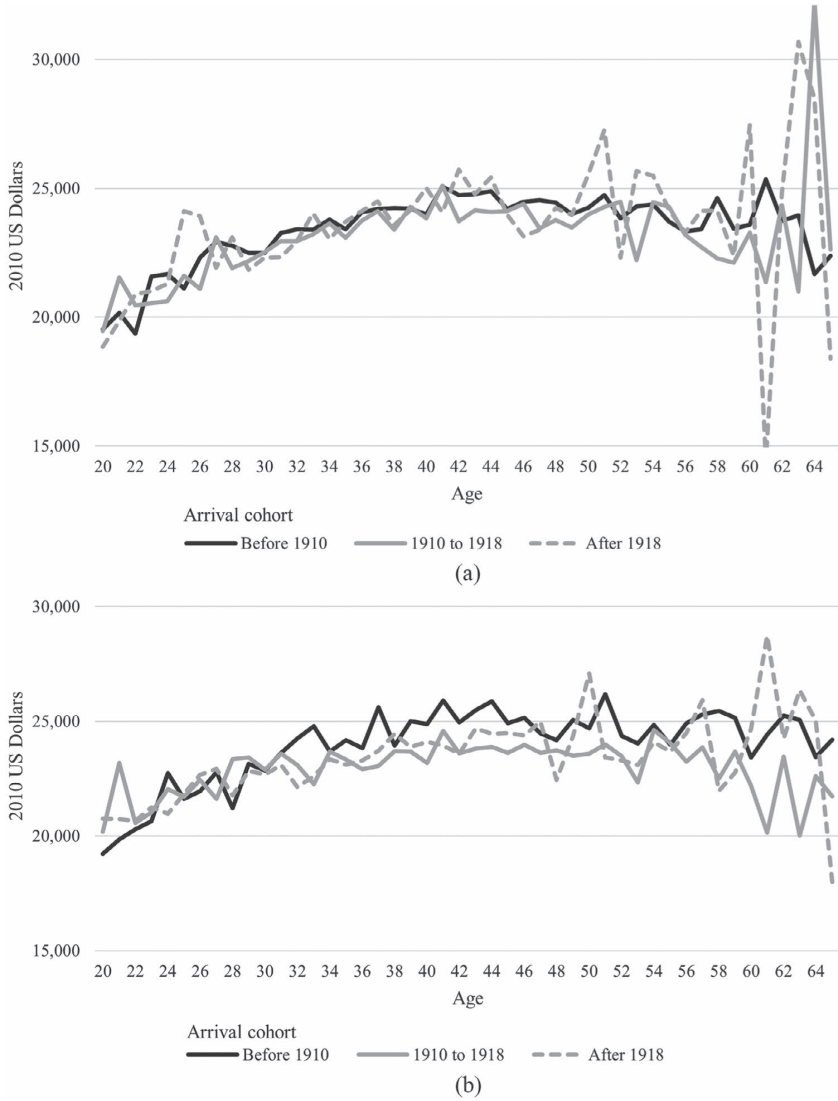


Figure 3. Age-income profiles using LIDO scores for Ottoman immigrants split by year of arrival and region of origin. (a) Syria-Lebanon. (b) Turkey.

is smaller for immigrants from Turkey, with new arrivals having an occupational score that is not significantly different from those of natives. However, their premium grows to over $1000 after 5 years in the USA and reaches over $2000 after 20 years in the country. The European immigrants show different patterns depending on the country of origin, with the

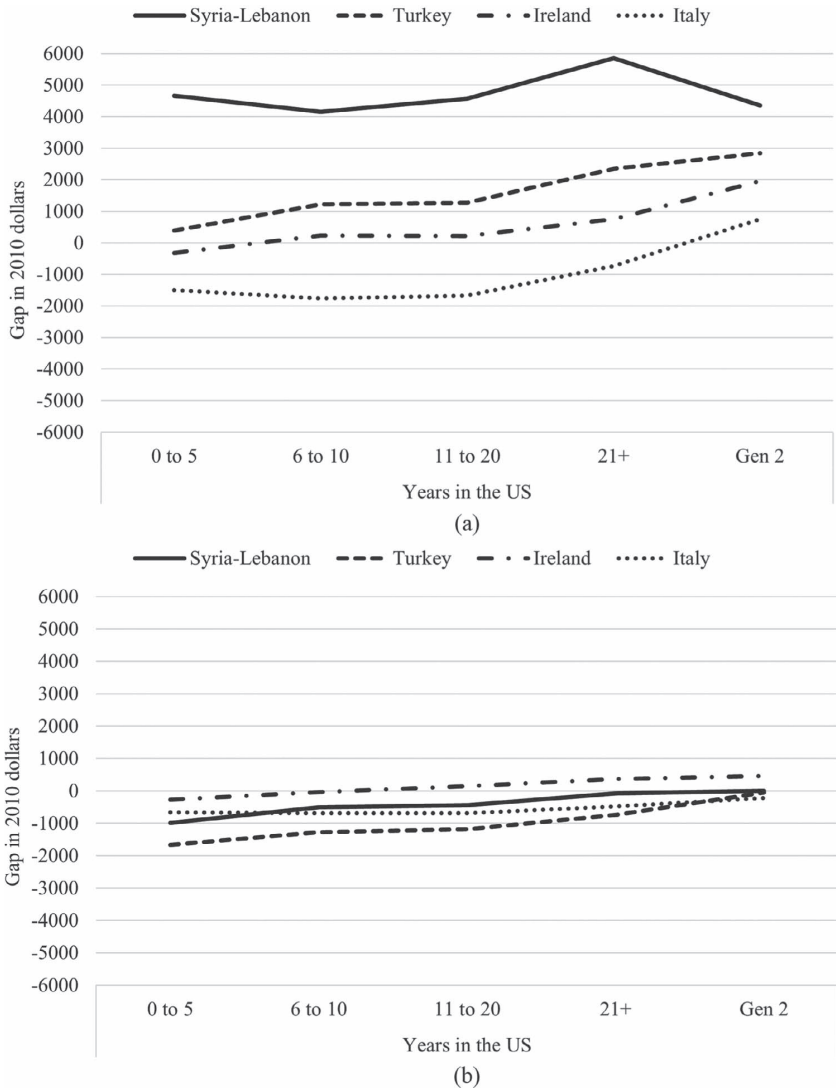


Figure 4. Estimated gaps in occupational scores between immigrants and natives based on the OLS regression models. The models control for age and state of residence. All gaps over $500 are significant at the 0.05 level. Full regression estimates are available in Online Appendix, table A2. (a) Unadjusted. (b) Controlling for occupation and industry groups.

Irish immigrants having no significant gaps in occupational scores on arrival but developing a small premium in occupational score which measures at $800 after 20 years in the USA, while the Italian immigrants arrive with penalties in occupational scores measuring at $1500 which decline to around $700 after 20 years in the USA. The premiums are largely maintained or increased in the second generation, measuring at $4350 for immigrants from Syria-Lebanon, and increasing to $2800 and $2000 for immigrants from Turkey and Ireland. While the penalty held by the Italian immigrants becomes a premium in the second generation measuring at $800.

While the value of the occupational score premium (or penalty) changes with time spent in the USA, the size of the change is relatively small and seldom surpasses $1000 even after 20 years in the USA. This suggests that immigrants experience similar career developments as natives, which is in line with earlier findings on the economic assimilation of European immigrants in the USA (Abramitzky et al. 2014). For the second generation, occupational score premiums indicate upward mobility in the case of Turkey, and inherited premiums in the case of Syria-Lebanon.

Figure 4B shows the occupational score gaps after controlling for occupation and industry groups. The premiums enjoyed by immigrants from Syria-Lebanon are reduced to small penalties that begin at $1000 on arrival but disappear after 20 years in the USA, with the second generation showing no gap with the native occupational scores. A similar pattern can be seen for the other immigrant groups as well. As previously shown the Ottoman immigrants were much more likely to enter into occupation groups with high average rewards as compared to their US-native counterparts, such as managers and proprietors, and sales workers. As these occupational groups are more highly paid on average as compared to other occupational groups, the occupational scores of immigrants will be higher than those of their native counterparts.

Figure 5A shows the gaps in LIDO scores between immigrants and natives. The results from the LIDO score regression give a similar picture, albeit with smaller premiums (or larger penalties) in all cases. Immigrants from Syria-Lebanon appear to arrive in the USA with premiums that are closer to $2000, with the second generation maintaining a premium of $1600. Overall, the premiums using the conventional occupational score are more than twice as large as the premiums using the LIDO scores. The occupational score premiums may be driven by strong selection into service-sector and white-collar occupations among Syrian immigrants. The LIDO score accounts for selection into industry and may provide a more accurate representation of the gap in earnings between immigrants and natives, conditional on industry as well as occupation. Ottoman immigrants are strongly over-represented in the wholesale and retail trade sector as compared to natives and as compared to other immigrant groups, which may explain the large disparity between the occupational score and LIDO score premiums for this group. Controlling for occupation and industry groups in figure 5B shows that within occupation and industry groups, immigrants tend to hold occupations that are on average slightly lower paid than their native counterparts.

While the gaps in occupational scores and LIDO scores are based on the full sample spanning the censuses from 1900 to 1940, the results from the 1940 sample are in accordance with the full sample results (see Online Appendix table A2). In the 1940 sample, we find significant occupational score and LIDO score premiums for Ottoman immigrants, which grow in the second generation. And these premiums largely disappear when we control for occupation and industry groups. Next, we turn to the estimates of actual earnings gaps based on the 1940 sample.

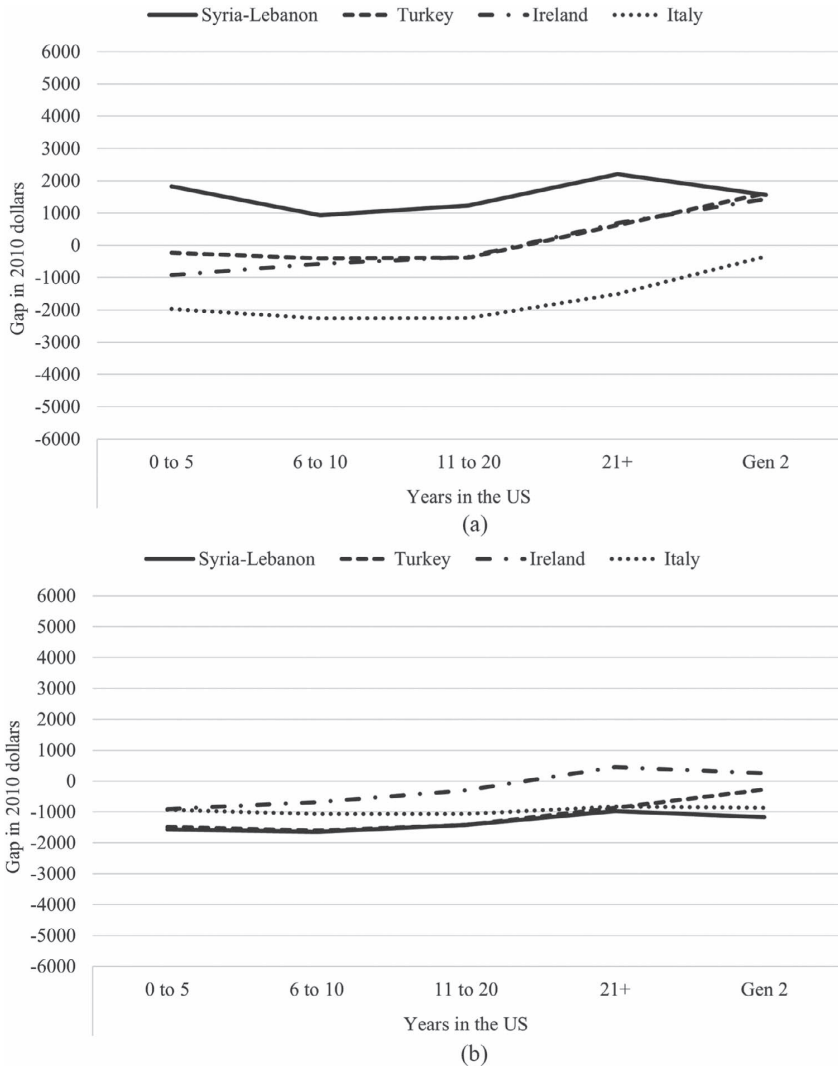


Figure 5. Estimates gaps in LIDO scores between immigrants and natives based on the OLS regression models. All models control for age and state of residence. All gaps over $500 are significant at the 0.05 level. Full regression estimates are available in Online Appendix, table A2. (a) Unadjusted. (b) Controlling for occupation and industry groups.

6.2 *Gaps in actual earnings between immigrants and natives*

Figure 6A shows the gaps in actual earnings between immigrants and natives in 1940. They indicate large penalties for the Ottoman immigrants as compared to the native born. Immigrants from Syria-Lebanon earn on average $3700 less than their native counterparts after 10 years in the USA, and $5500 below natives after more than 20 years in the country. For immigrants from Turkey the corresponding penalties are about $3400 and $4200, respectively. The growth in these penalties with time spent in the USA suggests that the earnings of immigrants grew at a slower pace as compared to those of natives. These earnings penalties are also present in the case of Italian immigrants (around $3300) and appear to be relatively stable with time spent in the USA. In contrast, immigrants from Ireland earn higher wages compared to natives with a gap of $4200 after 10 years in the USA and growing to $6000 after 20 years.

The second generation converges toward the native average earnings, closing most of the gap in the case of immigrants from Syria-Lebanon with a remaining penalty of $860 and achieving earnings parity with natives in the case of immigrants from Turkey. However, there is less convergence in the case of second-generation Italian immigrants where a penalty of $2000 is found. Second-generation Irish immigrants appear to maintain a smaller premium that is around half of that held by the first generation at $2900.

Controlling for occupation and industry groups in figure 6B does not alter the picture significantly. Here, we find slightly larger penalties for Ottoman immigrants, but smaller penalties for Italian immigrants and smaller premiums for Irish immigrants. It appears that the second generation converges toward the earnings of natives in this case as well, though a penalty of $2600 remains for the second-generation immigrants from Syria-Lebanon and $900 for the Turkey second generation.

The gaps estimated using actual earnings data contrast sharply with those estimated from occupational scores and LIDO scores. The earnings scores present a picture of a high degree of labor market assimilation among Ottoman immigrants, where immigrants from Syria-Lebanon out-earn their US native counterparts and those from Turkey match native earnings on arrival and out-grow them with time spent in the USA, and where the second generation appears to inherit the premiums that their fathers held. The actual earnings data, on the other hand, indicate that immigrants arrive in the USA at a significant disadvantage and experience slower earnings growth as compared to natives. Meanwhile, the second generation achieves a remarkable convergence toward the earnings of natives and spurns the large earnings penalties that their fathers endured.

To further test the robustness of our imputed earnings results we estimate individual fixed-effects regressions using the LIDO score as the dependent variable. The results reveal little growth in LIDO scores with time spent in the USA for Ottoman and Italian immigrants, with the Irish immigrants showing larger premiums due to time spent in the USA as compared to new arrivals (Online Appendix, table A4). Within-occupation and industry growth is small for all immigrant groups, with change in LIDO scores below $500 in almost all cases.

6.3 *Comparing earnings within occupation categories*

To explore the drivers behind the earnings penalties, we estimate a separate model for each occupational category. Additionally, we control for the presence of non-wage earnings by including a dummy variable which takes the value one if the individual earns non-wage income over $50. This is meant to account for the fact that immigrants were more likely to

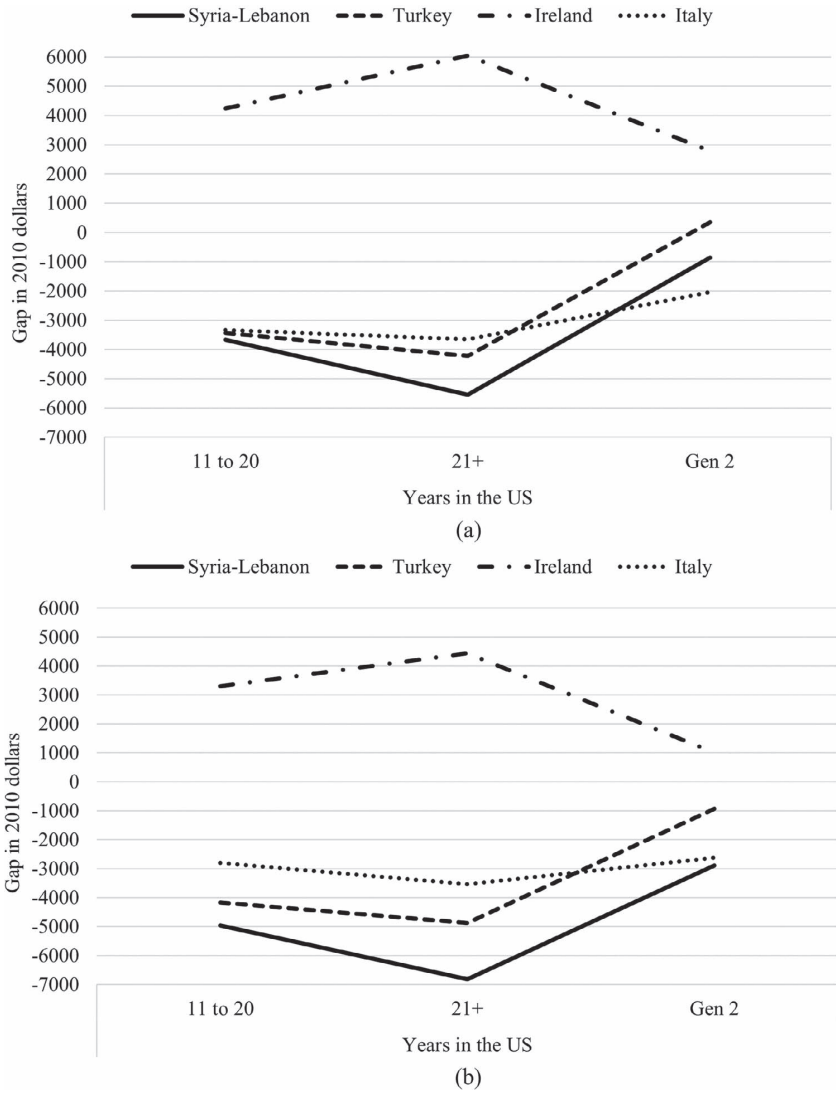


Figure 6. Estimated gaps in earnings between immigrants and natives based on the OLS regression models. All models control for age and state of residence. All gaps over $500 are significant at the 0.05 level. Full regression estimates are available in Online Appendix, table A2. (a) Unadjusted. (b) Controlling for occupation and industry groups.

be self-employed and therefore to earn income that is not captured in the earnings variable. Figure 7 shows the earnings gaps between immigrants and US natives within occupation groups (for full model estimates, see Online Appendix, table A5). The negative gaps are largest for the most common occupational categories amongst the Ottoman immigrants, namely managers/proprietors and sales workers. There are also negative earnings gaps for other less common occupations, while the laborers have income on par with, or even higher than, natives. The penalties within occupation groups are similar across Ottoman and Italian immigrants reaching over $10,000 in the case of managers/proprietors and between $5000 and $7000 for sales workers, whereas Northern European immigrants exhibit earnings premiums in most cases, or earnings parity with natives as in the case of managers and proprietors.

While overall earnings penalties are largest for immigrants from Syria-Lebanon, this immigrant group shows a striking similarity in within-occupation penalties to immigrants from Turkey and Italy, suggesting that the occupational selection plays an important role in shaping the overall penalties and the earnings trajectories with time spent in the USA. Ottoman and Italian immigrants working as operatives, craftsmen, and laborers achieved a higher degree of labor market assimilation, matching the earnings of natives or surpassing them in some cases. Northern European workers likewise achieved their highest premiums in these industrial occupations. Selection into industrial (secondary sector) occupations may explain the premiums for Northern European immigrants, while the high degree of selection of Ottoman immigrants into service-sector occupations, such as wholesale and retail trade, may have contributed to their earnings penalties. The descriptive statistics in table 3 show that the second generation of Ottoman immigrants saw a large reduction in the proportion working in the wholesale and retail trade amounting to 10 percentage points, but without a significant increase in the proportion of workers in industrial occupations. The second generation appears to enter management and proprietorship roles at a much lower rate compared to the immigrant generation, which could point to a process of assimilation into native firms and away from migrant-owned businesses. The largest increase in occupation types among Ottoman and Italian workers in the second generation compared to the first is due to work in clerical occupations.

6.4 The role of educational attainment

Table 4 shows a Gelbach decomposition of the gap in actual earnings between immigrants and natives. The unadjusted model is the same basic model we outline above with origin dummy variables, age, and state fixed effects, while the adjusted model adds a control variable for years of education. In accordance with the Gelbach decomposition method, we estimated a set of auxiliary variables where the dependent variable of interest (years of education) is regressed on a dummy variable for origin, and the contribution of years of education is then calculated as the origin coefficient multiplied by the coefficient on education in the adjusted model (see Gelbach 2016).

We find that years of education account for a substantial part of the gap in earnings for Ottoman immigrants, with the earnings penalty reduced significantly when controlling for years of education. The contribution of years of education to explaining the gap in earnings varies from 45.6% of the unadjusted gap for newly arrived immigrants from Syria-Lebanon, to 38% for those with 21 or more years in the USA. The contribution of education is similar in the case of Turkey immigrants and reaches 43.8% for those with 21 or more years in the USA. These results show that human capital played a crucial role

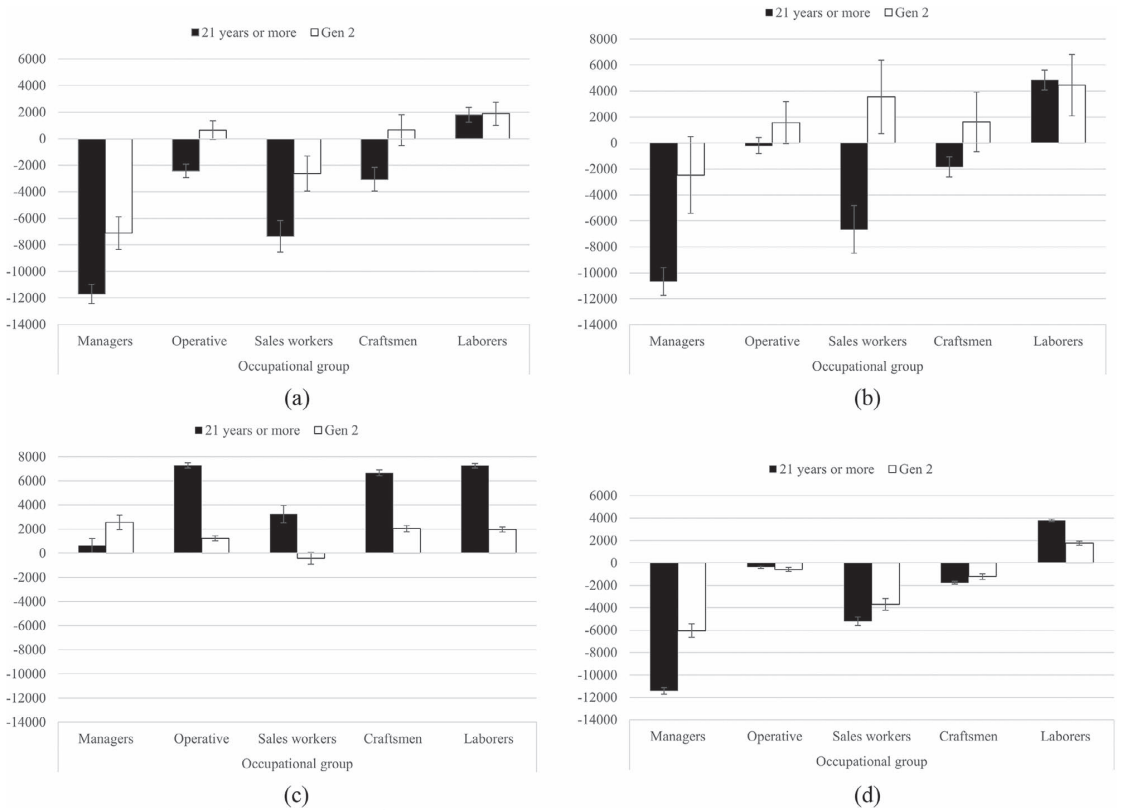


Figure 7. Regression estimates of earnings gaps for Ottoman immigrants by occupation group (2010 US$). The estimates show gaps in earnings based on five models that compare native and immigrant workers within the same occupational groups. All models control for age and state of residence. 95% confidence intervals are shown. Full model estimates are available in Online Appendix, table A5. (a) Syria-Lebanon. (b) Turkey. (c) Ireland. (d) Italy.

