

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

Rami Zalfou*

[Please click here for the latest version](#)

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 human and physical capital, and a higher share of workers in the primary sector. I find that these effects persist primarily due to lower population density, which explains a significant portion of the economic gaps. This persistence reflects path-dependent consequences: emigration and low population density, initiated by statelessness, continue to hinder economic development in these areas today.

Keywords: Stateless, Path Dependence, Spatial Regression Discontinuity, Ottoman Syria.

JEL codes: N15, N95, R12, O43.

* Department of Economic History, Centre for Economic Demography, and Centre for Advanced Middle Eastern Studies at Lund University. (email: rami.zalfou@ekh.lu.se). I am grateful to Martin Dribe and Anna Tegunimataka for their guidance and support, to Gunes Gokmen for detailed discussion and advice, and to participants at the Economic History Workshop of Developing Regions at Carlos III University of Madrid and at the Early-Stage Researcher Meeting at the Centre for Economic Demography at Lund University.

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, 1991, Bleakley and Lin 2012, 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.

This paper examines the long-term economic impact of historical statelessness, focusing on the 'desert line' in Syria — a border that historically divided areas governed by the Ottoman state from those that remained outside state control until the middle of the 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.

A regression discontinuity design is used to estimate the effect of a legacy of statelessness on economic development today, comparing Syrian towns and villages on two sides of the desert

line using contemporary census data. I use a rich array of village-level data from the Syrian population and housing census of 2004 to measure income, physical and human capital, and industrialization. The analysis finds that areas that were historically outside of state control have lower incomes and lower levels of human and physical capital, and a larger proportion of workers in agriculture. The results are not driven by geographic differences across the study region and hold under an array of geographic controls and alternative specifications, including spatial autoregressive lag models.

I find that population density is a strong predictor of the gaps in economic development across the border line, where density explains around half of the income and human capital gap, and most of the gap in infrastructure development. Furthermore, historical statelessness affects contemporary migration, with the effect detected in the age structure of the population due to a notable gap in the male working age population in historically stateless areas. Together, these results suggest an important role for path dependence. The effect of statelessness persists by having shaped population density and migration patterns across the border line, which in turn shape economic development. This evidence is in line with historical narratives that emphasize the effect that statelessness had on land abandonment and emigration from areas outside of Ottoman state control.

The analysis also tests for competing hypotheses. Using data on informality in home ownership, I detect an effect of historical statelessness on informality which points to weaker private property institutions in historically stateless areas. However, the gap in informality across the border line is a very weak predictor of economic development, explaining almost none of the economic gaps across the border line. Furthermore, I use data on village-level ethnic and religious composition (Khaddour and Mazur 2018) and show that the economic gaps are not due to ethnic or religious differences across the border line.

The paper contributes to the literature on the persistent effect of historical state institutions (Michalopoulos and Papaioannou 2013, Dell et al. 2018). Previous literature focuses on comparing regions with different degrees of state centralization in the past. This paper takes this idea further by testing for the effect of state absence on contemporary economic development outcomes. Previous research also focuses on the roles of institutions and culture as mechanisms of persistence of historical state capacity, while this paper takes a different approach by showing an important and overlooked role of path dependence as a mechanism of persistence, further contributing the literatures on the roles of population and path dependence in shaping economic development (Boserup 1965, Krugman 1991, Bleakley and Lin 2012). 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

The region of present-day Syria came to be part of the Ottoman Empire in 1516, having been previously ruled by the Mamluk Empire centred in Egypt. As Ma'oz (2013) describes, Syria under the Ottomans was not a unified political entity, but rather subdivided into provinces that were loosely governed by the central authority in Istanbul, with local autonomous forces developing within these provinces such as mountain chiefs, tribal sheikhs, and feudal overlords. The region extending from Aleppo to Damascus was the centre of Ottoman political power in Syria, whilst outlying regions like the Syria desert were inhabited by religious and ethnic minorities and formed centres of centrifugal political forces that resisted centralization. The Syrian desert, also known as the Syrian steppe or Badia, is primarily an arid steppe characterized by rocky and gravelly terrain with sparse vegetation including

shrubs and grasses. The region receives some rain during the winter season making it suitable for grazing and was used for this purpose by nomadic tribes.

Throughout Ottoman rule, the state faced the problem of de-population of its provinces. Population numbers 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 foreign 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. One of the measures taken by the government to alleviate the depopulation problem was through replacing tax farming tenancies with life leases to give tax farmers a long-term interest in the wellbeing of peasants, as well as through settlement projects.

The Ottoman state also carried out settlement projects across Anatolia and Syria with the aim of bringing more land into cultivation and to expand the tax base. These projects took on various form and included land grants to incoming refugees from former Ottoman territories in the Balkans and the Caucasus or involved forcible displacement of populations from within the Ottoman domain including nomadic populations such as those in the east of Syria and their sedenterization. Ma'oz (1968, Chapter 9) describes the settlement activity as sporadic, though it did move forward in the region around Aleppo where authorities encouraged settlement by granting land, seeds, and tools and tax exemptions. Though these efforts likely

faced difficulties as the government did not protect the newly sedentary tribes against attacks from nomads.

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. The Syrian desert is located further east of the line and formed the heartland of the nomadic tribes in Syria. Figure 1 also shows the cities, towns, and villages in the north-west of Syria. 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.

Figure 1 here

Archaeological evidence from the region east of the desert line suggests a degree of relative prosperity in the pre-modern era. The cities and towns on the Syrian Euphrates, which were largely ruined and abandoned by the time of Ottoman rule, had prospered from Antiquity well into the Middle Ages. The frontier of settlement in the study area moved significantly across time, reaching its furthest extent towards the east under Byzantine rule in the 5th and 6th centuries (Geyer 2011), a part of a cycle of intensification and abatement common to agricultural-pastoral frontier regions in the MENA (Wickham 2006, p. 19). The 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 east of Syria continued to play a significant role under the Islamic Caliphates in the following centuries. However, the entire area of present-day Syria was devastated by the Mongol invasions in the middle of the 13th century. In the present day, the only cities 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. The

district of Salamiyah was described by Robinson and Smith (1856) as ‘entirely deserted’ even though it was described as ‘exceeding even the neighbourhood of Homs and Hama, in the fertility of its soil.’