Table 4. *Gelbach decomposition of the earnings gap according to years of education.*

	Syria-Lebanon			**Turkey**		
	10 to 20 years	21 or more	Gen 2	10 to 20 years	21 or more	Gen 2
Unadjusted model	−3750	−5644	−854	−3409	−4252	187
	(462.8)	(163.0)	(229.0)	(407.5)	(201.5)	(479.3)
Adjusted model	−2213	−3734	−924	−2264	−2718	−1058
	(449.6)	(158.5)	(218.2)	(388.6)	(192.7)	(456.8)
Education contribution	−1709	−2143	−59	−1427	−1861	890
	(123.0)	(44.1)	(59.7)	(108.1)	(53.7)	(127.3)
% of unadjusted gap	45.6%	38.0%		41.9%	43.8%	
	Ireland			**Italy**		
	10 to 20 years	21 or more	Gen 2	10 to 20 years	21 or more	Gen 2
Unadjusted model	4280	6073	2807	−3333	−3650	−2039
	(111.1)	(65.5)	(63.4)	(66.7)	(35.4)	(62.2)
Adjusted model	5828	7133	2746	−317	−794	−719
	(107.5)	(65.0)	(62.7)	(66.5)	(38.8)	(61.6)
Education contribution	−1841	−1345	−239	−3271	−3017	−1505
	(29.7)	(17.9)	(16.8)	(19.8)	(12.4)	(16.8)
% of unadjusted gap	−43.0%	−22.2%		98.1%	82.7%	

Note: The table presents the gaps in earnings between immigrants and US-natives and the contribution of years of education to explaining these gaps. The base model controls for age and state fixed effects, while the adjusted model a control for years of education. The contribution of education to explaining earnings gaps is calculated according to Gelbach (2016). Standard errors are in parentheses. *Source:* Authors' calculations based on US census data.

in driving the observed differences in assimilation across immigrant origin groups, even without accounting for unmeasured aspects of human capital (for example, differences in schooling quality across countries). The picture is even starker for Italian immigrants, where adjusting for education accounts for 98.1% of the earnings gap between immigrants with 10 to 20 years in the USA and those of natives. Meanwhile, controlling for education increases the premium earned by Irish immigrants by 43% for immigrants with 10 to 20 years in the USA and declines to 22.2% for those with 21 years or more in the USA, suggesting that they out-earned native individuals with similar educational attainment by a significant margin.

The educational attainment of the second generation clearly contrasts with that of their fathers. A substantial proportion of Ottoman immigrants attended school to grade 4 or less (34%, 29%, and 48% for Syria-Lebanon, Turkey, and Italy respectively) while only a small proportion of the second generation shows such low levels of education (5%, 3%, and 3% respectively). This large intergenerational upgrading in human capital likely underlies the occupational upgrading seen in the previous section where the second generation appears

to take on more skilled occupations as compared to the immigrant generation, leading to assimilation in earnings as compared to US natives.

6.5 Heterogeneity and sensitivity analysis

We examine the heterogeneity in the Ottoman immigrant population by using the mother tongue instead of the country-of-origin variables. This breaks down the Ottoman immigrant sample into seven groups (see table 2) and we estimate the gap in each outcome variable comparing each group of immigrants with the outcomes of natives using our main model (Online Appendix, table A6). Since the data on mother tongue was only collected in 1920 and 1930, our sample includes individuals identified in the panels 1910–1920, 1920–1930, and 1930–1940. As shown in table 2, Arabic was the predominant language among immigrants from Syria-Lebanon. However, immigrants from Ottoman Turkey spoke a variety of languages with Armenian and Greek being the most common.

The results for Arabic speakers reflect those of immigrants from Syria-Lebanon, as the two samples are roughly the same. In the population that migrated from Turkey, we find some differences in the size of the occupational score premium, which reflects different occupation selection patterns across groups. The LIDO score penalties are small and only reach around $1000 in most cases. The actual earnings penalty estimated using the 1940 sample appear very similar across groups, reaching between $3500 and $4800 in most cases. Overall, the group-level results are similar to our main results, with small differences between groups and little evidence of important heterogeneities. The same can be said about the results controlling for occupation and industry groups. We do not find different assimilation trajectories for refugee populations (Armenians and Greeks) as compared to other Ottoman immigrants. Both occupational score gaps and LIDO score gaps are stable for Armenian immigrants, while occupational score premiums for Greek immigrant grow with time spent in the USA, but their LIDO score gaps appear to be stable. These results are in line with those of the Arabic- and Turkish-speaking immigrants.

We test the robustness of the main results by using a sub-sample which excludes individuals with non-wage earnings (see Online Appendix, table A7). Since our main results rely on the use of earnings data, misreported earnings could bias our estimates. It may be that immigrant workers are more likely to be self-employed as compared to natives and make a substantial portion of their incomes through business profits, which would make it appear as if they have much lower earnings as compared to natives. In this sub-sample, we find that the occupational score and LIDO score premiums are much lower as compared to our main sample, pointing to a lower degree of occupational selection into management occupations in the service sector as compared with the main sample. However, the earnings regressions tell a similar story as our main analysis: earnings penalties are significant and grow with time spent in the USA, but they are closed in the second generation, and we find earnings premiums of $900 and $3000 for second-generation immigrants from Syria-Lebanon and Turkey, respectively, in the sub-sample. We cannot rule out that the group of immigrants in this sub-sample is differently selected from the overall sample, and the occupational score and LIDO score results may point to this fact, although the results are in line with the main findings.

Finally, we exclude farmers and farm workers from our samples (see Online Appendix, table A8). In this case, the occupational score premiums held by the immigrants from Syria-Lebanon are reduced but remain at around $1800 after 21 years in the USA in the 1940 sample. Yet these premiums are in sharp contrast to the actual earnings penalties, which measure at $8600 after 21 years in the USA in the same sample. Overall, the exclusion of

the farmers and farm laborers shifts down our estimate of the occupational score and LIDO score premiums, but concurrently increases the actual earnings gaps.

7. Conclusion

We present new evidence on the economic assimilation of an early immigrant wave from the Middle East to the United States, which has not been considered previously in the research on the transatlantic migration of the early 20th century. The analysis shows that Ottoman immigrants in the period 1900 to 1940 had markedly higher occupational scores than comparable US natives. The premiums were higher for immigrants from Syria-Lebanon than for immigrants from Turkey. There was no sign of either gain or loss in occupational scores over time for immigrants from Syria-Lebanon, but we found that immigrants from Turkey experienced growth in their premiums with time spent in the USA. Compared to earlier results on the occupational scores premiums of immigrants from Europe, these estimates suggest a favorable status of immigrants from the Ottoman Empire as compared to their European counterparts in the first half of the 20th century.

We went beyond the current literature by evaluating the labor market assimilation of immigrants using actual earnings data for immigrants present in the 1940 census. These findings showed that Ottoman immigrants had lower earnings and did not experience assimilation even after several decades in the USA. There was a clear divergence in earnings, meaning that the earnings of immigrants grew more slowly than those of US natives. In contrast, we found that second-generation immigrants close most of the gap in earnings relative to natives. Immigrants working in certain occupations such as managers/proprietors and sales workers had much lower earnings than comparable natives, while those working as laborers out-earned native laborers. For managers/proprietors, earnings differentials were large, with a gap that reached $11,000, which should be compared to the average earnings in this occupation group of around $22,000.

These findings suggest that Ottoman immigrants selected themselves into relatively high-earning occupational groups, such as managers/proprietors and sales personnel, but they held low-paying occupations within these broader groups and had low earnings both compared to the average native in the labor market, and especially in comparison with natives with similar occupations. This pattern was especially pronounced for immigrants from Syria-Lebanon but was also clear for immigrants from Turkey. Ottoman immigrants were at a disadvantage when arriving in the USA and their situation did not improve overall with more years spent in the US labor market. Self-employment was the main explanation for the low earnings, as our robustness analysis showed similar patterns when individuals earning at least $50 in non-wage income were excluded.

The second generation achieved earnings parity with natives, as in the case of immigrants from Turkey, or at least closed most of the gap as in the case of immigrants from Syria-Lebanon. Despite this, we found evidence of residual earnings penalties when we compared individuals within the same occupation and industry groups. When exploring the mechanism behind the assimilation process, we found that education played an important role in explaining the earnings gaps for Ottoman immigrants (about one third or more of the gap), while the second generation entirely closed the education gap with natives. The second generation appears to have diversified their occupations away from proprietorship in the wholesale and retail trade sector which was strongly favored by the first generation. Overall, the labor market assimilation of the first- and second-generation Ottoman immigrants contrasted with that of

Northern European immigrants, who held occupations in the industrial sector at a much higher rate and appear to have arrived in the USA with an educational attainment profile that resembled that of natives.

As in the previous literature, we found that the degree of assimilation of immigrants depended on the country of origin. However, the imputed earnings measures typically used in the literature do not give an accurate picture of actual earnings, overestimating the earnings of Ottoman immigrants while strongly underestimating that of Northern European immigrants. Since imputed earnings underestimate the penalties faced by Ottoman immigrant groups, they mask the remarkable process of assimilation that occurs in the second generation.

Acknowledgements

We are grateful to Matt Nelson and Jonas Helgertz for assistance in acquiring the data, and to Björn Eriksson and seminar participants at the European Historical Economics Society conference in Groningen (2022), the IUSSP International Seminar on Migration in the Past (2022), and the Economic History seminar at Lund University for valuable comments and suggestions. Rami Zalfou's research was funded by the Centre for Advanced Middle Eastern Studies at Lund University.

Data availability

The data underlying this article was provided by IPUMS USA by permission. Data will be shared on request to the corresponding author with permission of IPUMS USA.

Supplementary material

Supplementary material is available at *European Review of Economic History* online.

References

Abraham, S.Y. and N. Abraham. 1983. *Arabs in the New World: Studies on Arab-American Communities*. Detroit, MI: Wayne State University, Center for Urban Studies.

Abramitzky, R. and L.P. Boustan. 2017. "Immigration in American Economic History". *Journal of Economic Literature* **55**, pp. 1311–45.

Abramitzky, R., L.P. Boustan, J. Elisa, and P. Santiago. 2021. "Intergenerational Mobility of Immigrants in the United States Over Two Centuries". *American Economic Review* **111**, pp. 580–608.

Abramitzky, R., L.P. Boustan, and K. Eriksson. 2014. "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration". *Journal of Political Economy* **122**, pp. 467–506.

Abramitzky, R., L.P. Boustan, K. Eriksson, and S. Hao. 2020. "Discrimination & the Returns to Cultural Assimilation in the Age of Mass Migration". *AEA Pap Proc* **110**, pp. 340–6.

Akarli, E. 1993. *The Long Peace: Ottoman Lebanon, 1861–1920*. Berkeley: University of California Press.

Algan, Y., C. Dustmann, A. Glitz, and A. Manning. 2010. "The Economic Situation of First and Second-Generation Immigrants in France, Germany and the United Kingdom". *The Economic Journal* **120** (542), pp. 4–30.

Baycar, K. 2020. "Between the Sublime Porte and Uncle Sam: The Ottoman-Syrian Immigration to the United States, 1880–1914". *American Nineteenth Century History* **21**, pp. 1–23.

Borjas, G.J. 1985. "Assimilation, Changes in Cohort Quality, & the Earnings of Immigrants". *Journal of Labor Economics* **3**, pp. 463–89.

Borjas, G.J. 1993. "The Intergenerational Mobility of Immigrants". *Journal of Labor Economics* **11**, pp. 113–35.
Borjas, G.J. 2000. The Economic Progress of Immigrants. In G.J. Borjas (ed), *Issues in the Economics of Immigration*. Chicago: University of Chicago Press, pp. 15–49.
Borjas, G.J. 2014. *Immigration Economics*. Cambridge, MA: Harvard University Press.
Brand, T., Caesar Affair in: 1914-1918-online. International Encyclopedia of the First World War, ed. by Ute Daniel, Peter Gatrell, Oliver Janz, Heather Jones, Jennifer Keene, Alan Kramer, and Bill Nasson, issued by Freie Universität Berlin, Berlin 2015-02-23. https://doi.org/10.15463/ie1418.10560.
Chiswick, B.R. 1978. "The Effect of Americanization on the Earnings of Foreign-born Men". *Journal of Political Economy*, **86** (5), pp. 897-921.
Connor, D.S. 2020. "Class Background, Reception Context, & Intergenerational Mobility: a Record Linkage and Surname Analysis of the Children of Irish Immigrants". *International Migration Review* **54**, pp. 4–34.
Escamilla-Guerrero, D., E. Kosack, and Z. Ward. 2021. "Life After Crossing the Border: Assimilation During the First Mexican Mass Migration". *Explorations in Economic History* **82**, 101403.
Fawaz, L.T. 1983. *Merchants & Migrants in Nineteenth-Century Beirut*. Cambridge, MA: Harvard University Press.
Ferrara, A. and P. Fishback. 2020. *Discrimination, Migration, and Economic Outcomes: Evidence from World War I*. NBER Working Paper No. 26936. Cambridge: MA.
Fittante, D. 2017. "But why Glendale? A History of Armenian Immigration to Southern California". *California History* **94**, pp. 2–19.
Gelbach, J.B. 2016. "When Do Covariates Matter? And Which Ones, and How Much? *Journal of Labor Economics* **34**, pp. 509–43.
Goldstein, J.R. and G. Stecklov. 2016. "From Patrick to John F. Ethnic Names & Occupational Success in the Last Era of Mass Migration". *American Sociological Review* **81**, pp. 85–106.
Gratien, C. and E.K. Pope-Obeda. 2020. "The Second Exchange: Ottoman Greeks and the American Deportation State During the 1930s". *Journal of Migration History* **6**, pp. 104–28.
Grogger, J. and H.H. Gordon. 2011. "Income Maximization & the Selection & Sorting of International Migrants". *Journal of Development Economics* **95**, pp. 42–57.
Gualtieri, S. 2001. "Becoming "White": Race, Religion and the Foundations of Syrian/Lebanese Ethnicity in the United States". *Journal of American Ethnic History* **20**, pp. 29–58.
Gualtieri, S. 2009. *Between Arab & White: Race & Ethnicity in the Early Syrian American Diaspora*. Berkeley & Los Angeles: University of California Press.
Hashimoto, K. 1992. Lebanese Population Movement 1920–1939. In A. Hourani, and N. Shehadi (eds), *Lebanese in the World: A Century of Emigration*. London: I.B. Tauris, pp. 77–107.
Hatton, T.J. 1997. "The Immigrant Assimilation Puzzle in Late Nineteenth-Centuty America". *The Journal of Economic History* **57**(1), pp. 34-62.
Helgertz, J., S. Ruggles, J.R. Warren, C.A. Fitch, R. Goeken, J.D. Hacker, M.A. Nelson, J.P. Price, E. Roberts, and M. Sobek. 2020. *IPUMS Multigenerational Longitudinal Panel: Version 1.0 [dataset]*. Minneapolis, MN: IPUMS.
Inwood, K., C. Minns, and F. Summerfield. 2019. "Occupational Income Scores & Immigrant Assimilation. Evidence from the Canadian Census". *Explorations in Economic History* **72**, pp. 114–22.
Issawi, C. 1988. *The Fertile Crescent, 1800–1914: A Documentary Economic History*. Oxford: Oxford University Press.
Karpat, K.H. 1985a. *Ottoman Population. Demographic & Social Characteristic*. Madison, WI: University of Wisconsin Press.
Karpat, K.H. 1985b. "The Ottoman Emigration to America, 1860–1914". *International Journal of Middle East Studies* **17**, pp. 175–209.
Khater, A.F. 2001. *Inventing Home: Emigration, Gender, & the Middle Class in Lebanon, 1870–1920*. Berkeley, CA: University of California Press.
Lindner, C.B. 2009. *Negotiating the Field: American Protestant Missionaries in Ottoman Syria, 1823 to 1860*. Doctoral Thesis, The University of Edinburgh. http://hdl.handle.net/1842/3687.

Minns, C. 2000. "Income, Cohort Effects, and Occupational Mobility: A New Look at Immigration to the United States at the Turn of the 20th century". *Explorations in economic history*, **37**(4), pp.326-350.
Naar, D.E. 2015. "Turkinos Beyond the Empire: Ottoman Jews in America, 1893 to 1924". *Jewish Quarterly Review* **105**, pp. 174–205.
Naff, A. 1985. *Becoming American: The Early Arab Immigrant Experience*. Carbondale & Edwardsville: Southern Illinois University Press.
Ruggles, S., C.A. Fitch, R. Goeken, J.D. Hacker, M.A. Nelson, E. Roberts, M. Schouweiler, and M. Sobek. 2021. *IPUMS Ancestry Full Count Data: Version 3.0 [dataset]*. Minneapolis, MN: IPUMS.
Saavedra, M. and T. Twinam. 2020. "A Machine Learning Approach to Improving Occupational Income Scores". *Explorations in Economic History* **75**, 101304.
Saliba, N.E. 1981. "Emigration from Syria". *Arab Studies Quarterly* **3**, pp. 56–67.
Sassmannshausen, C. 2019. "Educated with Distinction". *Journal of the Economic and Social History of the Orient* **62**, pp. 222–56.
Sobek, M. 1995. "The Comparability of Occupations & the Generation of Income Scores". *Hist Methods* **28**, pp. 47–51.
United Nations. 1969. *Growth of the World's Urban & Rural Population, 1920–2000*. New York: United Nations.
Ward, Z. 2020. "The Low Return to English Fluency During the Age of Mass Migration". *European Review of Economic History* **24**, pp. 219–42.
Ward, Z. 2023. "Intergenerational Mobility in American History: Accounting For Race & Measurement error". *The American Economic Review* **113**, pp. 3213–48.

Appendix

Table A1: Variable availability by census for selected variables

Census year	Occupation	Wage income	Year of arrival	Mother tongue
1900	x	-	x	-
1910	x	-	x	-
1920	x	-	x	x
1930	x	-	x	x
1940	x	x	-	-

Table A2: Regression estimates of gaps in occupational scores, LIDO scores, and earnings between immigrants and US natives.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent variable	Occscore	Occscore	Occscore	Occscore	Lido	Lido	Lido	Lido	Earnings	Earnings	Earnings	Earnings
Sample	1900 - 1940	1900 - 1940	1940	1940	1900 - 1940	1900 - 1940	1940	1940	1940	1940	1940†	1940†
Syria-Lebanon												
0 to 5 years	4663.4***	-990.1***	-	-	1830.2***	-1563.7***	-	-	-	-	-	-
	(497.5)	(234.9)			(390.8)	(238.3)						
6 to 10 years	4149.6***	-510.1***	-	-	932.4***	-1650.5***	-	-	-	-	-	-
	(192.5)	(90.92)			(151.2)	(92.20)						
11 to 20 years	4569.9***	-441.6***	4551.3***	-645.8***	1236.4***	-1417.0***	1585.0***	-1391.3***	-3668.1***	-4956.1***	-373.2	-1242.0*
	(102.3)	(48.32)	(292.4)	(149.3)	(80.34)	(49.00)	(233.9)	(150.3)	(454.5)	(413.0)	(563.4)	(495.6)
21 or more	5857.0***	-71.36*	4524.2***	-527.3***	2205.4***	-979.4***	1480.3***	-1286.9***	-5544.3***	-6813.1***	-3380.7***	-3505.3***
	(74.66)	(35.28)	(103.0)	(52.79)	(58.65)	(35.78)	(82.45)	(53.16)	(160.2)	(146.1)	(208.7)	(183.7)
Gen 2	4351.3***	2.264	5188.0***	444.1***	1566.9***	-1171.8***	2524.0***	-560.8***	-862.0***	-2892.9***	1060.1***	-535.0*
	(119.3)	(56.33)	(143.4)	(73.25)	(93.69)	(57.13)	(114.7)	(73.76)	(222.9)	(202.7)	(259.6)	(228.4)
Turkey												
0 to 5 years	390.7	-1668.8***	-	-	-227.6	-1486.1***	-	-	-	-	-	-
	(446.3)	(210.8)			(350.6)	(213.8)						
6 to 10 years	1218.7***	-1278.6***	-	-	-412.8*	-1607.4***	-	-	-	-	-	-
	(211.5)	(99.90)			(166.2)	(101.3)						
11 to 20 years	1269.7***	-1183.3***	1854.9***	-1149.2***	-379.7***	-1428.0***	191.5	-1356.8***	-3433.6***	-4170.5***	-853.2	-1017.1*
	(117.9)	(55.68)	(254.2)	(129.8)	(92.60)	(56.46)	(203.4)	(130.7)	(395.1)	(359.0)	(467.9)	(411.6)
21 or more	2346.3***	-745.3***	1827.4***	-1072.7***	630.3***	-863.1***	108.4	-1311.7***	-4226.5***	-4865.9***	-2187.9***	-1821.5***
	(107.7)	(50.85)	(125.6)	(64.23)	(84.57)	(51.57)	(100.5)	(64.68)	(195.3)	(177.7)	(235.6)	(207.4)
Gen 2	2836.3***	-49.94	4216.9***	718.1***	1616.9***	-270.3*	2624.4***	373.0*	356.0	-936.2*	3661.1***	1564.9***
	(232.7)	(109.9)	(300.3)	(153.3)	(182.8)	(111.5)	(240.3)	(154.4)	(466.8)	(424.1)	(535.4)	(470.9)
Sweden												
0 to 5 years	-196.6*	-862.5***	-	-	-816.4***	-962.3***	-	-	-	-	-	-
	(93.51)	(44.17)			(73.46)	(44.79)						

Table A2 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
6 to 10 years	-48.63	-946.4***	-	-	-572.1***	-899.6***	-	-	-	-	-	-
	(50.72)	(23.97)			(39.84)	(24.31)						
11 to 20 years	-382.3***	-840.1***	262.8**	-1552.8***	-688.7***	-768.0***	-197.8*	-1334.1***	4658.3***	3348.0***	2096.4***	2918.0***
	(26.71)	(12.63)	(98.37)	(50.34)	(20.98)	(12.81)	(78.71)	(50.69)	(152.9)	(139.3)	(155.2)	(136.9)
21 or more	-177.3***	-558.3***	860.5***	-756.9***	-342.8***	-473.4***	546.0***	-643.5***	3193.0***	2096.8***	2388.5***	2455.3***
	(17.41)	(8.236)	(42.87)	(22.03)	(13.68)	(8.352)	(34.30)	(22.19)	(66.65)	(60.96)	(74.09)	(65.67)
Gen 2	-87.03***	-137.4***	-43.24	-245.6***	-197.0***	-251.5***	196.9***	-157.4***	388.6***	136.8	914.4***	352.8***
	(17.14)	(8.095)	(52.71)	(26.91)	(13.46)	(8.209)	(42.17)	(27.10)	(81.93)	(74.46)	(92.92)	(81.77)
England												
0 to 5 years	956.0***	-823.5***	-	-	210.1***	-1041.2***	-	-	-	-	-	-
	(59.87)	(28.28)			(47.03)	(28.68)						
6 to 10 years	1112.2***	-767.5***	-	-	309.7***	-1013.7***	-	-	-	-	-	-
	(41.96)	(19.83)			(32.96)	(20.11)						
11 to 20 years	1533.9***	-569.0***	1502.9***	-701.3***	822.4***	-613.9***	1399.8***	-454.6***	5707.3***	3310.5***	2855.7***	2073.9***
	(23.58)	(11.15)	(61.68)	(31.57)	(18.53)	(11.31)	(49.35)	(31.79)	(95.88)	(87.34)	(96.60)	(85.14)
21 or more	2069.7***	-186.8***	2178.6***	-295.5***	1676.2***	113.5***	2127.4***	179.9***	5783.5***	3470.4***	4256.0***	3117.9***
	(15.23)	(7.207)	(37.04)	(19.00)	(11.96)	(7.309)	(29.64)	(19.14)	(57.58)	(52.58)	(60.92)	(53.75)
Gen 2	1596.7***	154.7***	1781.7***	134.4***	1192.2***	122.4***	1503.1***	120.2***	2832.8***	1353.9***	2446.3***	1474.8***
	(10.84)	(5.124)	(52.92)	(27.03)	(8.514)	(5.196)	(42.34)	(27.22)	(82.26)	(74.79)	(87.13)	(76.67)
Ireland												
0 to 5 years	-321.5***	-275.4***	-	-	-929.2***	-912.1***	-	-	-	-	-	-
	(65.60)	(31.00)			(51.54)	(31.43)						
6 to 10 years	228.4***	-36.42	-	-	-570.0***	-687.4***	-	-	-	-	-	-
	(46.65)	(22.05)			(36.65)	(22.36)						
11 to 20 years	214.9***	142.8***	-283.3***	-215.6***	-358.6***	-298.0***	-187.8***	-235.5***	4232.2***	3306.1***	931.8***	2097.7***
	(23.94)	(11.33)	(70.24)	(35.96)	(18.81)	(11.49)	(56.20)	(36.21)	(109.2)	(99.49)	(107.8)	(95.05)
21 or more	756.9***	364.2***	1093.4***	774.6***	686.7***	448.3***	1224.2***	798.4***	6050.9***	4444.4***	3182.0***	3758.8***
	(15.40)	(7.306)	(42.43)	(21.83)	(12.09)	(7.409)	(33.95)	(21.98)	(65.95)	(60.40)	(68.15)	(60.34)

Table A2 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Gen 2	1961.8***	455.7***	1765.5***	585.0***	1436.4***	255.5***	1598.8***	362.2***	2786.3***	1018.7***	2349.7***	1438.8***
	(8.205)	(3.891)	(41.09)	(21.05)	(6.446)	(3.946)	(32.88)	(21.20)	(63.87)	(58.24)	(66.82)	(58.96)
Italy												
0 to 5 years	-1495.0***	-668.3***	-	-	-1968.0***	-932.1***	-	-	-	-	-	-
	(51.16)	(24.19)			(40.19)	(24.53)						
6 to 10 years	-1761.6***	-685.7***	-	-	-2266.0***	-1063.1***	-	-	-	-	-	-
	(26.83)	(12.71)			(21.08)	(12.89)						
11 to 20 years	-1675.0***	-688.0***	-1495.9***	-1098.4***	-2254.7***	-1066.6***	-1888.6***	-1088.9***	-3329.1***	-2799.2***	-5078.6***	-2015.2***
	(15.68)	(7.457)	(42.86)	(22.07)	(12.32)	(7.562)	(34.29)	(22.23)	(66.63)	(61.07)	(70.38)	(62.48)
21 or more	-731.8***	-479.3***	-880.9***	-811.6***	-1501.1***	-826.7***	-1451.0***	-980.9***	-3649.1***	-3537.9***	-5130.0***	-2484.2***
	(12.96)	(6.167)	(24.58)	(12.83)	(10.18)	(6.254)	(19.67)	(12.92)	(38.22)	(35.49)	(41.89)	(37.61)
Gen 2	751.8***	-216.9***	1190.3***	55.56**	-339.7***	-868.2***	178.8***	-517.0***	-2032.2***	-2625.0***	-1720.9***	-1183.0***
	(23.71)	(11.21)	(40.35)	(20.65)	(18.62)	(11.36)	(32.28)	(20.80)	(62.72)	(57.14)	(67.38)	(59.40)
Year dummies	Y	Y	-	-	Y	Y	-	-	-	-	-	-
Age controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Occupation & Industry	-	Y	-	Y	-	Y	-	Y	-	Y	-	Y
State	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
_cons	1439.1***	13756.8***	3859.9***	10790.1***	3432.0***	-4170.0***	5026.6***	7683.5***	-18871.3***	1439.1***	13756.8***	-31224.3***
	(71.73)	(37.16)	(581.1)	(297.2)	(56.35)	(37.69)	(464.9)	(299.3)	(903.3)	(71.73)	(37.16)	(879.3)
N	30302856	30302856	1463780	1463780	30302856	30302856	1463780	1463780	1463780	30302856	30302856	1082869
R^2	0.076	0.794	0.053	0.753	0.207	0.705	0.152	0.650	0.097	0.076	0.794	0.304
adj. R^2	0.076	0.794	0.053	0.753	0.207	0.705	0.152	0.650	0.097	0.076	0.794	0.304

†: Columns 11 and 12 exclude individuals with zero earnings.
Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A3: Regression estimates of gaps in occupational scores, LIDO scores, and earnings between immigrants and US natives (using log outcomes)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dependent variable	Occscore	Occscore	Occscore	Occscore	Lido	Lido	Lido	Lido	Earnings	Earnings	Earnings	Earnings
Sample	1900 - 1940	1900 - 1940	1940	1940	1900 - 1940	1900 - 1940	1940	1940	1940	1940	1940†	1940†
Syria-Lebanon												
0 to 5 years	0.229***	-0.0683***	-	-	0.130***	-0.0687***	-	-	-	-	-	-
	(0.0223)	(0.00841)			(0.0191)	(0.0106)						
6 to 10 years	0.210***	-0.0317***	-	-	0.0872***	-0.0706***	-	-	-	-	-	-
	(0.00865)	(0.00326)			(0.00742)	(0.00412)						
11 to 20 years	0.218***	-0.0228***	0.219***	-0.0297***	0.0991***	-0.0525***	0.111***	-0.0468***	-1.451***	-1.427***	0.0482	-0.0184
	(0.00460)	(0.00173)	(0.0148)	(0.00669)	(0.00394)	(0.00219)	(0.0118)	(0.00676)	(0.127)	(0.108)	(0.0359)	(0.0317)
21 or more	0.271***	-0.00228	0.208***	-0.0267***	0.149***	-0.0224***	0.108***	-0.0360***	-1.669***	-1.740***	-0.103***	-0.136***
	(0.00336)	(0.00126)	(0.00522)	(0.00237)	(0.00288)	(0.00160)	(0.00417)	(0.00239)	(0.0446)	(0.0381)	(0.0133)	(0.0118)
Gen 2	0.232***	0.00310	0.263***	0.0163***	0.114***	-0.0494***	0.155***	-0.0155***	-0.651***	-0.876***	0.150***	0.0305*
	(0.00534)	(0.00201)	(0.00726)	(0.00328)	(0.00458)	(0.00254)	(0.00579)	(0.00332)	(0.0620)	(0.0528)	(0.0166)	(0.0146)
Turkey												
0 to 5 years	0.0591**	-0.0896***	-	-	0.0217	-0.0852***	-	-	-	-	-	-
	(0.0200)	(0.00754)			(0.0172)	(0.00953)						
6 to 10 years	0.0799***	-0.0658***	-	-	0.00647	-0.0862***	-	-	-	-	-	-
	(0.00951)	(0.00358)			(0.00815)	(0.00452)						
11 to 20 years	0.0708***	-0.0579***	0.0895***	-0.0616***	0.00432	-0.0708***	0.0378***	-0.0542***	-1.137***	-1.253***	0.0575	0.0113
	(0.00530)	(0.00199)	(0.0129)	(0.00582)	(0.00454)	(0.00252)	(0.0103)	(0.00587)	(0.110)	(0.0935)	(0.0298)	(0.0263)
21 or more	0.116***	-0.0348***	0.0828***	-0.0576***	0.0568***	-0.0364***	0.0349***	-0.0475***	-1.133***	-1.305***	-0.00387	-0.0204
	(0.00483)	(0.00182)	(0.00636)	(0.00288)	(0.00414)	(0.00230)	(0.00508)	(0.00291)	(0.0544)	(0.0463)	(0.0150)	(0.0133)
Gen 2	0.133***	-0.00199	0.188***	0.0211**	0.0784***	-0.0239***	0.127***	0.00773	-0.754***	-0.607***	0.214***	0.0876**
	(0.0103)	(0.00387)	(0.0152)	(0.00687)	(0.00881)	(0.00489)	(0.0121)	(0.00694)	(0.130)	(0.110)	(0.0341)	(0.0301)
Sweden												
0 to 5 years	0.0556***	-0.0303***	-	-	-0.00904*	-0.0502***	-	-	-	-	-	-
	(0.00420)	(0.00158)			(0.00360)	(0.00200)						
6 to 10 years	0.0642***	-0.0329***	-	-	0.00295	-0.0478***	-	-	-	-	-	-
	(0.00228)	(0.000859)			(0.00196)	(0.00109)						
11 to 20 years	0.0384***	-0.0284***	0.0623***	-0.0658***	-0.00351***	-0.0364***	0.0172***	-0.0596***	1.405***	0.534***	0.273***	0.274***
	(0.00120)	(0.000453)	(0.00498)	(0.00226)	(0.00103)	(0.000573)	(0.00398)	(0.00228)	(0.0426)	(0.0363)	(0.00990)	(0.00876)

Table A3 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
21 or more	0.0370***	-0.0154***	0.0798***	-0.0302***	0.0152***	-0.0146***	0.0502***	-0.0251***	0.804***	0.261***	0.258***	0.213***
	(0.000784)	(0.000296)	(0.00217)	(0.000988)	(0.000672)	(0.000374)	(0.00173)	(0.000997)	(0.0186)	(0.0159)	(0.00472)	(0.00420)
Gen 2	0.00853***	-0.00449***	0.00648*	-0.0144***	-0.00395***	-0.0136***	0.0145***	-0.00745***	-0.0814***	-0.116***	0.0921***	0.0411***
	(0.000762)	(0.000287)	(0.00267)	(0.00121)	(0.000654)	(0.000363)	(0.00213)	(0.00122)	(0.0228)	(0.0194)	(0.00592)	(0.00523)
England												
0 to 5 years	0.103***	-0.0298***	-	-	0.0380***	-0.0546***	-	-	-	-	-	-
	(0.00269)	(0.00101)			(0.00231)	(0.00128)						
6 to 10 years	0.105***	-0.0270***	-	-	0.0390***	-0.0520***	-	-	-	-	-	-
	(0.00189)	(0.000710)			(0.00162)	(0.000898)						
11 to 20 years	0.120***	-0.0179***	0.101***	-0.0316***	0.0617***	-0.0321***	0.0810***	-0.0244***	1.533***	0.609***	0.265***	0.179***
	(0.00106)	(0.000400)	(0.00312)	(0.00142)	(0.000909)	(0.000505)	(0.00249)	(0.00143)	(0.0267)	(0.0228)	(0.00616)	(0.00545)
21 or more	0.134***	-0.00177***	0.122***	-0.0146***	0.102***	0.00615***	0.110***	0.00397***	1.230***	0.449***	0.299***	0.198***
	(0.000685)	(0.000258)	(0.00188)	(0.000852)	(0.000587)	(0.000327)	(0.00150)	(0.000860)	(0.0160)	(0.0137)	(0.00388)	(0.00344)
Gen 2	0.0959***	0.00748***	0.108***	0.00639***	0.0691***	0.00302***	0.0809***	0.00130	0.547***	0.126***	0.167***	0.0892***
	(0.000452)	(0.000170)	(0.00268)	(0.00121)	(0.000388)	(0.000216)	(0.00214)	(0.00122)	(0.0229)	(0.0195)	(0.00555)	(0.00491)
Ireland												
0 to 5 years	0.0525***	-0.00621***	-	-	-0.0144***	-0.0520***	-	-	-	-	-	-
	(0.00295)	(0.00111)			(0.00253)	(0.00140)						
6 to 10 years	0.0718***	0.00398***	-	-	-0.00264	-0.0416***	-	-	-	-	-	-
	(0.00210)	(0.000790)			(0.00180)	(0.000999)						
11 to 20 years	0.0651***	0.0138***	0.0265***	-0.00601***	0.00574***	-0.0191***	0.0109***	-0.00969***	1.541***	0.720***	0.206***	0.235***
	(0.00108)	(0.000406)	(0.00356)	(0.00161)	(0.000923)	(0.000514)	(0.00284)	(0.00163)	(0.0304)	(0.0259)	(0.00687)	(0.00608)
21 or more	0.0837***	0.0239***	0.0767***	0.0357***	0.0628***	0.0264***	0.0726***	0.0380***	1.618***	0.667***	0.294***	0.281***
	(0.000693)	(0.000262)	(0.00215)	(0.000979)	(0.000594)	(0.000332)	(0.00171)	(0.000988)	(0.0184)	(0.0157)	(0.00435)	(0.00386)
Gen 2	0.121***	0.0202***	0.114***	0.0252***	0.0850***	0.00772***	0.0893***	0.0121***	0.527***	-0.0484**	0.179***	0.0940***
	(0.000351)	(0.000133)	(0.00208)	(0.000944)	(0.000301)	(0.000168)	(0.00166)	(0.000953)	(0.0178)	(0.0152)	(0.00426)	(0.00377)
Italy												
0 to 5 years	0.00607**	-0.0237***	-	-	-0.0545***	-0.0491***	-	-	-	-	-	-
	(0.00230)	(0.000867)			(0.00197)	(0.00110)						
6 to 10 years	-0.0169***	-0.0248***	-	-	-0.0719***	-0.0535***	-	-	-	-	-	-
	(0.00121)	(0.000456)			(0.00103)	(0.000576)						
11 to 20 years	-0.0237***	-0.0228***	-0.0252***	-0.0446***	-0.0727***	-0.0468***	-0.0542***	-0.0370***	0.210***	-0.406***	-0.120***	0.0226***
	(0.000706)	(0.000268)	(0.00217)	(0.000989)	(0.000605)	(0.000338)	(0.00173)	(0.000999)	(0.0185)	(0.0159)	(0.00449)	(0.00400)

Table A3 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
21 or more	0.0124***	-0.0108***	-0.00864***	-0.0333***	-0.0319***	-0.0257***	-0.0347***	-0.0285***	-0.0114	-0.660***	-0.138***	-0.0211***
	(0.000584)	(0.000221)	(0.00125)	(0.000575)	(0.000500)	(0.000280)	(0.000994)	(0.000581)	(0.0106)	(0.00925)	(0.00267)	(0.00241)
Gen 2	0.0786***	-0.00585***	0.0916***	-0.0000976	0.0110***	-0.0459***	0.0346***	-0.0226***	-0.302***	-0.677***	-0.0213***	-0.0240***
	(0.00106)	(0.000399)	(0.00204)	(0.000926)	(0.000909)	(0.000505)	(0.00163)	(0.000935)	(0.0175)	(0.0149)	(0.00430)	(0.00380)
Year dummies	Y	Y	-	-	Y	Y	-	-	-	-	-	-
Age controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Occupation & Industry	-	Y	-	Y	-	Y	-	Y	-	Y	-	Y
State	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
_cons	8.726***	9.624***	8.876***	9.394***	8.827***	8.381***	9.046***	9.312***	9.441***	6.727***	6.166***	6.492***
	(0.00327)	(0.00136)	(0.0294)	(0.0133)	(0.00281)	(0.00171)	(0.0235)	(0.0135)	(0.251)	(0.214)	(0.0636)	(0.0563)
N	29385477	29385477	1463780	1463780	29385477	29385477	1463780	1463780	1463780	1463780	1082869	1082869
R^2	0.108	0.874	0.078	0.812	0.307	0.786	0.201	0.739	0.076	0.332	0.101	0.300
adj. R^2	0.108	0.874	0.078	0.812	0.307	0.786	0.201	0.739	0.076	0.332	0.101	0.300

†: Columns 11 and 12 exclude individuals with zero earnings.
Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A4: Regression estimates from individual fixed-effects models.

	Years in the US		
Unadjusted	6 – 10	11 – 20	21 or more
Syria - Lebanon	900.7*	952.0*	773.8
	(417.0)	(389.1)	(395.5)
Turkey	571.3	850.6*	1129.1*
	(484.0)	(423.7)	(457.1)
Sweden	217.2**	400.7***	648.8***
	(79.89)	(72.56)	(74.87)
England	12.80	877.6***	2071.8***
	(62.57)	(51.31)	(54.86)
Ireland	349.0***	745.0***	1604.4***
	(68.63)	(59.05)	(62.22)
Italy	413.9***	352.9***	29.25
	(40.16)	(36.76)	(40.47)
N = 3684922, R2 = 0.049, adj. R2 = 0.049			
Controlling for occupation and industry			
Syria - Lebanon	-331.0	-316.4	-177.4
	(290.6)	(270.2)	(275.6)
Turkey	-344.0	-182.5	67.45
	(319.7)	(280.9)	(308.4)
Sweden	-96.65	-186.5***	-393.4***
	(52.91)	(48.24)	(49.93)
England	-127.9**	267.4***	267.4***
	(42.27)	(34.94)	(34.94)
Ireland	-104.6*	267.8***	709.5***
	(46.10)	(40.04)	(42.28)
Italy	-142.8***	-286.4***	-474.9***
	(27.15)	(24.87)	(27.48)
N = 3684922, R2 = 0.554, adj. R2 = 0.554			

Note: Standard errors in parentheses. Dependent variable is LIDO score. All regressions control for age and state of residence. Reference category is immigrants with 0 to 5 years in the US.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Authors' calculations based on US census data.

Table A5: Earnings gaps in 1940 by occupation group.

	(1) Managers	(2) Operatives	(3) Sales workers	(4) Craftsmen	(5) Laborers
Syria-Lebanon					
10 to 20 years	-9236.4***	1018.7	-4775.6***	-2377.3	3745.2***
	(1075.8)	(743.8)	(1385.9)	(1263.0)	(867.4)
21 or more	-11703.1***	-2443.8***	-7367.2***	-3068.9***	1800.3***
	(375.5)	(259.5)	(603.5)	(456.1)	(283.3)
Gen 2	-7124.1***	629.5	-2638.9***	645.8	1878.1***
	(624.3)	(359.3)	(678.3)	(592.5)	(442.8)
Turkey					
10 to 20 years	-9398.1***	1210.7	-4202.4*	-1680.7*	5121.1***
	(1186.6)	(655.6)	(1641.6)	(733.9)	(922.7)
21 or more	-10675.9***	-216.7	-6664.5***	-1852.4***	4837.5***
	(548.2)	(317.7)	(940.8)	(400.9)	(391.9)
Gen 2	-2469.2	1562.8	3536.8*	1622.1	4449.4***
	(1507.2)	(831.4)	(1447.3)	(1166.6)	(1212.1)
Sweden					
10 to 20 years	-2195.5*	5428.5***	2092.8	4361.7***	7866.9***
	(944.2)	(249.2)	(1375.7)	(198.7)	(269.1)
21 or more	-2810.1***	5300.5***	1761.6***	4577.1***	5932.2***
	(299.3)	(120.3)	(462.1)	(98.02)	(123.7)
Gen 2	-915.0*	1472.2***	427.3	1180.5***	1342.7***
	(405.8)	(145.9)	(368.6)	(165.7)	(152.4)
England					
10 to 20 years	2802.5***	4569.3***	4391.8***	5187.6***	6494.5***
	(428.9)	(144.1)	(487.4)	(150.1)	(204.1)
21 or more	3727.1***	3931.1***	5355.2***	5166.1***	4983.5***
	(222.7)	(93.39)	(253.8)	(97.75)	(117.7)
Gen 2	3407.9***	1517.1***	1875.5***	2408.2***	1742.9***
	(358.5)	(129.9)	(333.3)	(163.9)	(151.7)
Ireland					
10 to 20 years	-771.7	7063.0***	1361.9*	4719.1***	8015.1***
	(597.0)	(157.4)	(675.0)	(206.8)	(144.5)
21 or more	641.9*	7270.3***	3246.1***	6654.2***	7253.6***
	(300.8)	(104.8)	(371.9)	(122.5)	(95.65)
Gen 2	2553.2***	1233.5***	-412.5	2040.3***	1967.9***
	(306.2)	(101.1)	(254.2)	(132.2)	(106.9)
Italy					
10 to 20 years	-10836.4***	133.4	-5556.4***	-2184.2***	4533.9***
	(326.1)	(94.84)	(456.5)	(115.6)	(81.01)
21 or more	-11427.9***	-377.0***	-5201.6***	-1761.6***	3795.7***
	(147.3)	(56.28)	(196.0)	(67.56)	(53.15)
Gen 2	-6056.2***	-583.5***	-3710.9***	-1231.7***	1768.5***
	(311.6)	(84.75)	(266.8)	(131.2)	(93.50)
Age controls	Y	Y	Y	Y	Y
Non-wage income dummy	-18646.8***	-5889.8***	-9304.9***	-10126.0***	-3931.7***
	(108.1)	(49.96)	(133.5)	(55.07)	(46.86)
Constant	-51599.2***	-18565.2***	-36616.8***	-34029.8***	-8577.1***
	(4554.0)	(1474.6)	(4448.8)	(1948.2)	(1412.9)
N	139145	274659	80221	285486	207173
R^2	0.256	0.113	0.113	0.164	0.113
adj. R^2	0.255	0.113	0.113	0.164	0.113

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Source: Authors' calculations based on US census data.

Table A6: Regression estimates of earnings (occupational score and actual earnings) in different models and samples for language groups.

Variable:	(1) Occscore	(2) Occscore	(3) Occscore	(4) Occscore	(5) Lido	(6) Lido	(7) Lido	(8) Lido	(9) Earnings	(10) Earnings
Sample:	1910-1940	1910-1940	1940	1940	1910-1940	1910-1940	1940	1940	1940	1940
Arabic										
0 to 5 years	4329.1***	-1099.9***	-	-	1520.5***	-1799.6***	-	-	-	-
	(533.3)	(244.7)			(423.3)	(250.6)				
6 to 10 years	3934.0***	-436.2***	-	-	656.0***	-1800.4***	-	-	-	-
	(203.4)	(93.30)			(161.4)	(95.55)				
11 to 20 years	4211.5***	-433.1***	4371.6***	-435.7**	968.0***	-1539.7***	1610.1***	-1132.5***	-3453.2***	-4217.8***
	(107.5)	(49.33)	(301.8)	(154.0)	(85.30)	(50.52)	(242.5)	(155.1)	(475.9)	(430.1)
21 or more	5619.6***	-53.80	4422.1***	-207.3***	2071.9***	-1067.3***	1564.4***	-980.2***	-5639.6***	-6332.4***
	(78.41)	(36.00)	(104.8)	(53.60)	(62.23)	(36.87)	(84.20)	(54.00)	(165.2)	(149.8)
Turkish										
0 to 5 years	386.1	-278.0	-	-	-205.0	-428.5	-	-	-	-
	(1164.8)	(534.4)			(924.5)	(547.3)				
6 to 10 years	938.2	-1156.3***	-	-	-421.5	-1449.8***	-	-	-	-
	(561.0)	(257.4)			(445.3)	(263.6)				
11 to 20 years	327.3	-1377.3***	2711.6***	-344.9	-844.0***	-1438.8***	1007.4	-577.1	-3755.5***	-4380.8***
	(318.0)	(145.9)	(654.4)	(333.7)	(252.4)	(149.4)	(525.8)	(336.1)	(1031.9)	(932.3)
21 or more	2281.3***	-762.3***	1722.0***	-709.2***	932.5***	-653.6***	541.1*	-754.7***	-4089.5***	-4482.5***
	(299.8)	(137.5)	(336.9)	(171.8)	(237.9)	(140.9)	(270.7)	(173.1)	(531.2)	(480.0)
Armenian										
0 to 5 years	419.6	-1436.8*	-	-	-769.7	-1679.7**	-	-	-	-
	(1262.4)	(579.2)			(1002.0)	(593.1)				
6 to 10 years	-302.6	-1210.4***	-	-	-1245.6***	-1355.4***	-	-	-	-
	(404.7)	(185.7)			(321.2)	(190.2)				
11 to 20 years	-312.6	-978.4***	1467.1***	-619.8**	-1205.4***	-1153.5***	588.0	-525.4**	-3442.2***	-3810.5***
	(218.5)	(100.3)	(389.1)	(198.4)	(173.4)	(102.7)	(312.6)	(199.9)	(613.5)	(554.3)
21 or more	658.8***	-485.0***	911.6***	-384.4***	-316.4*	-589.4***	-133.4	-641.0***	-4755.6***	-4438.7***
	(184.6)	(84.71)	(214.8)	(109.5)	(146.5)	(86.75)	(172.6)	(110.3)	(338.7)	(306.0)
Greek										
0 to 5 years	72.79	-1488.3***	-	-	-855.8	-1729.9***	-	-	-	-
	(859.2)	(394.2)			(682.0)	(403.7)				
6 to 10 years	1197.1**	-1730.9***	-	-	-1019.7**	-2309.0***	-	-	-	-
	(459.8)	(210.9)			(364.9)	(216.0)				
11 to 20 years	1625.2***	-1507.2***	1763.6**	-1549.6***	-1004.7***	-2209.5***	-590.1	-2024.6***	-3352.8***	-3113.0***
	(238.3)	(109.3)	(599.6)	(305.7)	(189.1)	(112.0)	(481.8)	(308.0)	(945.4)	(854.2)
21 or more	2793.1***	-1150.7***	2181.3***	-1154.3***	289.3	-1614.8***	137.7	-1296.5***	-4663.6***	-4492.1***
	(221.4)	(101.6)	(235.9)	(120.4)	(175.7)	(104.0)	(189.5)	(121.2)	(372.0)	(336.3)