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 was responsible for 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 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 north-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 for re-defining 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 centralizations. Certain social classes such as urban elites and merchants played a 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 Ottoman modernization of land administration during the Hamidian period (1876-1909).

Several factors led to the Ottoman state modernization and reforms of the 19th century, and they mostly had to do with 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.

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 treasury¹) 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 in the north of western North-Africa usually 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

¹ The Arabic word 'Makhzan' translates to 'warehouse' but came to be denote the state in western North-Africa. The origin of this use may relate to the word's association with the treasury and the presence of state tax collectors in areas under state control.

according to Lindner (2017). The economic consequences of statelessness in the Middle East and North Africa (MENA) region remain understudied.

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 GCC countries, as well as some historical 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 stable governance.

3. Theory and Previous Research

A line of research focuses on the role of pre-colonial state history and institutions in explaining the economic inequalities observed the present. Studies of this kind measure state centralization in the past and find that it correlates with present day economic outcomes both across and within countries. Gennaioli and Rainer (2007) find that pre-colonial state

centralization in Africa at the country level is associated with better provision of public goods in the present. Their evidence suggests that the presence of more hierarchical chiefdom structures in the past causes local chiefs today to be more accountable to state authorities, leading to better implementation of state policies. Michalopoulos and Papaioannou (2013) extend this work using more granular data at the ethnic group level and find strong association between historical state centralization and economic development as proxied by nightlight density. More recently, Dell et al. (2018) tests the effect of historical state centralization in Vietnam across a historical border line that separated the centralized Dai Viet kingdom from the less centralized south, which was a peripheral tributary of the Khmer empire, finding that state centralization is associated with better economic outcomes, with village governance institutions playing a key role in explaining this persistent legacy.

Overall, these studies treat historical borders as natural experiments to test the effectiveness of state centralization in the past on economic development today. The mechanism suggested for this persistent relationship is mainly institutional, represented by local governance institutions such as chiefly hierarchy or village governance. Lowes et al (2017) uses a similar design to argue that past state centralization in Central Africa is related to cultural norms in the present, leading to weaker norms of rule following and a greater propensity to cheat as measured using behavioral experiments. While the study does not deal directly with economic development outcomes, it suggests another mechanism for persistence through cultural transmission.

A factor that has not been addressed in previous studies is that of economic and human geography. Historical evidence suggests that areas east of the desert line experienced very high emigration rates, with the abandonment of towns and villages becoming a common response to insecurity (Lewis 1987, Williams 1981). If differences in state centralization persisted for a long period, we would expect the emergence of large differences in population density between state-controlled and stateless regions in favour of state-controlled regions.

Differences in settlement patterns can persist long after differences in state centralization have disappeared. Path dependence can play an important role, and a large literature explores how initial conditions, and historical accidents can shape economic geography in the long run.

Krugman (1991) illustrates the effect of path dependence using the fact that one third of the US population today lives within the original thirteen colonies established by the British. Bleakley and Lin (2012) show that historical portage sites on rivers in the eastern US still determine the population distribution in that region, even though portage sites became obsolete long ago.

The effect of historical state centralization is not limited to shaping population. Population density is an important factor in economic development. Densely settled areas may experience urbanization at a higher rate and benefit from opportunities for specialization and investment. Boserup (1965) theorized that agricultural productivity is shaped by population density, with areas of higher density shifting to more intensive systems of land use. According to Kremer (1993), societies with larger initial populations have faster technological growth as well as faster population growth, all else being equal. Klasen and Nestmann (2006) incorporate both population size and population density in their analysis of the relationship between population and technological growth. According to their model, population density spurs technological change, particularly for countries with low levels of technology. Higher density also increases the returns to investment in public goods such as infrastructure, which in turn may affect technological change.

The link between population and economic development is a natural complement to the study of the legacy of state centralization, and that of the historical roots of economic development more broadly (see Nunn (2020) for a review). The analysis in this paper also relates to the study of state capacity and the consequences of weak rule of law and a lack of monopoly on violence for long term economic development (e.g. North et al. 2009).

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. The data is aggregated at the level of the city, town, and village. Information on the historical boundary of the desert line comes from the work of Lewis (1987) which uses traveller accounts to pinpoint the boundaries of state control prior to the mid-19th century.

The study 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:

$$outcome_i = \alpha + \gamma South_east_i + f(geographic\ location_i) + \beta to_Aleppo_i + \epsilon_i$$

Where $outcome_i$ is the outcome variable in village i , and $South_east_i$ 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 kilometers 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 results 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 geographic functions $f()$.

The estimation framework is also sensitive to selective sorting across the boundary. For this reason, the regression model additionally 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. Figure 2 shows a map of the study area according to ethnic group. 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.

Figure 2 here

Another threat to the estimation framework 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:

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

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:

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

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.

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 regular employment (as opposed to seasonal or intermittent work). 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.

5. Results

5.1. Descriptive statistics

Table 1 shows descriptive statistics for the main variables used in the analysis. The area of study is split into two, with towns and villages categorized into those that are located on the stateless side of the desert line and which were stateless until the middle of the 19th century, and those in the north-west which were under Ottoman state control throughout. The income indices measure at -1.26 and -0.80 standard deviations 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.

Table 1 here

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%.

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 respectively, 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.

Table 2 presents balance checks for the two sides of the border line using a set of geographic variables. Differences in slope, elevation, average yearly temperature, precipitation, and flow accumulation are presented alongside their mean values for the entire study region. The slope, presented as a percentage, is low overall which corresponds to a flat terrain on both sides of the border line, albeit the slope value is slightly higher in areas with a stateless legacy. The region sits on an elevation, with a relatively small difference in elevation between the two sides of the border line. Small differences are also found in terms of temperature and precipitation (with the latter favouring the stateless area), while the difference in flow accumulation is not statistically significantly different from zero.

Table 2 here

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). The columns show the RD effect of being located in the region with a stateless legacy, for different outcome measures. 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

Table 3 here

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 picture 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 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

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.

both considerable variation in the outcome variables across space, but also capture the discontinuity in outcomes across the border line.

Figure 3 here

To further test 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 here

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 may not be very 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 here

Panel B of Table 6 presents placebo outcomes tested using linear models in longitude and latitude. The placebo variables include the unemployment rate, female labour force participation (FLFP), the sex ratio in the local population, characteristics of dwellings in the local area (the proportions of the dwellings in refurbishment, rented dwellings, and empty dwellings), and the gender gap in educational attainment. The causal effects on the placebo outcomes are statistically insignificant (except for the proportion of empty dwellings, where the effect size is very small and statistically significant only at the 0.1 level). Overall, the placebo tests support the validity of the causal estimates found in the prior section, with no effects found for outcomes not directly measuring economic development.

Table 6 here

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 here

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 social 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 analyze 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 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.

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 here

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 here

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 here

To explore the role of population density, Table 11 shows the regression discontinuity effects on economic development outcomes, controlling for the population density measures. Comparing the regression discontinuity effects here to those in the unadjusted model (Table 3, Panel B), we find that the inclusion of the 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, perhaps due to government policy.

Table 11 here

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, through long-term 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 should disproportionately 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 relative age structure for males, and the ratio of males to females by age group, according to area. Measures for the study sample are included alongside those for cities of over 10,000 people (which are excluded from the analysis), and alongside all other areas in Syria outside of the study region. The relative age structure is calculated as the share of the males in each age group out of the entire male population, divided by the share of the same age group out of the entire population (both male and female). We find that for the population aged under 15 years, historically stateless and Ottoman areas show the same share (1.023 in the study sample) and a similar share in the comparison samples. When looking at the share of males in working age, we find a 0.109 gap in favour of historically Ottoman areas. The effect is only detectable in the study sample and cannot be seen in larger cities. A smaller but opposite effect is found in the relative share of the male population aged 65 and over. When looking at the ratio of males to females, a smaller effect is found in the same direction of 0.03. We can note that the ratio of males to females drops below 1 only in the working age group in historically stateless areas in the study sample, though it measures above 1 everywhere else.

Table 12 here

Using the age structure variables as dependent variables in Table 13 show negative and statistically significant effects of the regression discontinuity on the relative share of males aged 15 to 64 and the ratio of males to females in the same age group. For males under the age of 15, the effect is in the opposite direction, which may be due to worse infant survival favouring males in historically stateless areas. Therefore, the effect that historical statelessness has on the working age male share is even large if we take into account the ‘positive’ effect it has on the male share below age 15. Similar effects yield when excluding areas with majority non-Sunni Muslim populations and adding a dummy variable for tribal populations (See Appendix Table A9).

Table 13 here

7. Discussion and conclusion

The results confirm the presence of a negative effect of a legacy of statelessness on contemporary economic development in Syria. The finding supports earlier research on the topic. The results provide evidence of significant discrepancies in economic development on the two sides of the desert line. Areas exposed for a longer period to Ottoman state control have better economic outcomes in the present era, and the effects are found across different measures including income and physical and human capital. Towns and villages which experienced stable state rule in the western side of the desert line have higher average incomes, more developed physical infrastructure, and higher levels of education and a higher proportion of workers employed in white collar jobs as measured in the census data. The observed effects on economic development are robust to geographic confounding factors. The similar geography across the two sides of the historical border allows for clean identification of the causal effect of stateless legacy holding geography constant. The results are also robust for controlling for geographic factors and for various geographic robustness tests.

The effect of a legacy of statelessness is not likely to be transmitted through cultural norms. The causal effect is slightly larger when we control for ethnic and religious composition, and largest when we exclude non-Sunni Muslim groups. The effect is not driven by the gap between tribal and non-tribal communities in Syria either. The results of the analysis are unlikely to be simply a function of regional inequality during or before the Ottoman era. The infrastructure measures capture modern infrastructure that was constructed in the mid to late 20th century, while literacy and primary education spread in Syria in the course of the 20th century as well, well after the gap in state presence had been closed.

The analysis shows that the gap in population density across the border line is an important correlate of the gap in economic development. Rural insecurity prior to state expansion in the 19th century led to emigration into state-controlled areas west of the border line, with long term implications for the population distribution in the region. Low population density in areas with a legacy of statelessness explains around half the observed effects on income and education levels, and most of the effect on infrastructure variables and on the proportion of farmers in the labour force.

While the study finds evidence for a large negative effect of a legacy of statelessness on economic development today, the size of the effect may be conditional on government policy choices. The infrastructure effects disappear when we exclude tribal areas, which may suggest a tribal-specific penalty in infrastructure development projects or may point to an institutional mechanism even though the lack of data does not allow for analysis of local governance institutions. At the same time, the relatively smaller effects on the electricity network and the more developed status of this network on both sides of the border line reflects the Syrian state's prioritization of rural electrification after 1970 (Hinnebusch 2012).

The conflict between the ways of life of settled agrarian communities and nomadic groups is a recurring theme in the history of the MENA region and continues to influence the politics of

many developing countries today, particularly in Sub-Saharan Africa. In a world of climate change, increased conflict is expected in such frontier zones where nomadism survives, with recent research highlighting how worsening climatic conditions contribute to increased conflict by increasing pressures on pastoralist populations (McGuirk and Nunn 2024, Bai and Kung 2011). The analysis here shows that such conflict can have long term consequences for economic development, as well as for human settlement patterns and urbanization more broadly.

References

- Abu-Ismail, K., Abdel-Gadir, A. and El-Laithy, H., 2016. Poverty and Inequality in Syria.
- Baerlocher, D., Silva, D.F.C.D., Lambais, G., Reis, E. and Veras, H., 2024. Old But Gold: Historical Pathways and Path Dependence. Available at SSRN 4513384.
- Bai, Y. and Kung, J.K.S., 2011. Climate shocks and Sino-nomadic conflict. *Review of Economics and Statistics*, 93(3), pp.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.
- Bleakley, H. and Lin, J., 2012. Portage and path dependence. *The quarterly journal of economics*, 127(2), pp.587-644.
- Boserup, E., 2014. The conditions of agricultural growth: The economics of agrarian change under population pressure. Routledge.
- Dalgaard, C.J., Kaarsen, N., Olsson, O. and Selya, P., 2022. Roman roads to prosperity: Persistence and non-persistence of public infrastructure. *Journal of Comparative Economics*, 50(4), pp.896-916.
- Dell, M., 2010. The persistent effects of Peru's mining mita. *Econometrica*, 78(6), pp.1863-1903.