Table A6 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Spanish										
0 to 5 years	779.4	-1539.8**	-	-	-982.7	-2829.1***	-	-	-	-
	(1134.6)	(520.5)			(900.5)	(533.1)				
6 to 10 years	743.6	-1198.4***	-	-	-1028.3**	-2083.7***	-	-	-	-
	(487.3)	(223.5)			(386.7)	(228.9)				
11 to 20 years	1722.3***	-1219.7***	2363.6***	-720.8*	-298.0	-1901.3***	311.2	-1481.3***	-3643.4***	-3674.8***
	(303.3)	(139.2)	(598.3)	(305.1)	(240.7)	(142.5)	(480.8)	(307.3)	(943.5)	(852.4)
21 or more	2650.4***	-1082.8***	2111.7***	-1002.0***	850.3***	-1347.1***	733.8**	-1229.3***	-4833.0***	-5368.4***
	(303.1)	(139.1)	(319.5)	(163.0)	(240.6)	(142.4)	(256.7)	(164.2)	(503.9)	(455.3)
Hebrew or Yiddish										
0 to 5 years	298.2	-4004.7**	-	-	-1108.5	-4164.0***	-	-	-	-
	(2678.0)	(1228.6)			(2125.5)	(1258.2)				
6 to 10 years	334.9	-2187.7***	-	-	-547.1	-1600.2**	-	-	-	-
	(1106.6)	(507.6)			(878.3)	(519.9)				
11 to 20 years	2981.8***	-1413.3***	426.0	-2241.5**	546.5	-1787.3***	-62.62	-2100.6**	-3845.7	-5283.2*
	(631.2)	(289.6)	(1496.3)	(762.9)	(501.0)	(296.6)	(1202.2)	(768.5)	(2359.3)	(2131.5)
21 or more	3992.1***	-707.4**	3574.5***	-138.2	2150.9***	-708.7**	2206.2***	-442.5	-3169.2**	-4726.3***
	(550.0)	(252.3)	(725.3)	(369.8)	(436.5)	(258.4)	(582.7)	(372.5)	(1143.6)	(1033.2)
Other										
0 to 5 years	2280.4	231.9	-	-	580.4	-930.7	-	-	-	-
	(1315.7)	(603.6)			(1044.3)	(618.2)				
6 to 10 years	2692.7***	-645.6*	-	-	824.3	-1216.1***	-	-	-	-
	(662.2)	(303.8)			(525.5)	(311.1)				
11 to 20 years	3061.2***	-746.4***	3708.4***	217.7	652.7**	-1353.4***	1776.4*	-233.3	-1426.7	-2810.4*
	(308.7)	(141.6)	(949.1)	(483.9)	(245.0)	(145.0)	(762.5)	(487.5)	(1496.5)	(1351.9)
21 or more	4214.8***	-77.81	3458.6***	32.61	1946.9***	-451.4***	1531.6***	-353.0*	-1611.2**	-2521.4***
	(251.2)	(115.2)	(342.4)	(174.6)	(199.4)	(118.0)	(275.1)	(175.9)	(539.8)	(487.8)
Year dummies	Y	Y	-	-	Y	Y	-	-	-	-
Age controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Occupation & industry	-	Y	-	Y	-	Y	-	Y	-	Y
State	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Constant	3320.6***	13603.4***	4517.3***	10716.6***	5162.1***	-5543.4***	4190.4***	6934.0***	-25166.2***	-27831.7***
	(89.91)	(50.82)	(577.9)	(295.3)	(71.36)	(52.04)	(464.3)	(297.5)	(911.2)	(825.1)
N	20770079	20770079	1463355	1463355	20770079	20770079	1463355	1463355	1463355	1463355
R^2	0.067	0.804	0.043	0.751	0.204	0.721	0.135	0.647	0.061	0.233
adj. R^2	0.067	0.804	0.043	0.751	0.204	0.721	0.135	0.647	0.061	0.233

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Authors' calculations based on US census data.

Table A7: Sensitivity test – 1940 sample excluding individuals with non-wage earnings.

	(1) Occscore	(2) Occscore	(3) Lido	(4) Lido	(5) Earnings	(6) Earnings
Syria-Lebanon						
10 to 20 years	1196.2***	-622.6***	-323.9	-1592.3***	-1103.2	-2198.1***
	(338.7)	(183.0)	(302.0)	(200.6)	(566.8)	(510.8)
21 or more	562.8***	-624.8***	-454.9***	-1292.6***	-3632.2***	-4335.5***
	(126.0)	(68.17)	(112.4)	(74.70)	(210.9)	(190.2)
Gen 2	2413.0***	279.0***	1487.1***	-293.1**	878.3***	-939.7***
	(153.9)	(83.19)	(137.2)	(91.15)	(257.5)	(232.1)
Turkey						
10 to 20 years	-256.9	-1304.9***	-819.2**	-1450.7***	-46.09	-659.4
	(283.3)	(153.1)	(252.6)	(167.8)	(474.1)	(427.2)
21 or more	-616.8***	-1263.2***	-1000.6***	-1300.9***	-1823.1***	-2039.3***
	(145.3)	(78.59)	(129.6)	(86.11)	(243.2)	(219.3)
Gen 2	2207.4***	530.6**	1920.0***	392.7*	2975.1***	1221.0*
	(321.8)	(173.9)	(286.9)	(190.5)	(538.5)	(485.2)
Sweden						
10 to 20 years	47.95	-1609.6***	-513.7***	-1352.1***	3976.9***	3924.1***
	(96.75)	(52.41)	(86.27)	(57.43)	(161.9)	(146.3)
21 or more	1179.1***	-784.7***	706.6***	-647.2***	4128.8***	3406.3***
	(47.14)	(25.65)	(42.03)	(28.11)	(78.88)	(71.59)
Gen 2	505.6***	-186.9***	546.4***	-175.8***	895.5***	284.3***
	(55.77)	(30.15)	(49.73)	(33.03)	(93.33)	(84.13)
England						
10 to 20 years	1143.0***	-532.7***	1008.2***	-427.8***	4471.2***	3160.1***
	(60.37)	(32.69)	(53.83)	(35.82)	(101.0)	(91.22)
21 or more	1638.6***	-144.9***	1653.9***	131.4***	5199.7***	3704.0***
	(38.35)	(20.80)	(34.20)	(22.79)	(64.18)	(58.04)
Gen 2	1372.7***	135.6***	1282.1***	151.8***	2450.8***	1321.6***
	(52.96)	(28.63)	(47.22)	(31.38)	(88.63)	(79.90)
Ireland						
10 to 20 years	-512.3***	-162.2***	-500.9***	-173.2***	2764.4***	3154.6***
	(66.18)	(35.86)	(59.01)	(39.29)	(110.7)	(100.1)
21 or more	611.8***	819.1***	683.6***	732.4***	4649.4***	4458.3***
	(42.83)	(23.28)	(38.19)	(25.51)	(71.67)	(64.97)
Gen 2	1267.1***	560.0***	1302.0***	435.7***	1847.5***	727.6***
	(39.36)	(21.34)	(35.10)	(23.39)	(65.88)	(59.56)
Italy						
10 to 20 years	-2297.2***	-1166.7***	-2455.5***	-1058.1***	-3517.7***	-1497.3***
	(43.42)	(23.66)	(38.72)	(25.93)	(72.67)	(66.03)
21 or more	-1902.1***	-844.5***	-2187.9***	-1019.7***	-3812.4***	-2202.6***
	(26.00)	(14.30)	(23.18)	(15.67)	(43.51)	(39.90)
Gen 2	259.5***	-107.1***	-169.7***	-358.2***	-2043.1***	-1977.4***
	(38.83)	(21.04)	(34.62)	(23.05)	(64.98)	(58.70)
Age	Y	Y	Y	Y	Y	Y
State	Y	Y	Y	Y	Y	Y
Occupation & industry	-	Y	-	Y	-	Y
_cons	4776.6***	11265.4***	4303.9***	7510.7***	-29160.1***	-25252.1***
	(599.4)	(324.4)	(534.5)	(355.5)	(1003.1)	(905.3)
N	1013877	1013877	1013877	1013877	1013877	1013877
R^2	0.046	0.721	0.132	0.617	0.106	0.274
adj. R^2	0.046	0.721	0.132	0.617	0.106	0.274

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Authors' calculations based on US census data.

Table A8: Regression estimates of gaps in occupational scores, LIDO scores, and earnings between immigrants and US natives excluding farmers and farm workers.

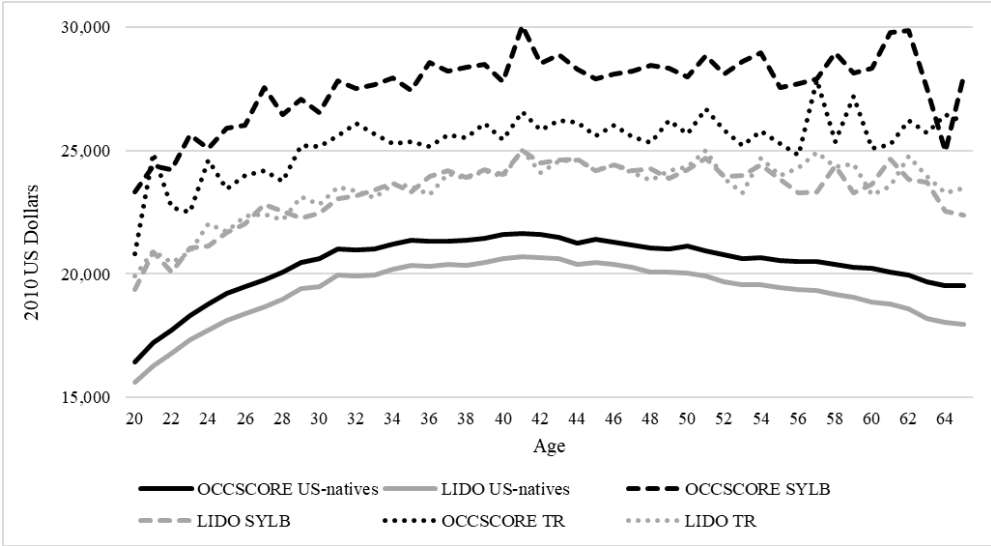
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent variable	Occscore	Occscore	Occscore	Occscore	Lido	Lido	Lido	Lido	Earnings	Earnings
Sample	1900 - 1940	1900 - 1940	1940	1940	1900 - 1940	1900 - 1940	1940	1940	1940	1940
Syria-Lebanon										
0 to 5	291.1	-964.8**	-	-	-976.2*	-1142.6***	-	-	-	-
	(466.2)	(303.7)			(380.8)	(291.5)				
6 to 10	591.8***	-473.9***	-	-	-1491.5***	-1298.4***	-	-	-	-
	(178.8)	(116.5)			(146.0)	(111.8)				
11 to 20	1015.6***	-505.9***	1762.9***	-697.2***	-1439.4***	-1473.5***	-583.5**	-1520.2***	-6542.8***	-5385.1***
	(95.19)	(62.02)	(260.4)	(162.8)	(77.76)	(59.54)	(212.2)	(160.6)	(465.9)	(446.7)
21 or more	2080.9***	-171.4***	1838.4***	-581.3***	-1033.2***	-1535.1***	-808.9***	-1650.3***	-8629.9***	-7362.5***
	(69.94)	(45.58)	(93.14)	(58.39)	(57.13)	(43.76)	(75.91)	(57.59)	(166.7)	(160.2)
Gen 2	1212.1***	20.88	2344.8***	430.5***	-533.0***	-814.8***	442.4***	-496.0***	-3428.8***	-3032.9***
	(111.2)	(72.45)	(129.1)	(80.73)	(90.86)	(69.56)	(105.2)	(79.62)	(231.0)	(221.5)
Turkey										
0 to 5	-2893.2***	-1575.5***	-	-	-2045.1***	-683.7**	-	-	-	-
	(412.9)	(268.9)			(337.3)	(258.2)				
6 to 10	-1467.0***	-1295.0***	-	-	-2119.9***	-1168.4***	-	-	-	-
	(196.5)	(128.0)			(160.5)	(122.9)				
11 to 20	-1206.7***	-1323.7***	-282.7	-1215.8***	-2234.4***	-1491.1***	-1569.1***	-1548.5***	-5783.2***	-4642.8***
	(110.5)	(71.97)	(228.3)	(142.7)	(90.24)	(69.09)	(186.0)	(140.7)	(408.5)	(391.5)
21 or more	-63.44	-899.3***	-288.7*	-1151.1***	-1526.6***	-1302.8***	-1754.5***	-1658.8***	-6792.4***	-5456.5***
	(101.7)	(66.23)	(113.8)	(71.19)	(83.05)	(63.59)	(92.71)	(70.21)	(203.6)	(195.3)
Gen 2	2005.2***	-70.39	3060.5***	750.8***	966.6***	-125.2	1601.8***	335.7	-734.8	-1022.7*
	(226.6)	(147.6)	(278.2)	(173.9)	(185.1)	(141.7)	(226.8)	(171.5)	(497.9)	(477.1)
Sweden										
0 to 5	-3268.0***	-974.2***	-	-	-2378.8***	-338.5***	-	-	-	-
	(94.43)	(61.52)			(77.13)	(59.06)				
6 to 10	-3261.0***	-1151.4***	-	-	-2449.5***	-556.9***	-	-	-	-
	(51.75)	(33.73)			(42.27)	(32.38)				
11 to 20	-3134.9***	-1169.9***	-2044.9***	-1732.0***	-2586.9***	-867.6***	-1954.1***	-1494.4***	2470.0***	3069.5***
	(28.39)	(18.52)	(90.01)	(56.38)	(23.19)	(17.78)	(73.36)	(55.61)	(161.1)	(154.7)
21 or more	-2192.1***	-877.7***	-711.6***	-938.3***	-1990.8***	-807.2***	-763.3***	-882.5***	2174.5***	2275.7***
	(19.29)	(12.59)	(42.03)	(26.44)	(15.76)	(12.09)	(34.25)	(26.08)	(75.20)	(72.56)
Gen 2	-149.3***	-177.9***	-117.0*	-247.2***	-180.2***	-194.0***	93.84*	-219.4***	681.9***	326.6***
	(19.53)	(12.73)	(53.96)	(33.73)	(15.96)	(12.22)	(43.97)	(33.27)	(96.55)	(92.55)

Table A8 (continued)

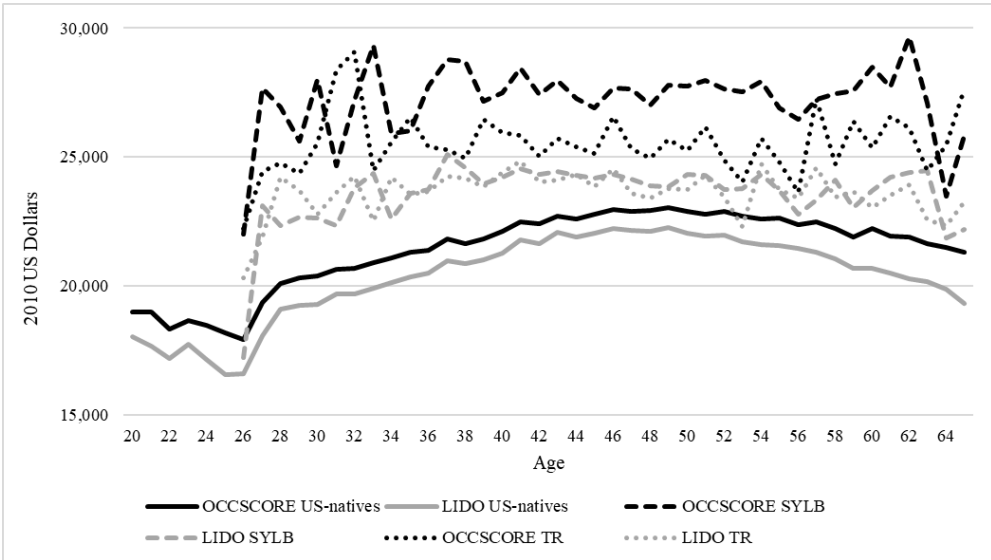
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
England										
0 to 5	-2296.1***	-852.5***	-	-	-1656.7***	-607.9***	-	-	-	-
	(57.02)	(37.15)			(46.58)	(35.66)				
6 to 10	-1999.1***	-823.4***	-	-	-1614.6***	-783.1***	-	-	-	-
	(39.98)	(26.05)			(32.66)	(25.01)				
11 to 20	-1751.4***	-656.8***	-790.5***	-794.3***	-1349.5***	-591.5***	-437.0***	-698.4***	3292.9***	2915.7***
	(22.70)	(14.80)	(55.80)	(34.92)	(18.55)	(14.21)	(45.48)	(34.44)	(99.85)	(95.82)
21 or more	-1002.8***	-289.0***	78.10*	-369.6***	-689.4***	-211.7***	426.0***	-89.65***	3733.6***	3309.2***
	(15.08)	(9.835)	(34.46)	(21.57)	(12.32)	(9.442)	(28.08)	(21.28)	(61.65)	(59.19)
Gen 2	227.7***	200.6***	399.3***	129.1***	233.6***	193.2***	500.1***	138.7***	1727.7***	1383.9***
	(11.49)	(7.485)	(49.66)	(31.04)	(9.384)	(7.186)	(40.47)	(30.61)	(88.85)	(85.15)
Ireland										
0 to 5	-3037.2***	-172.9***	-	-	-2158.6***	-78.35*	-	-	-	-
	(61.73)	(40.22)			(50.42)	(38.61)				
6 to 10	-2703.2***	28.99	-	-	-2016.0***	-28.89	-	-	-	-
	(43.77)	(28.53)			(35.75)	(27.39)				
11 to 20	-2943.7***	130.6***	-2207.7***	-287.0***	-2203.1***	-51.77***	-1671.2***	-361.8***	2086.2***	2827.4***
	(22.66)	(14.79)	(63.39)	(39.71)	(18.51)	(14.20)	(51.66)	(39.17)	(113.4)	(109.0)
21 or more	-2640.8***	330.4***	-955.4***	771.5***	-1923.0***	70.98***	-492.0***	503.2***	3894.3***	4216.0***
	(14.92)	(9.754)	(39.13)	(24.59)	(12.18)	(9.364)	(31.89)	(24.25)	(70.02)	(67.46)
Gen 2	-260.2***	546.8***	54.13	603.6***	-30.86***	392.2***	388.7***	443.5***	1319.0***	960.3***
	(8.327)	(5.433)	(37.73)	(23.62)	(6.802)	(5.216)	(30.75)	(23.30)	(67.51)	(64.82)
Italy										
0 to 5	-4799.7***	-621.7***	-	-	-3678.4***	-218.7***	-	-	-	-
	(48.22)	(31.44)			(39.39)	(30.19)				
6 to 10	-4413.8***	-694.5***	-	-	-3763.4***	-577.7***	-	-	-	-
	(25.27)	(16.51)			(20.65)	(15.85)				
11 to 20	-4212.9***	-787.3***	-3383.8***	-1193.2***	-3959.9***	-1009.4***	-3404.1***	-1275.8***	-5526.6***	-3345.5***
	(14.88)	(9.757)	(39.11)	(24.63)	(12.15)	(9.367)	(31.87)	(24.29)	(69.98)	(67.58)
21 or more	-3294.9***	-615.0***	-2771.2***	-896.4***	-3647.3***	-1255.3***	-3092.9***	-1324.6***	-5964.6***	-4104.9***
	(12.50)	(8.193)	(23.22)	(14.75)	(10.21)	(7.866)	(18.92)	(14.55)	(41.55)	(40.47)
Gen 2	-1001.5***	-187.6***	-480.9***	55.68*	-1310.4***	-391.4***	-994.2***	-374.3***	-3471.0***	-2729.9***
	(22.59)	(14.72)	(36.92)	(23.10)	(18.45)	(14.13)	(30.09)	(22.78)	(66.07)	(63.38)
Year dummies	Y	Y	-	-	Y	Y	-	-	-	-
Occupation & Industry	-	Y	-	Y	-	Y	-	Y	-	Y
State	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Constant	9272.4***	11380.7***	4368.2***	9422.6***	8433.0***	-8532.3***	3317.9***	3246.0***	-28977***	-30662***
	(84.55)	(58.02)	(570.3)	(356.9)	(69.07)	(55.70)	(464.8)	(352.0)	(1020.5)	(979.3)
N	30302856	30302856	1463780	1463780	30302856	30302856	1463780	1463780	1463780	1463780
R^2	0.076	0.794	0.053	0.753	0.207	0.705	0.152	0.650	0.097	0.255
adj. R^2	0.076	0.794	0.053	0.753	0.207	0.705	0.152	0.650	0.097	0.255

Figure A1: Age-income profiles for US-natives and Ottoman immigrants using Occupational Scores and LIDO scores.

A. Full sample (1900 to 1940)



B. 1940 Sample

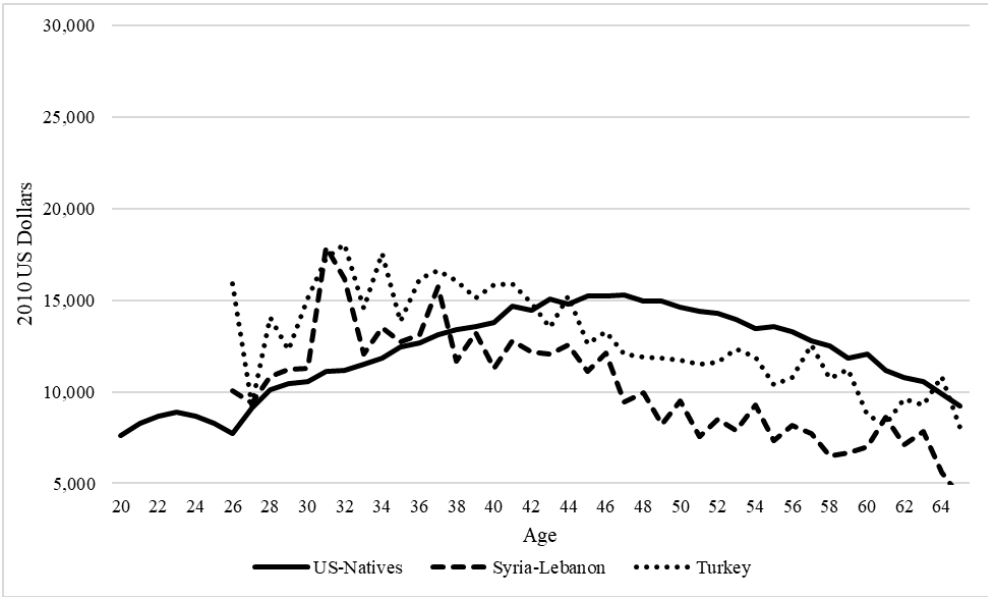


Note: SYLB and TR refer to immigrants from Syria-Lebanon and Turkey respectively.

Source: Authors' estimates based on US census data.

Figure A2: Age-income profiles for US-natives and Ottoman immigrants using actual wages.

A. 1940 sample



B. 1940 sample excluding individuals with non-wage earnings.

Source: Authors' estimates based on US census data.

Paper IV

Home-Leaving Patterns Among MENA-Origin Youth in Sweden: The Influence of Origin, Generation, and Neighbourhood

Rami Zalfou and Anna Tegunimataka

Abstract

Migrant integration extends beyond the labour market, including key life transitions such as leaving the parental home, which reflects economic independence and social autonomy. This study examines home-leaving patterns among Middle Eastern (MENA) immigrants and youth with MENA background in Sweden, using population register data (1998–2022) for individuals aged 17 to 35. Applying competing risks models, we analyse economic and neighbourhood factors influencing the transition to independent living. Findings reveal persistent gaps between MENA-origin youth and native Swedes, with the gaps growing among second-generation immigrants. Co-ethnic community size and neighbourhood unemployment matter significantly for transitions, while parental income has limited influence. Gender disparities are pronounced, with women experiencing larger gaps in leaving the parental home, reflecting the influence of cultural norms for integration trajectories.

1. Introduction

Leaving the parental home is a key milestone in the transition to adulthood, often associated with growing economic independence and social autonomy. In Western European societies, this transition tends to occur earlier and more frequently than in non-Western contexts. However, immigrant youth often follow distinct home-leaving trajectories influenced by a combination of economic, cultural, and social factors (Gillespie et al., 2020; Kleinepier & de Valk, 2017; Zorlu & Mulder, 2011). These differences are central to a broader understanding of migrant integration, which extend beyond the labour market into areas such as housing, family formation, and social inclusion.

This study focuses on Middle Eastern and North African (MENA)-origin youths in Sweden, a country with a generous welfare system that supports early transitions to adulthood through policies such as free university and universal student allowances. Swedish youth, on average, leave the parental home earlier than youths in other EU countries (Eurostat, 2024). However, these patterns do not necessarily extend to all segments of the population. Sweden has seen a significant increase in immigration from the MENA region in recent decades, and youths from these backgrounds likely have distinct home leaving patterns. Such differences may reflect persistent cultural expectations around family and co-residence, as well as unequal access to economic resources, often linked to higher unemployment and lower income among the parental generation (SCB 2024). These factors may also be different across migrant generations, with first-generation youth often facing more immediate constraints and second-generation often find themselves balancing their parents' traditions with the expectations of the host country, but also by gender. Girls from MENA backgrounds may face stronger expectations around family responsibility, protection, or modesty (Berry et al 2006), which can delay or complicate independent living. Boys, on the other hand, may experience greater freedom but also increased pressure to contribute economically to the household (Berry et al 2006). These patterns can also differ depending on country of origin, reflecting diverse cultural norms, migration histories, and integration.

An important consequence of the increasing migration to Sweden has been the emergence of large, ethnically concentrated neighbourhoods in major cities. Youth from MENA backgrounds often grow up in these areas, where strong ties to the co-ethnic community potentially has an influence on their home leaving decisions. The presence of a large, cohesive immigrant community may ease the pressure to leave home by providing strong social support networks and reinforcing norms around intergenerational co-residence (Portes & Rumbaut, 2001; McAvay & Pailhé, 2021). In contrast, youth living in areas with smaller or more dispersed co-ethnic communities may face greater cultural dissonance and instead pressure to adapt to mainstream expectations of autonomy (McAvay & Pailhé, 2021).

Using Swedish population register data from 1998 to 2022; this paper examines the home-leaving patterns of individuals aged 17 to 35. We track transitions from living with parents to either independent living or to partnership formation, focusing on first- and second-generation individuals from MENA backgrounds. We compare these patterns to those of native-born Swedes with two Swedish-born parents, as well as to selected European immigrant groups. Our analysis employs competing risks models to estimate the timing and likelihood of leaving the parental home, accounting for marriage as a competing destination. We include measures of generational status, gender, post-secondary education and employment status, own and parental socio-economic status, and contextual factors like the size of the co-ethnic community in the local area as well as unemployment levels in the local area.

Our results show that youth from MENA backgrounds are less likely to leave the parental home to live independently and are more likely to transition directly from the parental home to living in a married household. This pattern is particularly apparent in second generation immigrants with two foreign-born, even when compared with MENA-born youth that arrive in Sweden as children. This outcome is somewhat surprising, considering that the second generation is born in Sweden and likely more socially integrated, with greater exposure to Swedish cultural norms and institutions that generally emphasize independent living during young adulthood. We also find this pattern to be stronger amongst women as compared to men. Overall, we find that individual and parental socio-economic conditions are not major drivers of these gaps, and that they persist when controlling for individual contextual factors and socio-economic conditions, as well as parental socio-economic conditions, and when looking across parental income quantiles. This suggests that the observed differences are less about material constraints and more likely related to cultural preferences, family norms, or community-level expectations around the timing and pathways out of the parental home.

When looking at the role of the neighbourhood, we find that both the native share of the population in the local neighbourhood, and the share unemployed in the local neighbourhood, matter for the gaps in transitions between immigrants and natives. These findings suggest that norms around independence and early marriage persist across generations and are reinforced by local community contexts. However, there is also evidence of generational divergence, with youth who arrived in Sweden at younger ages showing more alignment with majority Swedish patterns of home-leaving as compared to later arrivals.

This study contributes to the literature on migrant integration by highlighting generational differences within immigrant-origin populations, gender gaps, and the role of neighbourhood context in producing diverse home-leaving trajectories among MENA-origin youth. It also addresses an unresolved question in the European context: whether fine-grained generational categories correspond to consistent differences in the timing of leaving the parental home.

2. Literature review

Research on home-leaving explores a number of factors that influence when young adults move out from their parental home. This transition is driven by a combination of personal decisions, economic conditions, and cultural expectations (Marini 1984). Important life events, such as securing employment, entry into higher education, and partnership formation, all can give the opportunity and motivation to move out among young people (Goldscheider & Goldscheider, 1987). Structural conditions, such as housing prices and governmental support for youth can either delay or speed up the process (Marini 1984). However, moving out triggers can vary significantly depending on the individual's background, access to resources, and expectations from the family. Cultural norms are particularly influential for youth from immigrant backgrounds, where ideas about family responsibilities, independence, and gender roles may contrast with those of natives in the host country (De Valk and Liefbroer 2007).

2.1. Structural and economic conditions

In the Swedish context, the transition out of the parental home is supported by the social democratic welfare regime, with a stable age at leaving the parental home even during economic downturns and spikes in youth unemployment rates (Buchmann & Kriesi 2011, Dribe & Stanfors 2005). Overall, the age at leaving the parental home changed little in Sweden over the second half of the 20th century, despite fundamental social and economic changes (Dribe & Stanfors 2005). However, recent literature shows that economic setbacks, such as becoming unemployed, are associated with increased return to the parental home (Olofsson et al. 2020). This underscores the importance of job availability and wages, as financial security increases likelihood of living independently (Olofsson et al. 2020). Parental resources also matter for young people's home-leaving opportunity as many young need the financial support of their parents to cover housing and living costs (Zorlu & Mulder 2011). Economic conditions are especially relevant for migrant youth. Hoolachan et al. (2017) highlight how young people face distinct financial challenges, such as higher unemployment and less job stability, which could hinder their ability to afford independent housing. This is frequently the case for immigrant youth and youth with immigrant background, who tend to experience greater economic uncertainty, higher unemployment, and labour market discrimination in many contexts (Gorodzeisky & Semyonov, 2017; Carlsson & Rooth, 2007).

In response to these challenges, education often becomes a central focus. Young people from immigrant backgrounds often prioritize higher education (Salikutluk 2016: Jonsson & Mood 2023), which can be an important path toward economic mobility in the host country (Borjas 2006). In Sweden, where student allowances

and housing options are available, attending university may provide an opportunity to leave the parental home.

2.2. Cultural influences, migrant contexts and gender

Cultural factors play an important role in influencing the home-leaving behaviour of migrant youth, particularly for those with origins in the Middle East and North Africa (MENA). In much of the MENA region, leaving the parental home is closely tied to marriage (Sonneveld, 2019). For women, marriage is often a prerequisite due to cultural, religious, and legal restrictions (Gebel & Heyne, 2014, p. 228), while men are typically expected to be able to provide financially for a new household before marrying (Kovacheva et al., 2018). Over recent decades, rising participation in higher education and delayed marriage have pushed the average age of home-leaving higher in the MENA region (Carmichael, 2011). Nonetheless, large gender gaps remain, with women marrying and thus leaving home, several years earlier than men (Koç, 2007).

These cultural norms can carry over to migrant populations, influencing behaviour in the host country. Compared to their native peers, migrant youth often face different circumstances and expectations. Immigrant youth may feel obligated to remain in the parental home to contribute financially or logistically to the household (Van Hook & Glick, 2007). They may also assume unique family responsibilities, such as language brokering, helping parents navigate linguistic and bureaucratic barriers (Villanueva & Buriel, 2010; Melander & Shmulyar Gréen, 2023). These roles can delay departure from the parental home as their presence becomes essential to family functioning, while intergenerational transmission of values further influence home-leaving patterns. De Valk and Liefbroer (2007), studying migrant families in the Netherlands, find that in many communities, especially those from collectivist cultural backgrounds, remaining at home until marriage or another major life milestone is a deeply embedded norm. This is not solely due to economic constraints but also reflects a cultural emphasis on familial support and unity. Gender norms play a role as well: young women in migrant families often face stronger expectations to stay at home, reflecting traditional caregiving roles even in context that could facilitate greater independence (Yalim et. al 2023).

Patterns in previous research are not uniform. Using Dutch administrative data, Zorlu and Mulder (2011) report that Turkish and Moroccan youth leave home at younger ages than native Dutch, most often for union formation and also for independent living, even though later home-leaving is common in the countries of origin. The authors point to factors such as the availability of welfare supports and migrant youths' position between origin and host-country norms as possible explanations.

Migrants' generational status also matters. Rumbaut (2004) distinguishes the second generation with two foreign-born parents, the 2.5 generation with one foreign-born and one native-born parent, the third generation with two native-born parents, and decimal generations among the foreign-born (for example, 1.75 for arrivals at ages 0–5 and 1.5 for arrivals at ages 6–12). Children with a native-born parent tend to have greater fluency in the official language and stronger identification with the host society, which can facilitate earlier home-leaving than among first-generation youth. Recent United States evidence links these gradations to home-leaving: Gillespie, Bostean, and Malizia (2020) find the highest hazards of departure among third-plus generation youth, followed by the second generation, then the 1.75 and 1.5 generations, with parental region of origin and non-English home language further reducing exit risks, particularly among youth with Latin American-born parents.

In Canada, analyses show that both the timing and destination of exits differ by immigrant generation and visible-minority status, although not in a uniform or strictly monotonic way across groups (Haan, Cheng, and Yu, 2023). Yet it remains unclear whether such nuanced generational categories correspond to consistent differences in when young people leave the parental home in European settings.

In this study, we distinguish between two main routes out of the parental home: on the one hand, independent living or nonmarital cohabitation, and on the other, direct transition into marriage. The distinction is crucial because the prevalence and social significance of cohabitation and marriage differ across societies and migrant generations. Across Europe, cohabitation serves varied roles, and its meaning varies by country (Hiekel, Liefbroer, & Poortman, 2014). Immigrants and their descendants also follow distinct union pathways: in the United Kingdom, origin groups differ in propensities to cohabit or marry (Hannemann & Kulu, 2015) and in the Netherlands, union choices among the Turkish and Moroccan second generation are closely tied to social embeddedness and networks (Huschek, de Valk, & Liefbroer, 2011).

2.3. Neighbourhood and ethnic community influence

The neighbourhood ethnic composition may have a direct influence on home-leaving decisions, yet few studies have investigated the direct influence of neighbourhood composition on home-leaving behaviour (McAvay & Pailhé 2022). The neighbourhood is important for the socialization of youth, as it influences behaviours, peer interactions, exposure to role models, access to social capital and social control. These factors contribute to the development of behaviours, attitudes, and aspirations (Sampson 2012). Immigrant youth that live in neighbourhoods with a large co-ethnic group is more likely to have their day-to-day interactions within the ethnic groups (van Tubergen & Maas 2007) and a greater exposure to the norms and expectations of that particular group also regarding the timing of life events

such as leaving the parental home Moreover, as highlighted by research on residential behaviour among international migrants (e.g. Bowes et al., 1997; Aradhya et al 2017), the presence of co-ethnics and family members in the neighbourhood significantly affect residential choices. Migrants often choose to settle near people from their country of origin upon arrival, with family proximity being important in location decisions. These preferences are strengthened by cultural norms of family solidarity, which are often stronger among non-Western migrant groups compared to the native-born population (Dykstra, P. A., & Fokkema, T. 2012). Such norms can foster a reluctance to move away from family members, reinforcing intergenerational co-residence and delaying home-leaving among youth.

McAvay & Pailhé (2022) extend this discussion to the French context, showing that neighbourhood composition plays a role for home-leaving decisions among immigrant youth. Their study finds that young people from immigrant backgrounds living in immigrant-dense areas tend to leave home later, net of individual, family, and contextual controls. The association differs by origin and gender: for example, the negative link is especially clear for leaving to unmarried cohabitation, while for some groups (e.g., North African women) higher neighbourhood immigrant shares are linked to more exits into marriage. These findings highlight the importance of considering local residential environments when analysing the transition to adulthood, particularly among immigrant populations.

The neighbourhood and its composition also matter for the opportunities in terms of jobs and education, as well as finding independent housing. There may also be discrimination directed towards those residing in a particular neighbourhood which may impact their possibilities of finding housing elsewhere (Pager & Shepherd 2008).

Summarizing the above literature review, we find that studies point to marked patterns in home-leaving between native and MENA-background youth. Youth from MENA backgrounds, especially women, tend to leave the parental home later due to cultural expectations that link home-leaving to marriage and emphasize family responsibilities. Young women in migrant families often face stronger expectations to stay at home, while men's departure is more often linked to their ability to provide economically. Patterns across migrant generations are less clear, as previous studies reach different conclusions. Nonetheless, some find that second-generation youth, particularly those with one native-born parent, tend to leave home earlier than first-generation peers, though often still later than natives, with pathways varying by context and origin, although fewer studies have focused on the European context. Neighbourhoods with a high concentration of co-ethnics can reinforce cultural norms and family ties, making youth more likely to remain at home longer. In these areas, limited access to housing and jobs may also delay the possibility of moving out.

3. Data and methods

We use Swedish longitudinal administrative data that provides annual observations on all individuals legally residing in Sweden. The data includes comprehensive demographic characteristics, such as age, sex, country of birth, and place of residence, as well as detailed information on individuals' linked parents. For individuals and their parents, we observe educational attainment, labour market participation, income, and occupation, allowing us to account for a wide range of background variables that may influence young people's transition out of the parental home. Our analytical sample consists of individuals born between 1981 and 1992. These cohorts are tracked from age 17 until the year they turn 35, providing a long observational window that captures both early and delayed transitions out of the parental home.

The primary outcome of interest is the timing and type of departure from the parental home. We define "leaving the parental home to live independently" as a situation in which the individual no longer resides with either parent and is not registered as part of a married household. This definition captures individuals who move out to live alone, with roommates, or with a partner outside of marriage. Due to limitations in the administrative data, we are unable to observe cohabitation directly; however, since cohabiting unions are common in Sweden and indistinguishable from other independent living arrangements in register data (unless the partners have common children), we classify such cases as independent living.

We focus primarily on the transition from the parental home to independent living, distinguishing it from direct transitions into marriage. The latter is treated as a competing pathway, as individuals may exit the parental home either to live independently or to form a marital household. This distinction is crucial, particularly in the context of intergenerational differences and cultural variation, where norms around marriage and cohabitation may differ significantly (discussed above). Accordingly, our analysis adopts a competing risks framework to model these two mutually exclusive pathways out of the parental home.

We focus on immigrants and their offspring with origin in the MENA region and include individuals from the top seven sending countries in that region (Afghanistan, Iraq, Iran, Lebanon, Somalia, Syria, and Turkey). We split both the immigrant and the second-generation samples into several groups based on age at arrival in Sweden for immigrants and based on parental origin for the second generation. The sample is split into six generations:

Generation 1: arrive in Sweden at age 17.
Generation 1.25: arrive at ages 13 to 16.
Generation 1.5: arrive at ages 6 to 12.
Generation 1.75: arrive at ages up to 5.
Generation 2: born in Sweden to two foreign-born parents.

Generation 2.5: born in Sweden to one foreign-born and one Sweden-born parent.

We compare the outcomes in these groups to those of native-born Swedes with two Sweden-born parents, and we also use a comparator group of migrants originating in Eastern European countries (Poland, Russia, Romania, Bulgaria, Hungary, and Ukraine), which constitute important immigrant sending countries in Sweden.

We use competing risks models which measure the relative risk of leaving home into independent living, with transitioning directly to marriage/cohabitation treated as a competing risk. This method follows the Fine & Gray (1999). The model allows for estimating the effects of different covariates on the cumulative probability of the event of interest while accounting for competing risks. The hazard function is defined as:

$$\lambda_k(t|X) = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T < t + \Delta t,\ D = k \mid T \geq t, X)}{\Delta t}$$

Where T is the event time and k is the event type. The regression equation takes the following form:

$$\lambda_k(t|X) = \lambda_{k0}(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \cdots)$$

Where $\lambda_k(t|X)$ is the subdistribution hazard for cause k at time t given the covariates in X. $\lambda_{k0}(t)$ is the baseline subdistribution hazard for cause k, and the X's refer to the included covariates.

We treat the native group with two native-born parents as our baseline and add a dummy variable for each region of origin and generation. We control for a set of individual factors, which include post-secondary education, employment status, total yearly income, and occupation. We further control for parental socio-economic status by adding controls for the sum of parental income, and the employment status and occupation of each parent. Finally, we control for county fixed effects, and commune fixed effects in an alternative specification, to capture geographic heterogeneity. To explore the role of the neighbourhood, we construct variables at the postcode level that capture the share of Sweden-born individuals in each postcode, and the share of unemployed individuals in each postcode.

The analysis relies on the Fine and Gray (1999) competing risks framework, which models the subdistribution hazard of each event type. A key implication of this approach is that individuals remain in the risk set even after they experience a competing event. For example, individuals who exit the parental home through marriage are retained in the risk set for independent living, albeit with appropriately weighted risk contributions (Fine and Gray 1999). This feature allows the model to estimate the cumulative incidence of independent living while correctly accounting for marriage as a competing pathway, rather than censoring such cases. At the same

time, it means that the risk sets used in the estimation do not correspond exactly to the set of individuals who are "at risk" in a substantive sense after experiencing a competing event. Readers should therefore interpret the coefficients as pertaining to differences in the relative incidence of outcomes, not to literal hazards conditional on remaining unmarried or at home.

Another limitation of the Fine & Gray model is that it is restricted to the first observed transition and does not capture subsequent moves (for instance, returning to the parental home or moving between independent living and cohabitation). Our focus is therefore on the timing and type of first departure from the parental home. The assumption of mutually exclusive and exhaustive event types simplifies the complexity of residential trajectories, while events such as international migration or informal cohabitation are not accounted for in the models.

Overall, the Fine & Gray approach is well suited to our research question, as it allows us to directly compare the relative likelihood of leaving the parental home through independent living versus marriage, while accounting for the fact that one type of transition precludes the other. However, it also entails that results should be interpreted as differences in cumulative incidence across groups, rather than as instantaneous risks in the presence of dynamic, sequential life-course events.

4. Results

4.1. Descriptive statistics

Table 1 shows the descriptive statistics for the analytical sample. The majority of individuals in the sample transition to living independently by the end of the observation period at 35 years of age (97.8% in the native group). Though the proportion differs by region of origin and by generation. We find that the proportion is generally lower in the MENA origin sample, measuring between 81% and 88% in the first generation and declining to 78% in generation 2. The mixed 2.5 generation on the other hand appears to be more similar to the natives with a share at 92%. The share transitioning directly from the parental home to living in a married household is at only 1.5% in the native sample and appears to be higher in the MENA first generation measuring between 9% and 13% and increasing to 14% in the second generation but remains at a much lower 3% in the 2.5 generation. The proportion remaining in the parental home without any transitions is at below 1% in the native sample, and measures between 3% and 5% in the first generation, increasing to 8% in the second. The MENA first-generation transition shares are similar to those in the Eastern European first generation, but they diverge in the second generation, with a comparatively high share of MENA-origin second generation youths transitioning directly to marriage.

Table 1: Descriptive statistics

Origin		Middle East						Eastern Europe					
Generation	Natives	2.5	2	1.75	1.5	1.25	1	2.5	2	1.75	1.5	1.25	1
Observations	5,651,598	16,125	114,966	39,553	45,687	16,332	3,231	32,003	27,584	9,443	7,879	3,965	525
Individuals	886,374	2,244	14,986	6,276	8,415	3,738	684	4,655	3,676	1,377	1,203	594	91
Sex (%)													
Male	52	52	52	54	56	60	55	52	52	53	53	56	43
Female	48	48	48	46	44	40	45	48	48	47	47	44	57
Status at the end of observation (%)													
Living independently	97.8	92.5	77.8	87.7	85.3	81.5	88.4	96.4	91.4	93.0	91.3	83.6	95.5
Married	1.5	3.1	13.8	8.3	10.6	13.1	8.7	0.9	2.2	3.8	6.1	5.0	3.4
In parental home	0.7	4.4	8.4	4.0	4.1	5.4	2.9	2.7	6.4	3.2	2.6	11.4	1.1
Post-secondary education (%)	9.5	15.8	23.3	19.2	16.8	14.5	13.8	17.9	23.4	19.4	25.2	18.2	17.1
In employment (%)	42.0	37.2	47.3	39.2	40.1	38.8	26.6	36.3	34.7	34.2	33.2	37.1	23.0
Total earned income (mean, SEK)	73,100	73,200	117,000	87,800	89,500	96,500	61,800	68,000	72,400	70,400	67,600	88,000	45,400
Occupation (%)													
Farming	7.6	2.5	2.7	3.0	3.2	2.6	1.5	5.1	3.2	4.8	4.9	3.3	1.5
Managers	0.06	0.17	0.22	0.19	0.16	0.11	0.00	0.08	0.02	0.12	0.14	0.00	0.00
Military	0.13	0.13	0.04	0.03	0.02	0.03	0.00	0.13	0.08	0.15	0.01	0.03	0.00
Professional	2.6	2.9	3.2	3.5	2.8	1.3	3.2	3.5	3.5	4.0	4.5	1.8	2.3
Services	20.6	19.3	17.6	20.3	18.4	12.0	13.6	18.8	17.0	17.4	17.1	10.0	15.4
Unknown	69.0	75.1	76.2	72.9	75.4	83.9	81.6	72.3	76.3	73.5	73.3	84.8	80.1
Income quartile (%)													
Quartile 1	21.5	30.4	27.1	31.0	31.7	35.9	42.2	28.7	33.4	33.3	34.3	38.3	41.8
Quartile 2	23.7	20.0	14.4	18.3	17.5	15.7	19.9	21.9	19.2	21.3	20.6	12.0	23.8

Table 1 (continued)

Origin		Middle East						Eastern Europe					
Generation	Natives	2.5	2	1.75	1.5	1.25	1	2.5	2	1.75	1.5	1.25	1
Quartile 3	27.2	23.6	20.4	21.7	21.4	17.1	20.1	25.3	23.0	21.6	22.8	20.2	20.0
Quartile 4	27.5	25.9	38.0	29.0	29.4	31.3	17.8	24.0	24.4	23.8	22.4	29.5	14.3
Father in employment (%)	90.8	79.0	72.0	68.1	63.6	45.1	56.2	80.0	74.5	77.9	79.7	79.7	77.5
Mother in employment (%)	90.9	86.1	68.3	70.0	63.1	41.2	54.7	85.4	82.0	82.6	80.7	74.6	83.8
Total parental income (mean, SEK)	610,900	546,200	393,300	365,800	313,800	221,400	291,900	579,800	520,300	513,600	515,800	514,200	604,900
Parental income quartile (%)													
Quartile 1	10.5	19.0	38.5	41.7	49.8	67.8	56.7	18.3	22.7	22.9	25.0	28.4	25.1
Quartile 2	24.5	25.7	29.3	30.4	28.6	17.1	24.0	24.6	26.9	30.2	29.6	27.9	22.8
Quartile 3	31.8	25.7	18.5	17.7	14.6	9.3	11.1	25.1	24.2	23.6	21.3	21.1	20.0
Quartile 4	33.1	29.6	13.6	10.2	7.0	5.8	8.2	31.9	26.2	23.2	24.0	22.5	32.0
Number of postcodes	10,439	2,532	4,394	3,415	3,546	2,161	857	4,291	3,207	1,776	1,558	923	171
Postcode categories (%)													
Native >=75%	93.4	75.7	43.4	48.8	42.1	37.3	48.7	86.7	70.0	65.7	65.7	50.8	66.9
Unemployed <25%	59.7	63.9	54.2	62.4	62.5	63.3	75.5	65.2	66.5	67.3	68.9	63.8	78.1
Native >=75% & Unemployed <25%	37.2	26.0	18.3	17.3	14.9	13.2	12.3	29.4	22.5	20.1	19.7	16.5	12.2
Native >=75% & Unemployed >=25%	56.2	49.7	25.0	31.5	27.3	24.1	36.5	57.4	47.5	45.6	46.0	34.3	54.7
Native <75% & Unemployed <25%	3.0	10.0	27.4	20.3	22.6	23.5	12.2	5.4	11.0	12.6	11.4	19.7	9.7
Native <75% & Unemployed >=25%	3.6	14.3	29.2	30.9	35.2	39.2	39.1	7.9	19.0	21.7	22.9	29.5	23.4

Note: The table shows descriptive statistics for the sample used in the analysis. Measures are evaluated using all observations, except for sex and status at the end of observation which are evaluated based on individuals.