- Dell, M., Lane, N. and Querubin, P., 2018. The historical state, local collective action, and economic development in Vietnam. *Econometrica*, 86(6), pp.2083-2121.
- Esmer, T.U., 2014. Economies of violence, banditry and governance in the Ottoman empire around 1800. *Past & Present*, 224(1), pp.163-199.
- Gennaioli, N. and Rainer, I., 2007. The modern impact of precolonial centralization in Africa. *Journal of Economic Growth*, 12, pp.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), pp.7-22.
- Hinnebusch, R., 2012. Syria: from 'authoritarian upgrading' to revolution? *International Affairs*, 88(1), pp.95-113.
- Hoffman, B.G., 1967. The Structure of Traditional Moroccan Rural Society.
- Khaddour, K. and Mazur, K, 2018, Syria town database,
<https://doi.org/10.7910/DVN/YQQ07L>, Harvard Dataverse, V1,
 UNF:6:HGbxgRKL5n8ycbLwi4dSxw== [fileUNF]
- Khoury, P.S. and Kostiner, J. eds., 1990. Tribes and state formation in the Middle East. Univ of California Press.
- Klasen, S. and Nestmann, T., 2006. Population, population density and technological change. *Journal of Population Economics*, 19, pp.611-626.
- Kremer, M., 1993. Population growth and technological change: One million BC to 1990. *The quarterly journal of economics*, 108(3), pp.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), pp.62-96.
- Lindner, R.P., 2017. Nomads and Ottomans in medieval Anatolia. Routledge.
- Lowes, S., Nunn, N., Robinson, J.A. and Weigel, J.L., 2017. The evolution of culture and institutions: Evidence from the Kuba Kingdom. *Econometrica*, 85(4), pp.1065-1091.
- Ma'oz, M., 2013. The emergence of modern Syria. In *Syria Under Assad (RLE Syria)* (pp. 9-35). Routledge.

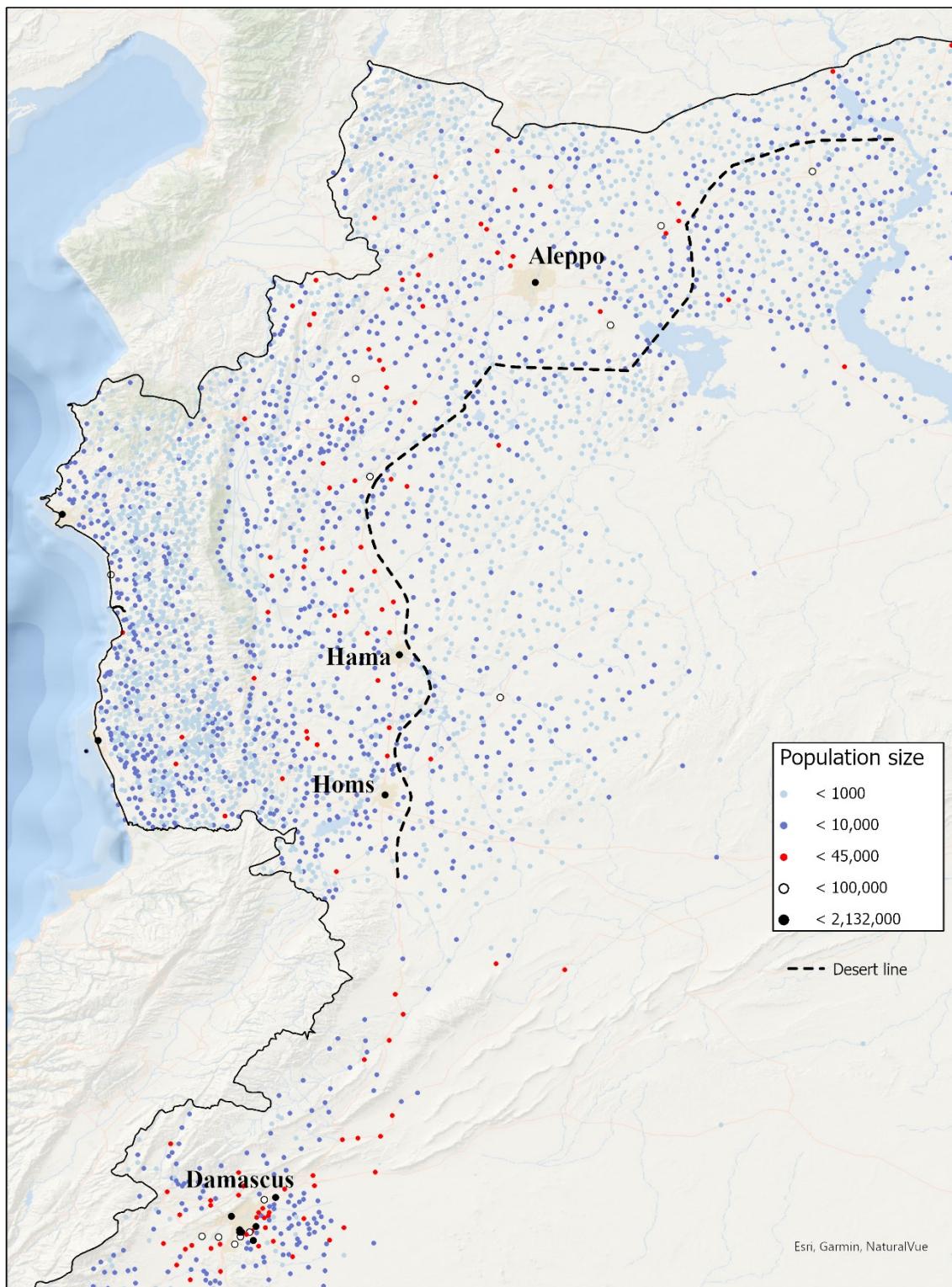
- Ma'oz, M., 1968. Ottoman reform in Syria and Palestine, 1840-1861: The impact of the Tanzimat on politics and society. Clarendon Press.
- McGuirk, E.F. and Nunn, N., 2024. Transhumant pastoralism, climate change, and conflict in Africa. *Review of Economic Studies*, p.rdae027.
- Michalopoulos, S. and Papaioannou, E., 2013. Pre-colonial ethnic institutions and contemporary African development. *Econometrica*, 81(1), pp.113-152.
- Norman N. Lewis, Nomads and Settlers in Syria and Jordan, 1800–1980, Cambridge Middle East Library (Cambridge: Cambridge University Press, 1987).
- North, D.C., Wallis, J.J. and 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), p.eaaz9986.
- Reinkowski, M. (2001). [Review of the book *Frontiers of the State in the Late Ottoman Empire: Transjordan, 1850-1921*, by E. L. Rogan]. *Die Welt des Islams*, New Series, 41(2), 266-268.
- 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). *Histoire économique et sociale de l'Empire ottoman et de la Turquie (1326-1960): actes du sixième congrès international tenu à Aix-en-Provence du 1er au 4 juillet 1992*, 8, p.251.
- 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), pp.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), pp.161-176.
- Wickham, C., 2006. Framing the early middle ages: Europe and the Mediterranean, 400-800. OUP Oxford.