The share of individual observations with post-secondary education measure at 9.5% in the native sample and appear to be elevated in the first-generation MENA-origin groups, measuring between 14% and 19%, and increasing to 23% in the second generation. The second generation also stands out in the high share in employment (47%) as compared to natives (42%), and in their average yearly incomes (117,000) as compared to the corresponding value in the native group (73,100), and as compared to the figure in the second-generation Eastern European sample (72,400).

The patterns in parental income contrast with own earned income, with the native sample having the highest mean of parental income at 610,900, while the first- and second-generation MENA samples have lower values ranging from 221,400 to 393,300. These figures also contrast with the corresponding values in the Eastern European sample, where parental income is much closer to the native mean and consistently above 500,000. These patterns are also reflected in the disparities seen in the parental income quartiles between MENA migrants and the native sample, and in parental employment figures.

We construct two variables at the postcode level to capture the conditions in the local neighbourhood: a dummy variable that takes the value 1 if the share of natives is at 75% or more, and 0 otherwise, and a dummy variable that takes the value 1 if the share of individuals who are unemployed is at 25% or more, and 0 otherwise. While the vast majority of natives Swedes live in neighbourhoods with a native share of 75% or more, less than half of the MENA population in our samples live in such neighbourhoods.[1]

4.2. Differences across groups

Table 2 presents the results from competing risks regressions, estimated separately for males and females. The coefficients are reported as subdistribution hazard ratios (SHRs), which have been exponentiated for easier interpretation. A coefficient of 1 indicates no difference in the subdistribution hazard compared to the reference group. A coefficient below 1 indicates a lower subdistribution hazard, meaning the group experiences the event of interest at a slower rate over time compared to the reference. For example, a coefficient of 0.5 suggests that the group's subdistribution hazard is half that of the reference group, although this does not mean the absolute risk is exactly halved.

[1] The 75% cutoff was selected based on the distribution of immigrants and the second generation across neighbourhood native shares, with 75% being near the midway point of the distributions. See Appendix figure A2.

Table 2: Coefficients of competing risks regression

	Male		Female	
	(1)		(2)	
Reference group	Natives		Natives	
Origin	MENA	EEURO	MENA	EEURO
Generation				
2.5	0.824***	0.846***	0.695***	0.836***
	(-6.81)	(-8.58)	(-12.16)	(-9.28)
2	0.511***	0.701***	0.376***	0.629***
	(-52.88)	(-15.62)	(-62.45)	(-18.65)
1.75	0.690***	0.770***	0.505***	0.696***
	(-19.79)	(-6.84)	(-28.20)	(-8.54)
1.5	0.657***	0.648***	0.464***	0.634***
	(-23.45)	(-10.02)	(-24.81)	(-9.22)
1.25	0.653***	0.606***	0.458***	0.576***
	(-15.10)	(-8.85)	(-19.30)	(-9.05)
1	0.670***	0.934	0.552***	0.634***
	(-6.02)	(-0.33)	(-9.32)	(-3.33)
Controls				
Unemployed	0.686***		0.907***	
Post-secondary education	1.338***		1.374***	
Income	1.000***		1.000***	
Occupation	Y		Y	
Parental controls	Y		Y	
County	Y		Y	
Observations	3,342,797		2,624,073	

Note: The table reports the exponentiated regression coefficients for the competing risks regression with the failure event defined as leaving the parental home to live independently and the competing risk event defined as leaving the parental home to live in a partnership. t statistics in parentheses. See appendix table A1 for full regression output.

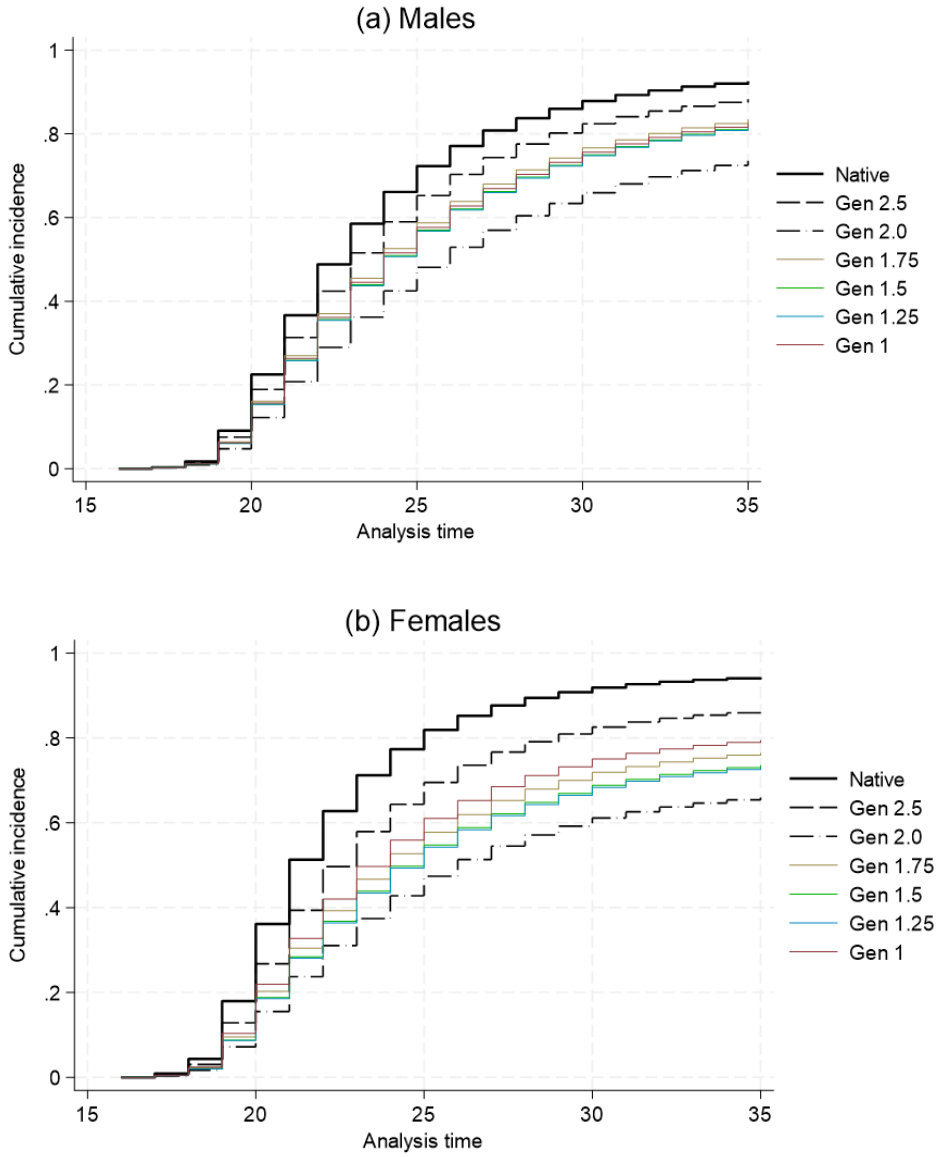
Overall, the coefficients in table 2 are at values below 1, which points to reduced subdistribution hazards for the transition to independent living, in both the Middle Eastern and Eastern European immigrant groups and in the second generation, for both males and females. The effect size measure between 0.65 and 0.69 in the MENA first generation males, but is lower at 0.51 in generation 2, while the coefficient is higher at 0.82 in generation 2.5. The gap with natives is larger in the MENA female population, with coefficients measuring between 0.46 and 0.55 in the first generation, decreasing to 0.38 in generation 2. Comparing these effect sizes to those of the Eastern European sample, they appear to be similar in the first-generation males but diverge in the case of females in generation 1, and both males and females in generation 2.

Looking at the control variables, unemployment decreases the subdistribution hazard, especially for males. Post-secondary education increases the subdistribution hazard and has a similar effect for both males and females, while income has a vanishingly small effect. The addition of individual and family controls only shifts the gap coefficients by a small measure of 0.1 or less, as compared to the unadjusted regression[2].

Figure 1 shows the cumulative incidence functions for MENA immigrants and the second generation alongside those of the native sample using the models estimated in table 2, for both males and females. The cumulative incidence functions show the proportion of individuals that have transitioned to living independently by a given age, giving a better view of both the timing of transitions and the proportion eventually making the transition. In the case of native males, over 80% transition by the age of 27. Gen 2.5 matches this group closely, with 80% completing the transition by age 29. The first-generation immigrants transition at a slower rate, reaching 80% at age 31. Meanwhile, the transitions are slowest in generation 2, where only around 60% transition by age 28, and only around 75% transition by the end of the observation period. In the case of females, the gaps are larger, with the share transitioning at less than 80% in the first generation, and only around 65% in generation 2.

[2] See appendix table A3 for unadjusted estimates. The results are robust to using commune fixed effects in table A4.

Figure 1: Cumulative incidence functions



Note: The cumulative incidence functions are based on the models controlling for individual controls, parental controls, and county fixed effects.

4.3. The role of parental income

To explore the role of parental income, we estimate the competing risks regression coefficients interacted with parental income quartiles, calculated using the sum of the two parents' incomes. The results are shown in Table 3. We find that the size of coefficients varies only slightly between the bottom three income quartiles, with the top income quartile showing overall smaller gaps as compared with natives across generations 1.25 to 2 in males, but less so for females. Overall, the gaps with natives are largely robust across parental income quartiles. This result suggests a weak role for parental income in shaping the transition to living independently.

Table 3: Competing risks regression coefficients according to parental income

	Male				**Female**			
	(1)				(2)			
Reference	Natives, Quartile 1				Natives, Quartile 1			
Origin	MENA				MENA			
Quartile	1	2	3	4	1	2	3	4
Generation								
2.5	0.755**	1.039	0.842**	1.265***	0.648***	0.766**	0.619***	0.762***
	(-3.03)	(0.47)	(-2.71)	(5.76)	(-4.68)	(-3.09)	(-6.53)	(-6.32)
2	0.483***	0.507***	0.549***	0.819***	0.370***	0.344***	0.367***	0.417***
	(-23.32)	(-17.95)	(-20.28)	(-8.73)	(-29.87)	(-25.71)	(-28.86)	(-24.73)
1.75	0.725***	0.716***	0.809***	1.025	0.539***	0.521***	0.470***	0.523***
	(-7.78)	(-6.69)	(-5.21)	(0.76)	(-12.70)	(-11.36)	(-14.19)	(-14.13)
1.5	0.703***	0.746***	0.717***	0.945	0.512***	0.473***	0.518***	0.403***
	(-10.02)	(-6.54)	(-8.18)	(-1.73)	(-16.32)	(-13.69)	(-13.61)	(-10.39)
1.25	0.701***	0.737***	0.670***	0.922	0.484***	0.431***	0.441***	0.452***
	(-7.66)	(-4.20)	(-5.78)	(-1.46)	(-12.36)	(-7.85)	(-7.73)	(-8.51)
1	0.764**	0.916	0.544**	0.853	0.590***	0.494***	0.320***	0.711**
	(-2.66)	(-0.58)	(-2.88)	(-1.14)	(-5.38)	(-3.59)	(-4.78)	(-2.98)
Observations	3,315,230				2,602,251			

Note: The table reports the exponentiated regression coefficients according to parental income quartile. The reference group is native Swedes in the bottom income quartile. The regressions control for employment status, post-secondary education, occupation, income, parental controls, and county fixed effects. t statistics in parentheses.

4.4. Heterogeneity by country of origin

To explore heterogeneity by country of origin, we estimate models that interact country of origin with generation, instead of treating all MENA-origin individuals as a single group. We plot the resulting cumulative incidence functions by country of origin for males and females in figure 2. We find substantial heterogeneity by country of origin for males, with some countries matching the native trajectories more closely, especially Iran and Iraq. The degree of the divergence in generation 2 differs by country of origin, and appears strongest in the case of Afghanistan, but is present to some extent in most origin countries. Overall, there is less variation by country of origin in the case of females.

Figure 2: Cumulative incidence functions according to country of origin

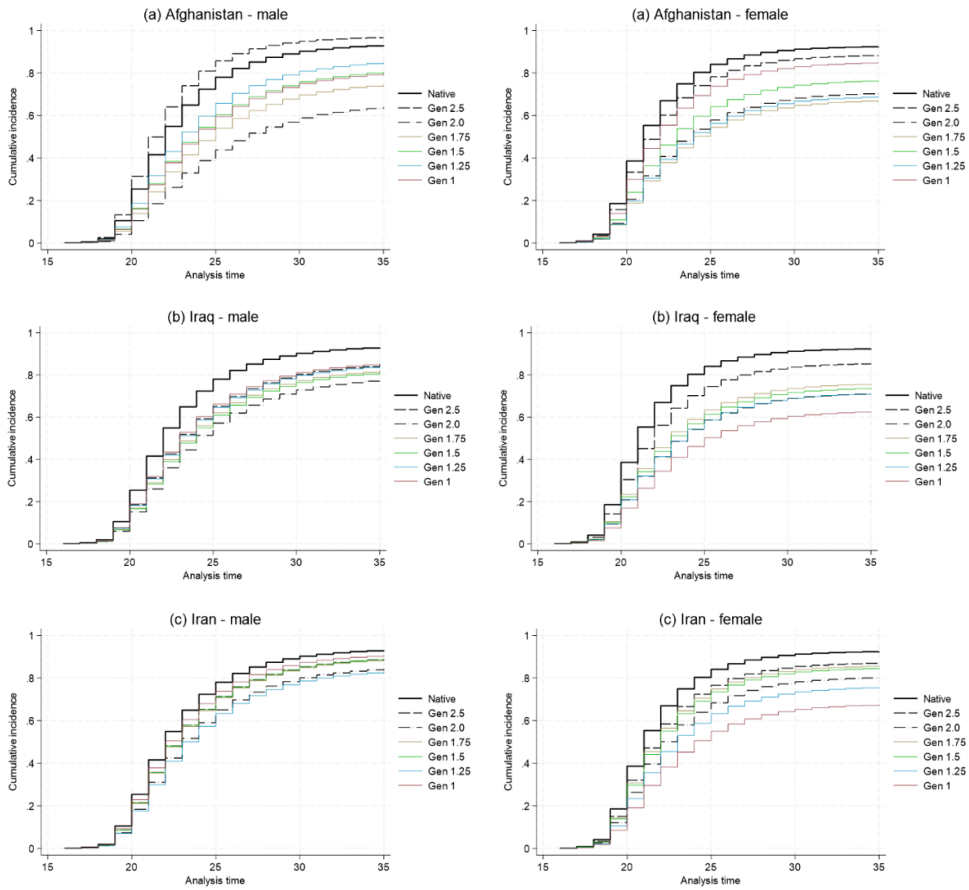
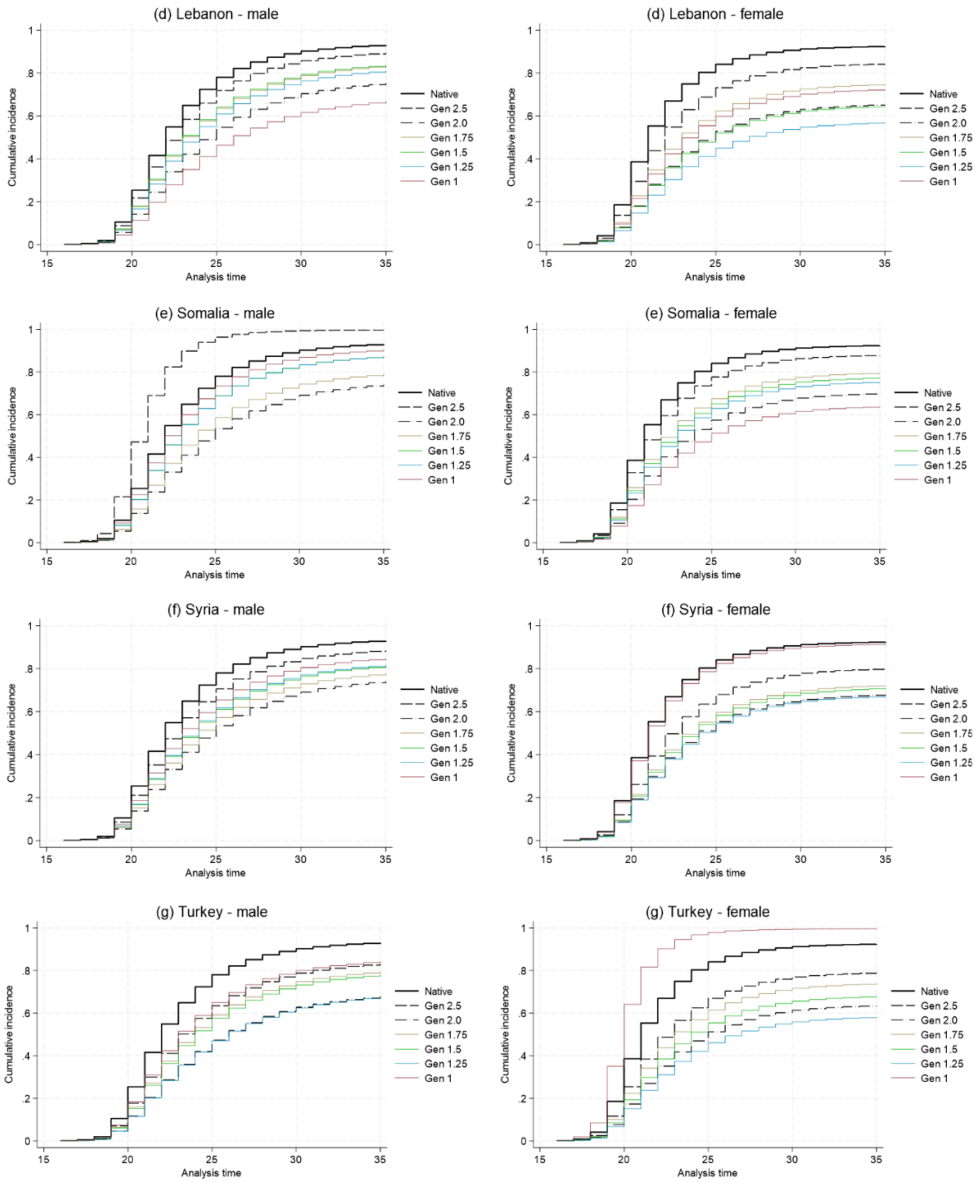


Figure 2 (continued)



4.5. The role of the neighbourhood

We test the contribution of two variables to explaining the gaps in transitions between immigrants and natives – the share of natives in the local postcode, and the share of unemployed individuals in the local postcode. Table 4 shows the size of the coefficients of competing risks regressions where we interact with a dummy variable that takes the value 1 if the share of natives is at 75% or more and 0 otherwise. We find significant differences in the size of the coefficients across neighbourhood native shares, in generations 1.5 to 2, for both males and females. In the case of generation 2 males, the effect sizes in neighbourhoods with high native shares is at 0.7 as compared to 0.54 in neighbourhoods with low native shares. The corresponding coefficients are 0.51 and 0.38 in the case of females.

Table 4: Coefficients of competing risks regression according to postcode native share

	Male		**Female**	
	(1)		(2)	
Reference	Natives, >=75%		Natives, >=75%	
Origin	MENA	MENA	MENA	MENA
Native share	>=75%	<75%	>=75%	<75%
Generation				
2.5	0.917*	1.057	0.777***	0.829**
	(-2.42)	(0.86)	(-7.10)	(-2.58)
2	0.703***	0.544***	0.510***	0.385***
	(-17.17)	(-30.68)	(-29.45)	(-40.98)
1.75	0.894***	0.742***	0.653***	0.524***
	(-3.75)	(-9.71)	(-12.08)	(-17.06)
1.5	0.848***	0.724***	0.609***	0.489***
	(-5.54)	(-11.91)	(-13.36)	(-14.94)
1.25	0.815***	0.759***	0.502***	0.556***
	(-3.88)	(-7.24)	(-9.57)	(-11.04)
1	0.839	0.745**	0.670***	0.614***
	(-1.79)	(-2.74)	(-4.62)	(-4.71)
Observations	3,297,430		2,588,050	

Note: The table reports the exponentiated regression coefficients for the competing risks regressions. The reference group is Sweden natives in postcodes with 75% or more native population. The regressions control for employment status, post-secondary education, occupation, income, parental controls, and county fixed effects. t statistics in parentheses.

Table 5 shows the coefficients interacted with a dummy variable that takes the value 1 if the share of unemployment is below 25% and 0 otherwise. The sizes of the coefficients are also significantly different between the two group. Males in high unemployment neighbourhoods transition to independent living at nearly half the rate of their counterparts in low unemployment neighbourhoods. The gaps are even larger in the case of females. Overall, the neighbourhood's unemployment share appears to be a much more significant determinant of transition rates as compared to the neighbourhood native share.

Table 5: Coefficients of competing risks regression according to unemployment

	Male		**Female**	
	(1)		(2)	
Reference	Natives, <25%		Natives, <25%	
Origin	MENA	MENA	MENA	MENA
Unemployed	<25%	>=25%	<25%	>=25%
Generation				
2.5	0.812***	0.409***	0.700***	0.119***
	(-6.07)	(-15.61)	(-9.37)	(-31.19)
2	0.494***	0.267***	0.341***	0.078***
	(-45.33)	(-36.90)	(-52.47)	(-50.33)
1.75	0.667***	0.354***	0.455***	0.102***
	(-16.99)	(-24.80)	(-24.39)	(-40.20)
1.5	0.626***	0.346***	0.403***	0.096***
	(-20.95)	(-26.41)	(-19.23)	(-42.37)
1.25	0.631***	0.332***	0.383***	0.090***
	(-12.89)	(-20.51)	(-15.50)	(-34.62)
1	0.624***	0.345***	0.584***	0.092***
	(-4.82)	(-10.55)	(-5.66)	(-22.94)
Observations	3,297,430		2,588,050	

Note: The table reports the exponentiated regression coefficients for the competing risks regressions. The reference group is Sweden natives in postcodes with less than 25% unemployment level. The regressions control for employment status, post-secondary education, occupation, income, parental controls, and county fixed effects. t statistics in parentheses.

We interact the two neighbourhood variables in Table 6, further splitting the sample according to both neighbourhood native share and neighbourhood unemployment level. We find the largest gaps in high-unemployment neighbourhoods, in both low-

and high- native share neighbourhoods. In the case of females, effect sizes are particularly stark in these neighbourhoods, measuring at around 0.1.

Table 6: Coefficients from competing risks regression according to unemployment level and native share at the zip code level

	Male				**Female**			
	(1)				(2)			
Reference	Natives, Unemployed <25%, Natives>=75%				Natives, Unemployed <25%, Natives>=75%			
Origin	MENA	MENA	MENA	MENA	MENA	MENA	MENA	MENA
Unemployed	<25%	>=25%	<25%	>=25%	<25%	>=25%	<25%	>=25%
Native share	>=75%	>=75%	<75%	<75%	>=75%	>=75%	<75%	<75%
Generation								
2.5	0.916*	0.440***	0.937	0.611***	0.795***	0.132***	0.762**	0.168***
	(-2.06)	(-12.21)	(-0.83)	(-4.55)	(-5.07)	(-26.82)	(-3.10)	(-14.91)
2	0.674***	0.363***	0.513***	0.294***	0.469***	0.103***	0.336***	0.0865***
	(-16.22)	(-22.86)	(-27.78)	(-27.77)	(-25.82)	(-40.29)	(-37.17)	(-42.84)
1.75	0.851***	0.463***	0.704***	0.386***	0.584***	0.135***	0.464***	0.110***
	(-4.33)	(-14.28)	(-9.15)	(-16.91)	(-11.49)	(-30.12)	(-15.51)	(-31.18)
1.5	0.788***	0.455***	0.681***	0.384***	0.522***	0.130***	0.418***	0.105***
	(-6.36)	(-14.96)	(-11.58)	(-18.54)	(-12.68)	(-30.80)	(-11.88)	(-35.10)
1.25	0.732***	0.458***	0.740***	0.375***	0.403***	0.107***	0.466***	0.112***
	(-4.67)	(-9.52)	(-6.33)	(-14.58)	(-8.26)	(-22.21)	(-9.42)	(-27.10)
1	0.838	0.403***	0.580**	0.431***	0.738*	0.113***	0.591**	0.111***
	(-1.31)	(-6.19)	(-3.18)	(-6.10)	(-2.25)	(-15.31)	(-3.24)	(-15.73)
Observations	3,297,430				2,588,050			

Note: The table reports the exponentiated regression coefficients for the competing risks regressions. The reference group is Sweden natives in postcodes with 75% or more native population and unemployment level below 25%. The regressions control for employment status, post-secondary education, occupation, income, parental controls, and county fixed effects. t statistics in parentheses.

5. Discussion and conclusion

Our results underscore three central findings: first, youth of MENA background are less likely to leave the parental home for independent living than their native

Swedish peers; second, these differences are especially pronounced among women and among the second generation; and third, neighbourhood context and particularly local unemployment, plays an important role in reinforcing these disparities, whereas parental economic background is of limited importance.

The finding that MENA origin youth are more likely to transition directly into marriage than into independent living is consistent with evidence from origin countries, where home leaving is closely tied to marital status, particularly for women (Koç 2007; Sonneveld 2019). At the same time, our results show that the majority of MENA origin youth in Sweden still leave for independent living, which contrasts with the dominant pattern in their countries of origin. This suggests partial adaptation to native norms alongside persistent differences. Prior studies also point to such a mixed picture. In the Netherlands, Turkish and Moroccan youth tend to leave home earlier than natives, yet their departures are more often linked to marriage or union formation rather than to independent living (Zorlu and Mulder 2011). In Canada, generational and visible minority differences further highlight this complexity, as second-generation youth sometimes converge with natives in timing, but pathways out of the parental home continue to diverge (Haan, Cheng and Yu 2023). Our findings fit into this broader evidence of both adaptation and divergence. In the Swedish case, we show that second generation MENA youth, particularly women, are less likely to leave independently than their first-generation peers, while the 2.5 generation more closely resembles natives.

The so called "second generation paradox" highlights that assimilation does not follow a simple, linear trajectory. Instead of steadily converging toward native patterns, certain outcomes show a persistence or even intensification of disadvantage across generations. In the Swedish context, recent studies point to such dynamics. Second generation immigrants have been found to face higher risks of unemployment (Aradhya et al. 2023) and a reversal of the mortality advantage often observed among first generation immigrants (Wallace 2022). Our findings on home leaving align with this broader pattern. Rather than moving closer to native norms, children of immigrants diverge. While first generation youth who arrived in childhood tend to assimilate the earlier their age at arrival, their second-generation peers are less likely to transition out of the parental home in ways that mirror the native population.

This divergence can be reinforced by gendered expectations and neighbourhood environments. Our results confirm that family values around co-residence and marriage are transmitted across generations (De Valk and Liefbroer 2007). We also find meaningful neighbourhood effects. Immigrant dense areas provide important support networks, but at the same time reproduce traditional expectations around staying at home (McAvay and Pailhé 2022). Crucially, our evidence shows that local unemployment has an even stronger influence than co-ethnic density. This indicates that structural constraints in the labour market may amplify the role of cultural

norms, making youth in disadvantaged neighbourhoods less able to leave home to live independently.

Women of MENA origin display larger gaps relative to native Swedes than men, and this is particularly visible in the second generation. This aligns with earlier research showing that young women in migrant families often face stricter expectations around family responsibility, modesty, and protection (Berry et al. 2006; Gebel & Heyne 2014). Remaining in the parental home until marriage may thus represent not only a cultural preference but also a form of constrained choice, influenced by familial expectations and broader community norms. For men, expectations may tilt in the opposite direction, with pressure to contribute financially before establishing their own household. Such gendered constraints help explain why integration in this domain is slower for women, even in a welfare-state context that otherwise supports early residential independence.

The contrast between the second generation and the 2.5 generation further highlights the importance of household-level socialization. Youth with one Swedish-born parent display home-leaving patterns much closer to those of natives, consistent with research linking intermarriage to faster assimilation (Tegunimataka 2021). This suggests that exposure to a mix of cultures within the household dilutes strict adherence to traditional norms and makes independent living a more feasible or acceptable option. By contrast, in households where both parents are foreign-born, the reinforcement of cultural expectations appears stronger, leading to delayed or foregone independent living.

Theoretically, these results challenge the view of assimilation as a unidirectional process that unfolds automatically across generations. Instead, they point toward a segmented trajectory in which cultural norms, structural barriers, and gender roles interact to produce heterogeneous outcomes. The fact that economic and parental resource controls do little to explain the observed gaps reinforces the conclusion that cultural and contextual mechanisms carry greater weight. This resonates with literature on segmented assimilation, which stresses that outcomes among the second generation depend critically on neighbourhood conditions, social capital, and the degree of discrimination and exclusion encountered (Portes & Rumbaut 2001).

From a policy perspective, our findings underline the importance of housing and neighbourhood contexts for integration trajectories. The Swedish welfare state has long been credited with equalizing opportunities for youth transitions, but our results show that local unemployment and residential segregation can sustain significant gaps in life-course milestones. Integration policy must address housing accessibility and the social environments in which immigrant youth grow up. Supporting affordable housing options for young adults, reducing residential segregation, and strengthening women's autonomy in immigrant communities are

crucial for ensuring that immigrant-origin youth have the same opportunities for independence as their native peers.

Finally, these results open several avenues for future research. One limitation of the competing risks framework is that it captures only the first transition out of the parental home. Future work could extend this by examining "boomerang" behaviour (Olofsson et al. 2020), serial residential moves, or the dynamics of cohabitation and partnership formation over time. Comparative studies across European welfare states could also shed light on whether the second-generation paradox we identify is unique to Sweden or part of a broader pattern in contexts with strong welfare support but persistent neighbourhood segregation. Further, qualitative work could illuminate how second-generation youth themselves perceive the tensions between family expectations, community norms, and personal aspirations regarding independent living.

References

Aradhya, S., Grotti, R., & Härkönen, J. (2023). Unemployment persistence among second-generation immigrants. *European Sociological Review, 39*(3), 433-448.

Aradhya, S., Hedefalk, F., Helgertz, J., & Scott, K. (2017). Region of origin: Settlement decisions of Turkish and Iranian immigrants in Sweden, 1968–2001. *Population, Space and Place, 23*(4), e2031.

Billari, F. C., & Liefbroer, A. C. (2007). Should I stay or should I go? The impact of age norms on leaving home. *Demography, 44*(1), 181-198.

Berry, J. W., Phinney, J. S., Sam, D. L., & Vedder, P. (Eds.). (2006). *Immigrant youth in cultural transition: Acculturation, identity, and adaptation across national contexts.* Lawrence Erlbaum Associates Publishers.

Borjas, G. J. (2006). Making It in America: Social Mobility in the Immigrant Population. *The Future of Children, 16*(2), 55–71.

Bowes, A., Dar, N., & Sim, D. (1997). Tenure preference and housing strategy: An exploration of Pakistani experiences. *Housing Studies, 12*(1), 63–84.

Buchmann, M. C., & Kriesi, I. (2011). Transition to adulthood in Europe. *Annual Review of Sociology, 37*(1), 481-503.

Carlsson, M., & Rooth, D-O. (2007). Evidence of ethnic discrimination in the Swedish labor market using experimental data. *Labour Economics, 14*(4), 716–729.

Carmichael, S. (2011). Marriage and power: Age at first marriage and spousal age gap in lesser developed countries. *The History of the Family, 16*(4), 416-436.

de Valk, H. A. G., & Liefbroer, A. C. (2007). Timing Preferences for Women's Family-Life Transitions: Intergenerational Transmission Among Migrants and Dutch. *Journal of Marriage and Family, 69*(1), 190–206.

Dribe, M., & Stanfors, M. (2005). Leaving home in post-war Sweden: A micro-level analysis of leaving the parental home in three birth cohorts. *Scandinavian Economic History Review, 53*(2), 30-49.

Dykstra, P. A., & Fokkema, T. (2012). Norms of filial obligation in the Netherlands. *Population, 67*(1), 97–122

Eurostat. (2024). *Age of leaving the parental home steady in 2023.* Eurostat News. Retrieved September 1, 2025, from https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20240930-2

Haan, M., Cheng, W., & Yu, Z. (2023). "Leaving Home Ain't Easy:" The timing and pathways of young immigrants' home-leaving transitions. *Canadian Review of Sociology/Revue Canadienne de Sociologie, 60*(2), 276-301.

Hannemann, T., & Kulu, H. (2015). Union formation and dissolution among immigrants and their descendants in the United Kingdom. *Demographic Research, 33,* 273–312.

Hoolachan, J., McKee, K., Moore, T., & Soaita, A. M. (2017). 'Generation rent' and the ability to 'settle down': economic and geographical variation in young people's housing transitions. *Journal of Youth Studies, 20*(1), 63–78.

Gebel, M., & Heyne, S. (2014). *Transitions to adulthood in the Middle East and North Africa: young women's rising?* Springer.

Gillespie, B. J., Bostean, G., & Malizia, S. (2020). Timing of Departure From the Parental Home: Differences by Immigrant Generation and Parents' Region of Origin. *Hispanic Journal of Behavioral Sciences, 42*(2), 165–190.

Glick, J. E., & van Hook, J. (2002). Parents' Coresidence with Adult Children: Can Immigration Explain Racial and Ethnic Variation? *Journal of Marriage and Family, 64*(1), 240–253.

Goldscheider, C., & Goldscheider, F. K. (1987). Moving out and marriage: What do young adults expect? *American Sociological Review, 52*(2), 278–285.

Gorodzeisky, A., & Semyonov, M. (2017). Labor force participation, unemployment and occupational attainment among immigrants in West European countries. *PLOS ONE, 12*(5), e0176856.

Hiekel, N., Liefbroer, A. C., & Poortman, A.-R. (2014). Understanding diversity in the meaning of cohabitation across Europe. *European Journal of Population, 30*(4), 391–410.

Huschek, D., de Valk, H. A. G., & Liefbroer, A. C. (2011). Does social embeddedness influence union formation choices among the Turkish and Moroccan second generation in the Netherlands? *Journal of Comparative Family Studies, 42*(6), 787–808.

Jonsson, J. O., & Mood, C. (2023) Karriärer och barriärer – En ESO-rapport om skolgång och etablering för unga med utländsk bakgrund. *Rapport till Expertgruppen för studier i offentlöig ekonomi, 2023:8*

Kleinepier, T., & Valk, H. A. G. de. (2017). Leaving Home of Migrant and Dutch Young Adults: Parent-Child and Peer Relations. *Journal of Comparative Family Studies, 48*(1), 15–39.

Koç, İ. (2007). The timing of leaving the parental home and its linkages to other life course events in Turkey. *Marriage & Family Review, 42*(1), 29-47.

Kovacheva, S., Kabaivanov, S., & Roberts, K. (2018). Interrogating waithood: Family and housing life stage transitions among young adults in North-West Africa countries. *International Journal of Adolescence and Youth, 23*(4), 441-456.

Marini, M. M. (1984). Age and sequencing norms in the transition to adulthood. *Social Forces, 63*(1), 229–244.

McAvay, H., & Pailhé, A. (2022). Leaving the nest in immigrant neighbourhoods: gender and origin differences in France. *Journal of Ethnic and Migration Studies, 48*(19), 4622–4647.

Melander, C., & Shmulyar Gréen, O. (2023). Language brokering as acts of care: experiences of young migrants born in Poland and Romania living in Sweden. *Nordic Social Work Research, 13*(4), 577–588.

Olofsson, J., Sandow, E., Findlay, A., & Malmberg, G. (2020). Boomerang behaviour and emerging adulthood: Moving back to the parental home and the parental neighbourhood in Sweden. *European Journal of Population, 36,* 919-945.

Pager, D., & Shepherd, H. (2008). The Sociology of Discrimination: Racial Discrimination in Employment, Housing, Credit, and Consumer Markets. *Annual Review of Sociology, 34*(1), 181–209.

Portes, A., & Rumbaut, R. G. (2001). *Legacies: The story of the immigrant second generation.* University of California Press.

Salikutluk, Z. (2016). Why Do Immigrant Students Aim High? Explaining the Aspiration–Achievement Paradox of Immigrants in Germany. *European Sociological Review, 32*(5), 581–592.

Sampson RJ (2012) *Great American City. Chicago and the Enduring Neighborhood Effect.* University of Chicago Press.

Sonneveld, N. (2019). Marriage and Divorce. In *Gender in the Middle East and North Africa* (pp. 51–72). Lynne Rienner Publishers.

Tegunimataka, A. (2021). The intergenerational effects of intermarriage. *Journal of International Migration and Integration, 22*(1), 311-332.

van Hook, J., & Glick, J. E. (2007). Immigration and living arrangements: Moving beyond economic need versus acculturation. *Demography, 44*(2), 225–249.

van Tubergen, F., & Maas, I. (2007). Ethnic intermarriage among immigrants in the Netherlands: An analysis of population data. *Social Science Research, 36*(3), 1065–1086.

Villanueva, C. M., & Buriel, R. (2010). Speaking on behalf of others: A qualitative study of the perceptions and feelings of adolescent Latina language brokers. *Journal of Social Issues, 66*(1), 197–210.

Wallace, M. (2022). Mortality advantage reversed: the causes of death driving all-cause mortality differentials between immigrants, the descendants of immigrants and ancestral natives in Sweden, 1997–2016. *European journal of population, 38*(5), 1213-1241.

Yalim, A. C., & Critelli, F. (2023). Gender roles among Syrian refugees in resettlement contexts: Revisiting empowerment of refugee women and needs of refugee men. *Women's Studies International Forum, 96,* Article 102670.

Zorlu, A. and Mulder, C.H., 2011. Ethnic differences in leaving home: Timing and pathways. *Demography, 48*(1), pp.49-72.

Appendix

Table A1: Full regression output with the transition into independent living defined as the main failure event

	(1) Males	(2) Females
Native Swedes	.ref	.ref
MENA_2.5	0.824***	0.695***
	(-6.81)	(-12.16)
MENA_2.0	0.511***	0.376***
	(-52.88)	(-62.45)
MENA_1.75	0.690***	0.505***
	(-19.79)	(-28.20)
MENA_1.5	0.657***	0.464***
	(-23.45)	(-24.81)
MENA_1.25	0.653***	0.458***
	(-15.10)	(-19.30)
MENA_1.0	0.670***	0.552***
	(-6.02)	(-9.32)
EEURO_2.5	0.846***	0.836***
	(-8.58)	(-9.28)
EEURO_2.0	0.701***	0.629***
	(-15.62)	(-18.65)
EEURO_1.75	0.770***	0.696***
	(-6.84)	(-8.54)
EEURO_1.5	0.648***	0.634***
	(-10.02)	(-9.22)
EEURO_1.25	0.606***	0.576***
	(-8.85)	(-9.05)
EEURO_1.0	0.934	0.634***
	(-0.33)	(-3.33)
Post secondary education	1.338***	1.374***
	(71.81)	(72.19)
Unemployed	0.686***	0.907***
	(-52.00)	(-5.09)
Unemployed mother	1.026***	1.066***
	(5.10)	(11.03)
Unemployed father	0.996	1.010
	(-0.86)	(1.93)
Occupation		
Farming, industrial, and transport workers	.ref	.ref
Managers and officials	1.054	0.995
	(1.60)	(-0.11)

Table A1 (continued)

	Males	Females
Military Personnel	1.030	1.127*
	(1.60)	(2.22)
Professionals	1.029***	1.008
	(4.28)	(0.81)
Service workers	0.948***	1.034***
	(-11.67)	(3.58)
Unknown	0.878***	0.995
	(-28.27)	(-0.49)
Mother's occupation		
Farming, industrial, and transport workers	.ref	.ref
Managers and officials	1.012	0.917***
	(1.17)	(-8.35)
Professionals	1.005	0.895***
	(0.60)	(-14.10)
Service workers	0.995	0.952***
	(-0.65)	(-6.76)
Unknown	1.054***	0.915***
	(3.80)	(-6.02)
Father's occupation		
Farming, industrial, and transport workers	.ref	.ref
Managers and officials	1.029***	0.945***
	(5.50)	(-10.77)
Professionals	1.027***	0.927***
	(6.71)	(-19.64)
Service workers	1.016***	0.966***
	(3.30)	(-7.13)
Unknown	0.975*	0.878***
	(-2.11)	(-10.62)
Income	1.000***	1.000***
	(5.98)	(11.98)
Mother's income	1.000***	1.000***
	(5.02)	(-7.02)
Father's income	1.000***	1.000***
	(6.93)	(-7.97)
County		
Lan=1	.ref	.ref
Lan=3	1.382***	1.456***
	(39.01)	(44.56)
Lan=4	1.239***	1.206***
	(20.57)	(16.83)
Lan=5	1.635***	1.539***
	(63.83)	(54.26)
Lan=6	1.136***	1.129***
	(14.74)	(13.25)
Lan=7	1.317***	1.347***
	(25.29)	(25.97)

Table A1 (continued)

	Males	Females
Lan=8	1.266***	1.281***
	(23.59)	(23.36)
Lan=9	0.907***	0.849***
	(-5.08)	(-8.04)
Lan=10	1.318***	1.181***
	(22.27)	(12.25)
Lan=12	1.179***	1.197***
	(30.84)	(32.69)
Lan=13	0.969***	0.944***
	(-3.48)	(-5.89)
Lan=14	1.146***	1.164***
	(28.00)	(29.76)
Lan=17	1.318***	1.340***
	(29.52)	(30.14)
Lan=18	1.409***	1.431***
	(37.03)	(37.79)
Lan=19	1.308***	1.269***
	(25.80)	(21.52)
Lan=20	1.105***	1.087***
	(10.40)	(8.09)
Lan=21	1.229***	1.224***
	(21.32)	(19.82)
Lan=22	1.151***	1.220***
	(14.04)	(18.50)
Lan=23	1.087***	1.192***
	(6.25)	(12.81)
Lan=24	1.519***	1.612***
	(47.68)	(54.28)
Lan=25	1.301***	1.275***
	(26.98)	(22.27)
Observations	3,342,797	2,624,073

Exponentiated coefficients; *t* statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A2: : Full regression output with the transition into marriage defined as the main failure event

	(1) Males	(2) Females
Native Swedes	.ref	.ref
GMENA_2.5	2.864***	4.202***
	(4.78)	(9.42)
GMENA_2.0	9.647***	11.97***
	(43.06)	(56.00)
GMENA_1.75	7.008***	9.831***
	(20.97)	(32.04)
GMENA_1.5	8.825***	11.33***
	(28.02)	(35.58)
GMENA_1.25	9.812***	16.54***
	(21.98)	(30.36)
GMENA_1.0	9.725***	8.477***
	(9.22)	(9.76)
EEURO_2.5	1.299	0.928
	(1.11)	(-0.31)
EEURO_2.0	2.067***	2.466***
	(3.64)	(5.72)
EEURO_1.75	5.068***	5.132***
	(7.06)	(8.10)
EEURO_1.5	8.381***	6.438***
	(10.66)	(9.57)
EEURO_1.25	3.767***	7.291***
	(3.95)	(8.33)
EEURO_1.0	5.224	2.264
	(1.69)	(0.81)
Post secondary education	1.752***	1.906***
	(13.69)	(18.45)
Unemployed	0.462***	1.031
	(-16.82)	(0.71)
Unemployed mother	1.348***	1.147**
	(5.87)	(2.97)
Unemployed father	0.997	1.010
	(-0.06)	(0.22)
Occupation		
Farming, industrial, and transport workers	.ref	.ref
Managers and officials	1.937*	2.589**
	(2.34)	(2.68)
Military Personnel	0.213**	3.67e-08***
	(-2.67)	(-131.79)
Professionals	1.571***	2.616***
	(6.18)	(8.41)
Service workers	0.903	1.284*
	(-1.85)	(2.36)
Unknown	0.644***	1.173
	(-7.84)	(1.49)

Table A2 (continued)

	Males	Females
Mother's occupation		
Farming, industrial, and transport workers	1	1
Managers and officials	0.869	1.085
	(-0.93)	(0.66)
Professionals	1.049	0.996
	(0.47)	(-0.05)
Service workers	1.091	1.193*
	(0.87)	(2.08)
Unknown	2.044***	2.149***
	(6.50)	(7.70)
Father's occupation		
Farming, industrial, and transport workers	.ref	.ref
Managers and officials	0.867	0.879*
	(-1.78)	(-1.97)
Professionals	0.932	0.922
	(-1.28)	(-1.81)
Service workers	1.169**	1.176***
	(2.65)	(3.49)
Unknown	1.899***	1.378***
	(9.68)	(5.06)
Income	1.000***	1.000***
	(8.12)	(12.88)
Mother's income	1.000	1.000
	(-0.65)	(-1.67)
Father's income	1.000	1.000
	(-0.51)	(-0.76)
County		
Lan=1	.ref	.ref
Lan=3	0.825*	0.728***
	(-2.14)	(-3.98)
Lan=4	0.856	0.924
	(-1.35)	(-0.82)
Lan=5	0.832*	0.752***
	(-2.14)	(-3.60)
Lan=6	1.692***	1.902***
	(7.25)	(10.75)
Lan=7	0.857	0.645***
	(-1.19)	(-3.36)
Lan=8	0.931	0.743*
	(-0.60)	(-2.54)
Lan=9	0.512*	0.458**
	(-2.20)	(-2.79)
Lan=10	0.638*	0.545***
	(-2.52)	(-3.38)
Lan=12	0.871*	0.858**
	(-2.34)	(-3.02)

Table A2 (continued)

	Males	Females
Lan=13	0.721**	0.812*
	(-2.77)	(-2.08)
Lan=14	1.016	1.044
	(0.32)	(1.03)
Lan=17	0.629***	0.581***
	(-3.54)	(-4.55)
Lan=18	0.969	0.899
	(-0.32)	(-1.27)
Lan=19	0.867	0.792*
	(-1.30)	(-2.30)
Lan=20	0.647***	0.679**
	(-3.30)	(-3.26)
Lan=21	0.740*	0.820
	(-2.44)	(-1.87)
Lan=22	0.887	0.732*
	(-0.96)	(-2.55)
Lan=23	0.652*	0.453***
	(-2.30)	(-4.09)
Lan=24	0.802*	0.606***
	(-2.01)	(-4.67)
Lan=25	0.999	1.050
	(-0.01)	(0.48)
Observations	3,342,797	2,624,073

Exponentiated coefficients; *t* statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A3: Coefficients of competing risks regression without control variables

	Male		**Female**	
	(1)		(2)	
Reference	Natives		Natives	
Origin	MENA	EEURO	MENA	EEURO
Generation				
2.5	0.766***	0.800***	0.674***	0.797***
	(-9.30)	(-11.47)	(-13.76)	(-11.96)
2	0.479***	0.644***	0.395***	0.606***
	(-62.01)	(-19.29)	(-67.29)	(-21.49)
1.75	0.637***	0.703***	0.514***	0.665***
	(-24.83)	(-9.25)	(-29.06)	(-9.62)
1.5	0.601***	0.600***	0.468***	0.589***
	(-29.38)	(-11.98)	(-33.99)	(-10.79)
1.25	0.589***	0.520***	0.418***	0.502***
	(-19.44)	(-11.67)	(-22.19)	(-11.15)
1	0.582***	0.833	0.474***	0.562***
	(-8.75)	(-0.87)	(-11.59)	(-4.87)
Controls				
Unemployment	N		N	
Post-secondary education	N		N	
Occupation	N		N	
Income	N		N	
Parental controls	N		N	
County FE's	N		N	
Observations	3,342,797		2,624,073	

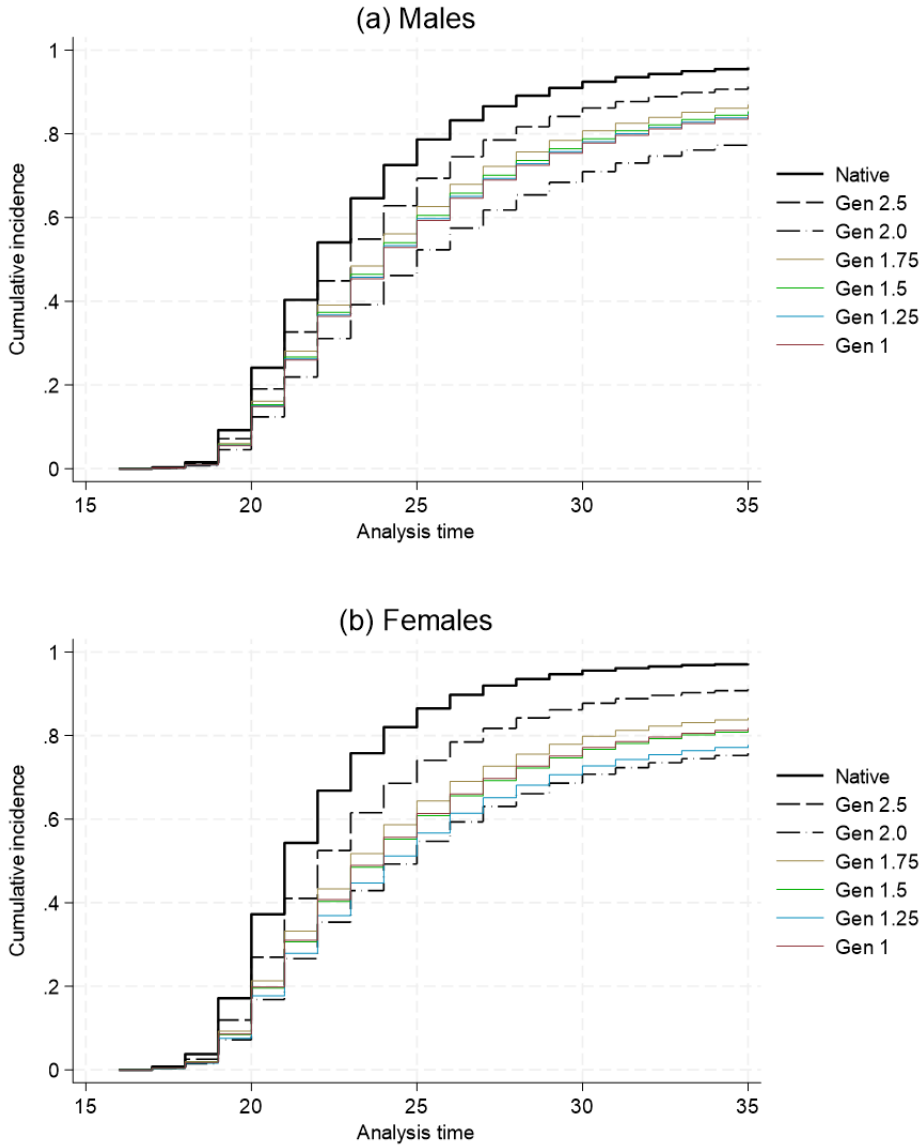
Note: The table reports the regression coefficients for the competing risks regression with the failure event defined as leaving the parental home to live independently and the competing risk event defined as leaving the parental home to live in a partnership.