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

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

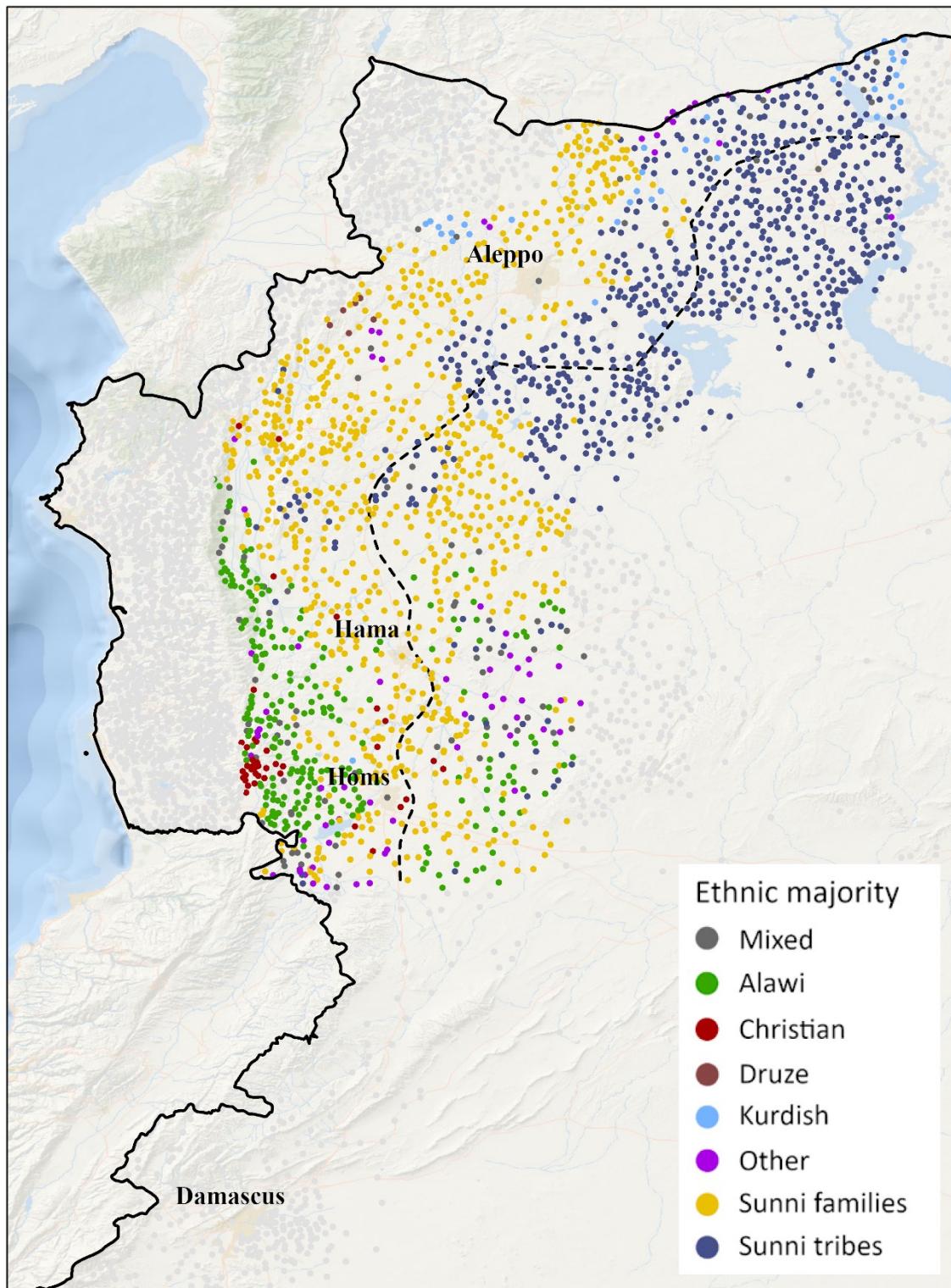
Figures and Tables

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.

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.