Table A4: Coefficients of competing risks regression with commune fixed effects

	Male		Female	
	(1)		(2)	
Reference	Natives		Natives	
Origin	MENA	EEURO	MENA	EEURO
Generation				
2.5	0.732***	0.809***	0.623***	0.797***
	(-10.35)	(-10.13)	(-15.01)	(-10.98)
2	0.443***	0.631***	0.333***	0.556***
	(-60.05)	(-19.19)	(-67.65)	(-22.38)
1.75	0.575***	0.679***	0.435***	0.591***
	(-28.11)	(-9.64)	(-34.15)	(-11.56)
1.5	0.552***	0.586***	0.397***	0.567***
	(-31.59)	(-11.68)	(-33.59)	(-11.15)
1.25	0.569***	0.565***	0.405***	0.580***
	(-19.37)	(-9.58)	(-21.84)	(-8.41)
1	0.652***	0.513**	0.514***	0.557**
	(-5.47)	(-2.76)	(-8.71)	(-2.84)
Controls				
Unemployment	Y		Y	
Post-secondary education	Y		Y	
Occupation	Y		Y	
Income	Y		Y	
Parental controls	Y		Y	
Commune FE's	Y		Y	
Observations	3,385,303		2,657,346	

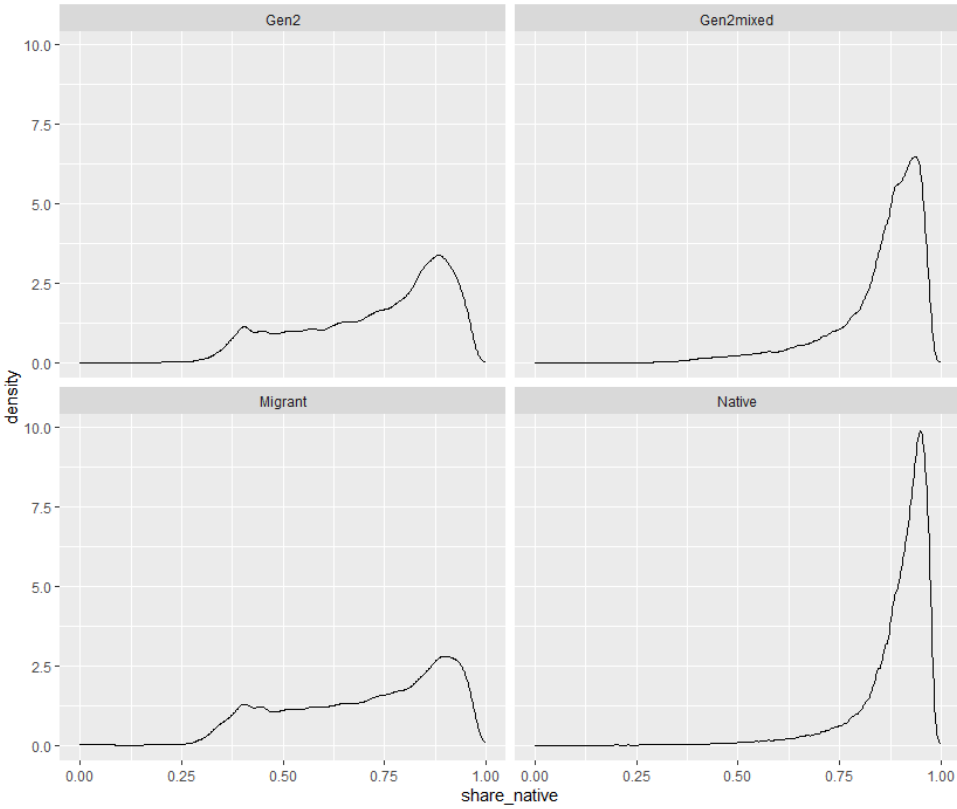
Note: The table reports the regression coefficients for the competing risks regression with the failure event defined as leaving the parental home to live independently and the competing risk event defined as leaving the parental home to live in a partnership.

Figure A1: Cumulative incidence functions (without parental controls)



Note: The cumulative incidence functions are based on the models controlling without parental controls.

Figure A2: Distribution of individuals in the sample according to the share of native Swedes in the same ZIP code



Note: The figures plot the distribution of individuals according to the share of native individuals in the local postcode, according to migrant status.

LUND STUDIES IN ECONOMIC HISTORY

123 Zalfou, Rami
Borders and Pathways: Essays on Syrian Economic Development and Migrant Integration, 2025.
122 Brolin, John
Studies in Earth-system economic history, 2025.
121 De Souza Maia, Vinicius
Familiar territory: neighborhood effects on union formation and fertility in Sweden 1905–2022, 2025.
120 Jayes, Jonathan
Like moths to a flame: an individual level approach to technological change in 20th century Sweden, 2025.
119 Dallier-Fortier, Clara
Studies on the effects of Mine Closures in Canada, 2025.
118 Falk, Marcus
Wealth, Consumption, and Industriousness: Evidence from southern Sweden, 1570–1860, 2025.
117 Cormack, Louise
Growing up. Early-life circumstances and wellbeing throughout life, Sweden 1905–2016, 2025.
116 Ternsjö, Linn
Garment Workers and the Labour Issue in Development: The Case of Mauritius, 2025.
115 Ford, Nicholas Martin
Origins of the knowledge economy: Higher education and Scandinavia's development, ca. 1800–1929, 2025.
114 Melles, Bertus Markus
Transforming Foundations: Structural Change and Development Dynamics in Sub-Saharan Africa (1960–2024), 2024.
113 Zhao, Liang
International Capital Flow, Monetary Policy, and Banking Instability: Lessons from Sweden, 1877–1939, 2024.
112 Castillo, Marcos
Roots Across the Atlantic: Social Networks and Economic Outcomes from the Swedish Mass Migration to the United States, 1880–1920, 2024.
111 Lukkari, Valeria
Economic disparities in colonial Kenya: income inequality and wage differentiation, 2024.
110 Ostermeyer, Vinzent Leon
Why Firms Grow: The Roles of Institutions, Trade, and Technology during Swedish Industrialization, 2023.
109 Bergvall, Martin
Untying the knot. Studies of the Swedish divorce transition, 1900–2020.
108 Dalman, Elien
Patterns of Persistence: Intergenerational mobility and Sweden's social structure 1865–2015, 2022.
107 Labbas, Elisa
Squeezed in Midlife: Studies of unpaid caregiving among working-age men and women across Europe, 2022.
106 Miladinovic, Luka
Us and Them: Economic Consequences of Ethno-confessional Diversity, 2022.

105 Baumert, Nicolai
For God's Sake – The Work and Long-Term Impact of Christian Missionaries in Cameroon 1844–2018, 2022.

104 Chanda Chiseni, Michael
Here I am Send Me: The Historical and Long-Term Impact of Christian Missionaries in Zambia 1924–2018, 2022.

103 Johansson, Tony
Sverige och den västeuropeiska arbetslösheten 1970–1990, 2021.

102 Gülsunar, Emrah
The State, Parliamentary Legislation and Economic Policy during the Structural Transformation of British Economy, 1700–1850, 2021.

101 Klocke, Sascha
Land, Labour, Legacies: Long-term Trends in Inequality and Living Standards in Tanzania, c. 1920–2020, 2021.

100 Önnerfors, Martin
Water for the Many. Health, neighbourhood change and equality of access during the expansion of Swedish urban water networks, 2021.

99 Rohne Till, Emelie
Agriculture for Development in the 21st Century: Evidence from Ethiopia, 2021.

98 Serratos-Sotelo, Luis
Disease, Death and Displacement: The long-term effects of early-life conditions on income, education, and health in Sweden, 1937–2011, 2021.

97 Martins, Igor
Collateral Effect: Slavery and Wealth in the Cape Colony, 2020.

96 Debiasi, Enrico
The Historical Origins of the Mortality Gradient. Socioeconomic inequalities in adult mortality over two centuries in Sweden, 2020.

95 Fibæk Maria
Rural Income Diversification, Employment, and Differentiation in Kenya and Implications for Rural Change, 2020.

94 Aboagye, Prince Young
Inequality and Uneven Development in Ghana, 2020.

93 Probst, Julius
Growth, Factor Shares, and Factor Prices, 2019.

92 Kulionis, Viktoras
Energy Embodied in Trade, 1970–2014, 2019.

91 Gary, Kathryn E.
Work, Wages and Income: Remuneration and Labor Patterns in Southern Sweden 1500–1850, 2018.

90 Chen, Zeyuan
Getting Old in a Changing World: Sons, Pensions, and the Wellbeing of the Elderly in Contemporary China, 2018.

89 Jenkin, Samuel
Foreign Investment in Agricultural Development: The Past of the Present in Zambia, 2018.

88 Karlsson, Omar
Growth and Survival: Child Health Disparities and Early Life Adversity in Sub-Saharan Africa, 2018.

87 Lennard, Jason
Economic Fluctuations in the United Kingdom, 1750–1938, 2018.

86 Andersson, Jens
State Capacity and Development in Francophone West Africa, 2018.

85 Aradhya, Siddartha
Diversity of Legacy: The Experience of Immigrants and their Descendants in Sweden, 2018.

84 Elwert, Annika
Will You Intermarry Me? Determinants and Consequences of Immigrant-Native Intermarriage in Contemporary Nordic Settings, 2018.

83 Lazuka, Volha
Defeating Disease: Lasting Effects of Public Health and Medical Breakthroughs Between 1880 and 1945 on Health and Income in Sweden, 2017.

82 Sztern, Sylvia
Russia on the Move: The Railroads and the Exodus from Compulsory Collectivism 1861–1914, 2017.

81 Tegunimataka, Anna
Trajectories of Integration: Naturalization, Intermarriage and Education in Denmark, 1980–2015, 2017.

80 Zheng, Yannu
How Immigrants Invent: Evidence from Sweden, 2017.

79 Nielsen, Hana
Coal, commerce and communism: Empirical studies on energy history in the Czech Republic, 2017.

78 Berger, Thor
Engines of Growth: Essays in Swedish Economic History, 2016.

77 Kenny, Seán
Empirical Studies in Money and Debt in Northern Europe 1840–2015, 2016.

76 Neilson, Jeffrey
Converging Times: Parenthood, Time Allocation and the Gender Division of Labor in Sweden 1990–2010, 2016.

75 Qi, Haodong
Live Longer, Work Longer? Evidence from Sweden's Ageing Population, 2016.

74 Lindström, Martin
All that's mine I carry with me: Early life disease and adult health in Sweden during 250 years, 2015.

73 Molitoris, Joseph
Life and Death in the City: Demography and living standards during Stockholm's industrialization, 2015.

72 Eriksson, Björn
Dynamic Decades: A micro perspective on late nineteenth century Sweden, 2015.

71 Persson, Mats R.
In good times and in bad: Immigrants, Self-employment and Social Insurance, 2015.

70 Borowiecki, Karol Jan
"Mournful and yet grand is the destiny of the artist." Some Economic History of Music Composers and their Creativity, 2015.

69 López Jerez, Montserrat
Deltas Apart. Factor Endowments, Colonial Extraction and Pathways of Agricultural Development in Vietnam, 2014.

68 Xiao, Jing
Entrepreneurial Dynamics and Acquisitions of New Technology-Based Firms, 2014.

67 Taalbi, Josef
Innovation as Creative Response. Determinants of Innovation in the Swedish Manufacturing Industry, 1970–2007, 2014.

66 Bergenfeldt, Fredrik
Roads to Market. Transport and Agricultural Change – the case of southern Sweden, 1750–1850, 2014.

65 Sjöö, Karolin
Innovation and Transformation in the Swedish Manufacturing Sector, 1970–2007, 2014.

64 Lind, Daniel
Value Creation and Structural Change During the Third Industrial Revolution. The Swedish Economy from a Vertical Perspective, 2014.

63 Silva, Ernesto
Essays on Income and Fertility. Sweden 1968–2009, 2014.

62 Dekhtyar, Serhiy
Early Life Origins of Health and Well-Being in Modern Sweden, 2013.

61 Holmberg, Kajsa
Behind the Counter. Female Inroads, Obstacles and Careers in the Swedish Commercial Bank Sector 1885–1937, 2013.

60 Palacio, Andrés
Mind the Gap. Inequalities in Infant and Child Mortality: the Case of Colombia 1967–2010, 2013.

59 Quaranta, Luciana
Scarred for life. How conditions in early life affect socioeconomic status, reproduction and mortality in southern Sweden, 1813–1968, 2013.

58 Hannemann, Tina
It Breaks a Man's Heart. Socioeconomic Differences in the Onset of Cardiovascular Disease in Contemporary Sweden, 2012.

57 Faran, Turaj S.
The Industrialization Process in the Chinese Mirror: Continuity and change in transition from collective agriculture to market economy in China, 2011.

56 Weisdorf, Jacob L.
Malthusian Progress, 2011.

55 Håkansson, Peter
Ungdomsarbetslösheten - om övergångsregimer, institutionell förändring och socialt kapital (Youth unemployment - transition regimes, institutional change and social capital), 2011.

54 Henriques, Sofia Teives
Energy transitions, economic growth and structural change. Portugal in a long-run comparative perspective, 2011.

53 Helgertz, Jonas
Immigrant Careers – Why Country of Origin Matters, 2010.

52 Bohman, Magnus
Bonden, bygden och bördigheten. Produktionsmönster och utvecklingsvägar under jordbruksomvandlingen i Skåne ca 1700–1870 (Peasants, production patterns and paths of development during the agricultural transformation in Southern Sweden (Scania) ca 1700–1870), 2010.

51 Waldenström, Daniel
Lifting All Boats? The Evolution of Income and Wealth Inequality over the Path of Development, 2009.

50 Roubert, Carl-Johan
Småländsk världsindustri - en studie om entreprenörer och kluster 1950–2002, 2009.

49 Örnberg, Lena
Bondens val - System, samspel och stödjande strukturer under moderniseringsprocessen i norra Thailand (Farmer's Choice - Systems, interaction and supporting structures during the modernisation process in Northern Thailand), 2008.

48 Rönnqvist, Sofia
*Från diversity management till mångfaldsplaner? Om mångfaldsidéns spridning i Sverige och Malmö (*From Diversity Management to Diversity Plans? On the Spread of the Diversity Idea in Sweden and Malmö), 2008.

47 Karlsson, Tobias
Downsizing: Personnel Reductions at the Swedish Tobacco Monopoly, 1915–1939, 2008.

46 Enflo, Kerstin
Regional and Industrial Growth Patterns in 20th Century Western Europe, 2008.

45 Axelsson, Tobias
Peasants and Policymakers, Agricultural transformation in Java under Suharto, 2008.

44 Gustafsson, Fredrik
The Visible Palm: Market Failures, Industrial Policy and the Malaysian Palm Oil Industry, 2007.

43 Forsberg, Le Thanh
Defining Strong Ownership: Institutional Determinants and Stakeholder Interests in Vietnamese Development Planning, 2007.

42 Andera, Jan
Driving under the Influence – Strategic Trade Policy and Market Integration in the European Car Industry, 2007.

41 Krantz, Olle & Schön, Lennart
Swedish Historical National Accounts 1800–2000, 2007.

40 Lundh Nilsson, Fay
Lönande lärande. Teknologisk förändring, yrkesskicklighet och lön i svensk verkstadsindustri omkring 1900 (Profitable Learning - Technological Change, Skills and Wages in the Swedish Engineering Industry around 1900), 2007.

39 Abou-Zeinab, Ali
Commercial Banking and Economic Development, The Lebanese Experience 1856–1974, 2006.

38 Jönsson, Per Ingvar
Chaebolföretagen – en"omöjlig" konfiguration i den sydkoreanska industrialiseringsprocessen (The Chaebol Conglomerate – An "Impossible" Configuration in the South Korean Industrialization Process), 2006.

37 Josephson, Camilla
Growth and Business Cycles – Swedish Manufacturing Industry 1952–2001, 2005.

36 Appelquist, Joakim
Informationsteknik och organisatorisk förändring – Teknik, organisation och produktivitet i svensk banksektor 1975–2003 (Information Technology and Organizational Change – Technology, Organization and Productivity in Swedish Banking Sector 1975–2003), 2005.

35 Green, Erik
Peasant production and limits to labour: Thyolo and Mzimba Districts in Malawi, mid-1930s to late-1970s, 2005.

34 Saarinen, Jani
Innovations and Industrial Performance in Finland 1945–1998, 2005.

33 Lindgren, Mattias
Does schooling make sense? A household perspective on the returns to schooling for self-employed, farmers and employees in Egypt, 2005.

32 Lundh, Christer, Jonas Olofsson, Lennart Schön & Lars Svensson (eds.)
Wage Formation, Labour Market Institutions and Economic Transformation in Sweden 1860–2000, 2004.

31 Rytkönen, Paulina
Fruits of Capitalism, Modernisation of Chilean Agriculture 1950–2000, 2004.

30 Johansson, Kent
Child Mortality during the Demographic Transition. A Longitudinal Analysis of a Rural Population in Southern Sweden, 1766–1894, 2004.

29 Holmquist, Svante
Kapitalbildning i svensk industri 1870–1930 (Capital Formation in Swedish Industry 1870–1930), 2004.

28 Carlsson, Ellen
To Have and to Hold: Continuity and change in property rights governing water resources among the Meru of Tanzania and the BaKgatla of Botswana; 1925–2000, 2003.

27 Abdulsomad, Kamaruding
Building technological capabilities of local auto parts firms under contrasting industrial policies: a comparative study of Malaysia and Thailand 1960–2000, 2003.

26 Bengtsson, Henrik
Ålder och ekonomisk omvandling. Sveriges tillväxthistoria i ett demografiskt perspektiv, 1890–1995 (Age and Economic Transformation. Sweden's History of Economic Growth in a Demographic Perspective, 1890–1995), 2003.

25 Andersson, Martin
Bending the Kuznets Curve – Wrenching and Levelling Forces during Agricultural Transformation in Taiwan and Chile, 2003.

24 Sandal, Jan-U.
Jakten på Entreprenøren. Kan Joseph A. Schumpeters teori benyttes til å identifisere og differensiere entreprenører i det 21.århundrede? (In Search of the Entrepreneur. Can Joseph A. Schumpeter's theory be used to identify and differentiate between entrepreneurs in the 21st century), 2003.

23 Öberg, Klas
Pharmacy Regulation in Sweden. A New Institutional Economic Perspective, 2003.

22 Stanfors, Maria
Education, Labor Force Participation and Changing Fertility Patterns. A Study of Women and Socioeconomic change in Twentieth Century Sweden, 2003.

21 Klinthäll, Martin
Return Migration from Sweden 1968–1996, 2003.

20 Olsson, Mats
Storgodsdrift. Godsekonomi och arbetsorganisation i Skåne från dansk tid till mitten av 1800-talet (Manorial Economy and Corvée Labour in Early Modern Scania), 2002.

19 Kander, Astrid
Economic growth, energy consumption and CO2 emissions in Sweden 1800–2000, 2002.

18 Rauhut, Daniel
Fattigvård, socialbidrag och synen på fattigdom i Sverige 1918–1997 (Poor Relief, Social Assistance and the View on Poverty in Sweden 1918–1997), 2002.

17 Schånberg, Ingela
Genus och utbildning. Ekonomisk-historiska studier i kvinnors utbildning ca 1870–1970 (Gender and Education. Economic History Studies in Women´s Education ca. 1870 to 1970), 2002.

16 Svensson, Patrick
Agrara entreprenörer. Böndernas roll i omvandlingen av jordbruket i Skåne ca 1800–1870 (Agrarian Entrepreneurs. The Role of the Peasants in the Transformation of Agriculture in Scania, 1800–1870), 2001.

15 Bevelander, Pieter *Immigrant Employment Integration and Structural Change in Sweden, 1970–1995*, 2000.

14 Lobell, Håkan
Växelkurspolitik och marknadsintegration. De utländska växelkurserna i Sverige 1834–1880 (Exchange Rate Policy and Market Integration. Foreign Exchange Rates in Sweden 1834–1880), 2000.

13 Dribe, Martin
Leaving Home in a Peasant Society. Economic Fluctuations, Household Dynamics and Youth Migration in Southern Sweden, 1829–1866, 2000.

12 Thorburn, Thomas
Economic of Transport: the Swedish Case 1780–1980, 2000.

11 Ankarloo, Daniel *"Institutions", What Is In a Word? A Critique of the New Institutional Economics*, 1999.

10 Iglesias, Edgar
Welfare, Security and the Level of Fertility: A Global Analysis, 1985 and 1995, 1999.

9 Scott, Kirk
The Immigrant Experience: Changing Employment and Income Patterns in Sweden, 1970–1993, 1999.

8 Siriprachai, Somboon
Control and Rent-Seeking, The Role of the State in the Thai Cassava Industry, 1998.

7 Pei, Xiaolin
The Institutional Root of China's Rural Industry and Gradual Reform, 1998.

6 Olofsson, Jonas
Arbetslöshetsfrågan i historisk belysning. En diskussion om arbetslöshet och social politik i Sverige 1830–1920 (The Unemployment Question in Historical Perspective. A Discussion on Unemployment and Social Policy in Sweden, 1830–1920), 1996.

5 Ahlström, Göran
Technological Development and Industrial Exhibitions 1850–1914. Sweden in an International Perspective, 1995.

4 Arlebäck, Sven Olof
Från ägarmakt under ansvar till företagsledarmakt upa. En studie över maktförskjutningen i tre svenska verkstadsaktiebolag (The Transition of Power from Shareholders to Top-Management. A Study of Three Swedish Engineering Firms), 1995.

3 Staffansson, Jan-Åke
Svenskt smör. Produktion, konsumtion och utrikeshandel 1861–1913 (Swedish Butter. Productivity, Consumption, and Foreign Trade, 1861–1913), 1995.

2 Bergström, Asta
Åtstramning och Expansion. Den ekonomiska politiken i Sverige 1971–1982 (Retrenchment and Expansion. Swedish Economic Policy 1971–1982), 1995.

1 Lundh, Christer (ed.)
Population, Economy and Welfare. Scandinavian Population Studies, vol 10, 1995.

0 Ahlström, Göran
Technological Development and Industrial Exhibitions 1850–1914. Sweden in an International Perspective, 1995.