Figure 3: Regression discontinuity graphs

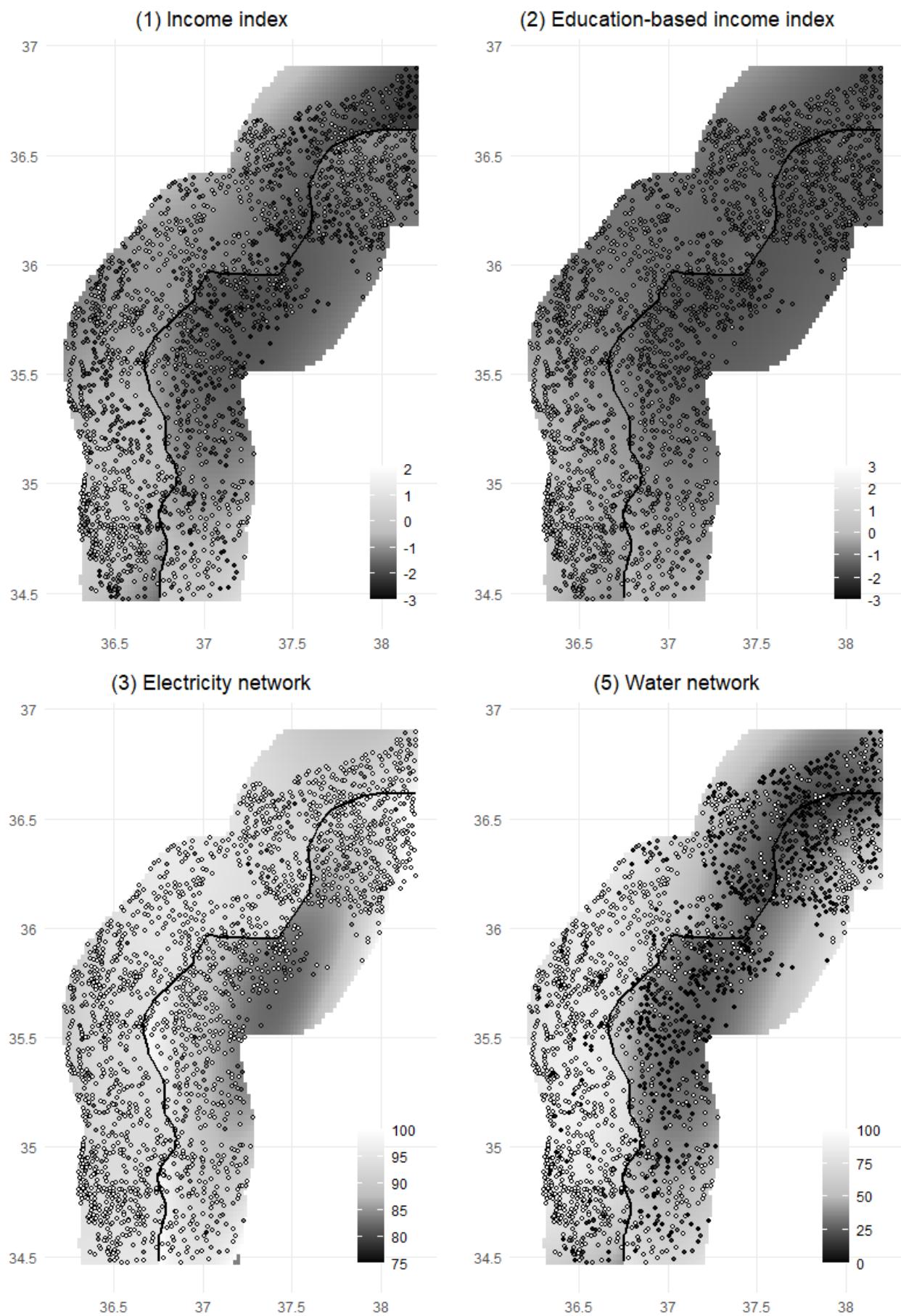


Figure 3 (continued)

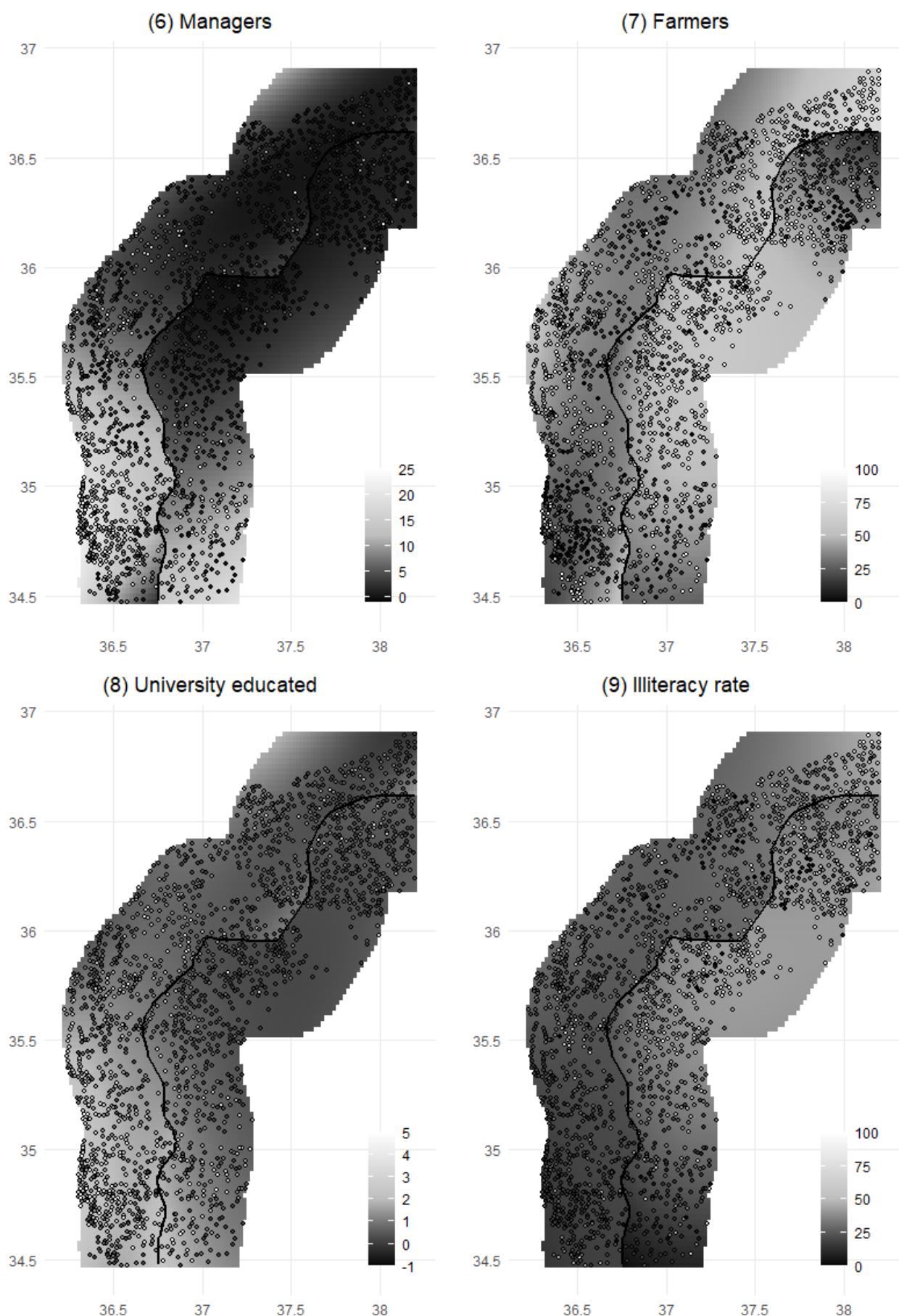


Table 1: Descriptive statistics

	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).

Table 2: Balance checks

<i>Dependent variable:</i>					
	Slope	Elevation	Temperature	Precipitation	Flow Accum.
	(1)	(2)	(3)	(4)	(5)
Stateless	0.718** (0.300)	-38.759*** (10.754)	0.125*** (0.045)	4.127** (1.795)	139.573 (233.060)
Mean	2.6	434.8	17.5	171.8	190.928

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

Note: The table shows the RD cutoff effect using a cubic polynomial in latitude and longitude for each variable, alongside the unadjusted mean value for the entire sample.

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

Table 4: Spatial autoregressive lag models

	<i>Dependent variable:</i>								
	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)
Stateless	-0.211** (0.103)	-0.191*** (0.066)	-6.144** (2.460)	-2.441 (3.885)	-13.954*** (4.032)	-1.625* (0.895)	3.751 (2.819)	-0.283** (0.143)	8.189*** (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

Note: The above table shows the RD cutoff effect for the spatial autoregressive lag models.

Table 5: Moran's I statistics

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

Table 6: RD effects using placebo outcomes

	<i>Dependent variable:</i>						
	Unemployment	FLFP	Sex ratio	Dwellings in refurbishment	Rented dwellings	Empty dwellings	Gender gap in schooling
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Stateless	0.845 (0.974)	-1.232 (1.091)	-0.014 (0.011)	-0.132 (0.441)	-0.224 (0.571)	1.270* (0.753)	0.223 (0.365)
Mean Y (control)	12.855	15.837	1.054	4.839	65.831	1.540	9.624
SD	15.271	16.701	0.192	6.108	39.261	6.669	10.437
Observations	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 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.

Table 7: RD effects according to ethnic group

	<i>Dependent variable:</i>								
	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)
<i>Panel A: Sunni Muslims</i>									
Stateless	-0.624*** (0.124)	-0.563*** (0.068)	-9.876*** (3.101)	-5.631 (4.603)	-28.045*** (4.807)	-3.885*** (0.642)	10.260*** (3.923)	-0.806*** (0.152)	14.511*** (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
<i>Panel B: Sunni Muslims excluding tribes</i>									
Stateless	-1.151*** (0.386)	-0.799*** (0.219)	14.797 (9.144)	-13.331 (14.423)	-5.628 (15.667)	-2.220 (2.283)	23.947** (11.190)	-1.360*** (0.459)	14.057** (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
<i>Panel C: Sunni tribes</i>									
Stateless	-0.132 (0.364)	-0.128 (0.160)	-7.481 (7.317)	-13.781 (13.355)	1.739 (17.025)	-1.425 (1.217)	3.485 (11.791)	-0.322 (0.326)	4.318 (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.

Table 8: RD effects on informality in property ownership

	<i>Dependent variable:</i>					
	Official deed		Agricultural deed		Public notary	
	(1)	(2)	(3)	(4)	(5)	(6)
Stateless	-8.376*** (2.169)	-6.710*** (2.197)	5.235** (2.314)	4.754** (2.392)	-0.431 (0.351)	-0.388 (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.

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

	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)
Stateless	-0.584*** (0.113)	-0.441*** (0.066)	-7.974*** (2.581)	-3.476 (3.794)	-28.096*** (4.194)	-3.628*** (0.800)	10.353*** (3.435)	-0.500*** (0.154)	13.889*** (1.938)
Official deed	0.002** (0.001)	0.001*** (0.0005)	0.150*** (0.016)	0.294*** (0.026)	0.210*** (0.027)	-0.00002 (0.007)	-0.039** (0.020)	0.005*** (0.001)	-0.010 (0.011)
Agricultural deed	0.001 (0.001)	-0.001* (0.0004)	0.126*** (0.017)	0.028 (0.021)	0.036 (0.026)	0.001 (0.006)	0.001 (0.021)	-0.0002 (0.001)	0.042*** (0.012)
Public notary	0.001 (0.002)	0.008*** (0.002)	0.194*** (0.042)	0.778*** (0.171)	0.665*** (0.113)	0.029* (0.015)	-0.017 (0.102)	0.007* (0.004)	-0.213*** (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.

Table 10: RD effect on population density

<i>Dependent variable:</i>						
	Total population		Population within 5 Km diameter		Population within 10 Km diameter	
	(1)	(2)	(3)	(4)	(5)	(6)
Stateless	-1,360.842*** (171.637)	-1,409.517*** (173.711)	-17,343.250*** (3,760.468)	-18,178.850*** (3,983.053)	-99,027.280*** (11,083.590)	-99,859.590*** (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.

Table 11: RD effects controlling for population density

	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)
Stateless	-0.297** (0.122)	-0.190*** (0.070)	0.592 (2.835)	19.027*** (4.225)	-7.949* (4.582)	-3.483*** (0.928)	2.837 (3.765)	-0.236 (0.155)	6.398*** (2.083)
ln total population	0.181*** (0.025)	0.152*** (0.016)	4.057*** (0.596)	11.728*** (0.844)	8.433*** (0.868)	1.013 *** (0.232)	-4.841 *** (0.699)	0.214 *** (0.035)	-2.754 *** (0.420)
ln population within 5 Km	0.064*** (0.018)	0.030** (0.013)	0.898* (0.462)	1.106 (0.763)	1.351* (0.753)	0.348 *** (0.134)	-1.806 *** (0.519)	0.014 (0.024)	-1.143 *** (0.309)
ln population within 10 Km	0.008 (0.034)	0.030 (0.026)	1.989*** (0.729)	4.888*** (1.286)	5.799*** (1.315)	-0.881 ** (0.344)	0.014 (1.056)	-0.004 (0.067)	-1.993 *** (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).

Table 12: Relative age structure and sex ratio

		<i>Relative age structure</i>					
		Under 15		15 to 64		65 and over	
		(Stateless)	(Ottoman)	(Stateless)	(Ottoman)	(Stateless)	(Ottoman)
Study sample		1.023	1.023	1.218	1.327	1.160	1.096
Cities over 10,000		1.003	1.003	1.360	1.373	1.090	1.026
Outside study area		0.999		1.371		1.034	

		<i>Ratio of males to females</i>					
		Under 15		15 to 64		65 and over	
		(Stateless)	(Ottoman)	(Stateless)	(Ottoman)	(Stateless)	(Ottoman)
Study sample		1.110	1.081	0.998	1.027	1.674	1.412
Cities over 10,000		1.061	1.063	1.041	1.066	1.318	1.141
Outside study area		1.063		1.0625		1.162	

Note: The table shows the proportion of each age group within the male population divided by the proportion of the age group in the total population (both males and females). The ratio is shown for the full analytical sample (study sample), and for two comparison samples: the sample of cities in the study region with a population over 10,000 people, and the sample of all Syria cities, towns, and villages outside the study area.

Table 13: RD effects on age structure in the male population and the sex ratio according to age

	<i>Dependent variable:</i>					
	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	0.013*** (0.003)	-0.156*** (0.024)	0.013 (0.015)	0.026** (0.011)	-0.048** (0.021)	0.079 (0.068)
Mean Y (Control)	1.023	1.327	1.096	1.081	1.027	1.412
SD	0.050	0.504	0.202	0.153	0.311	0.832
Observations	1,872	1,872	1,856	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.

Online Appendix

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

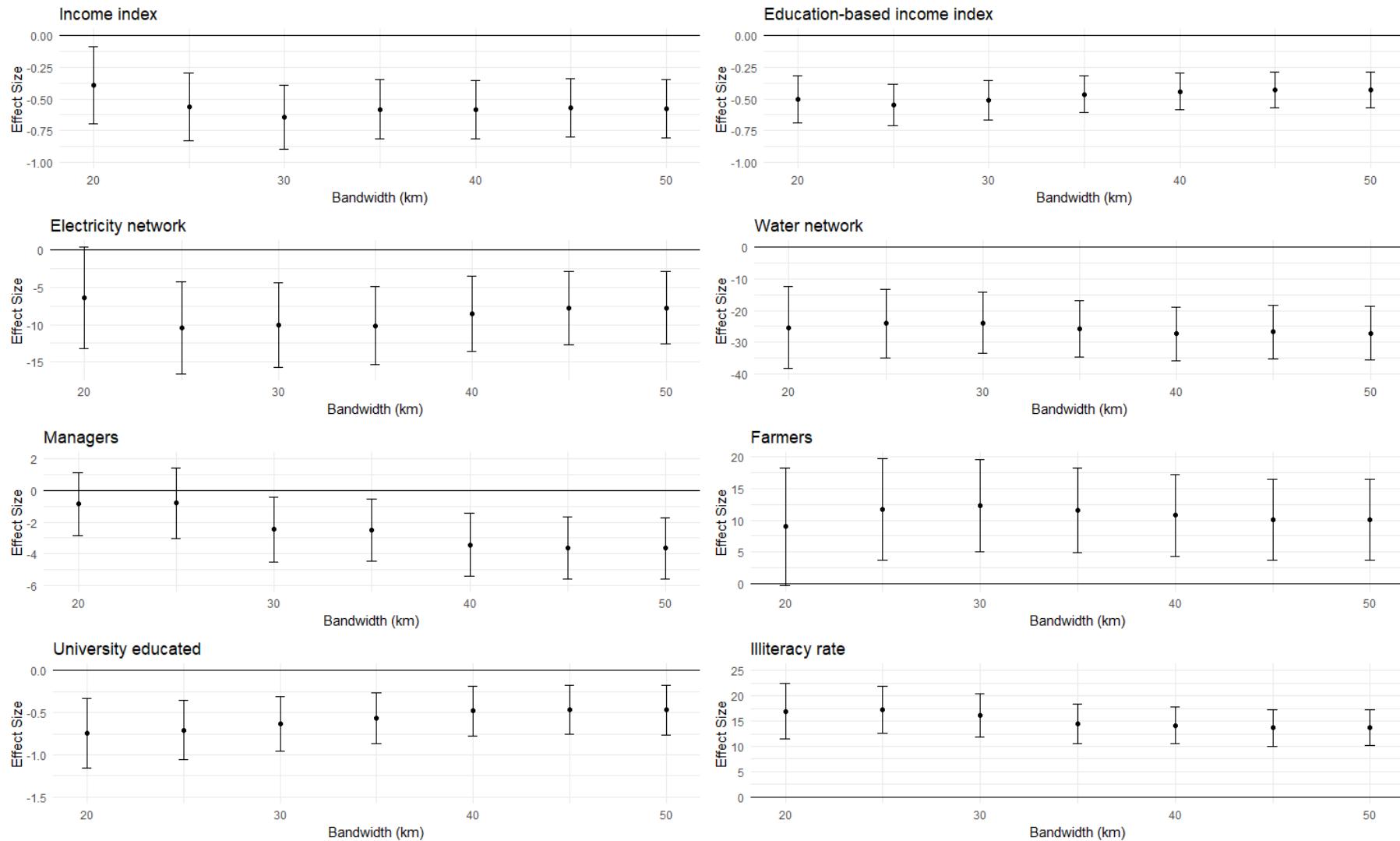


Figure A2: Bandwidth sensitivity with a 5k donut

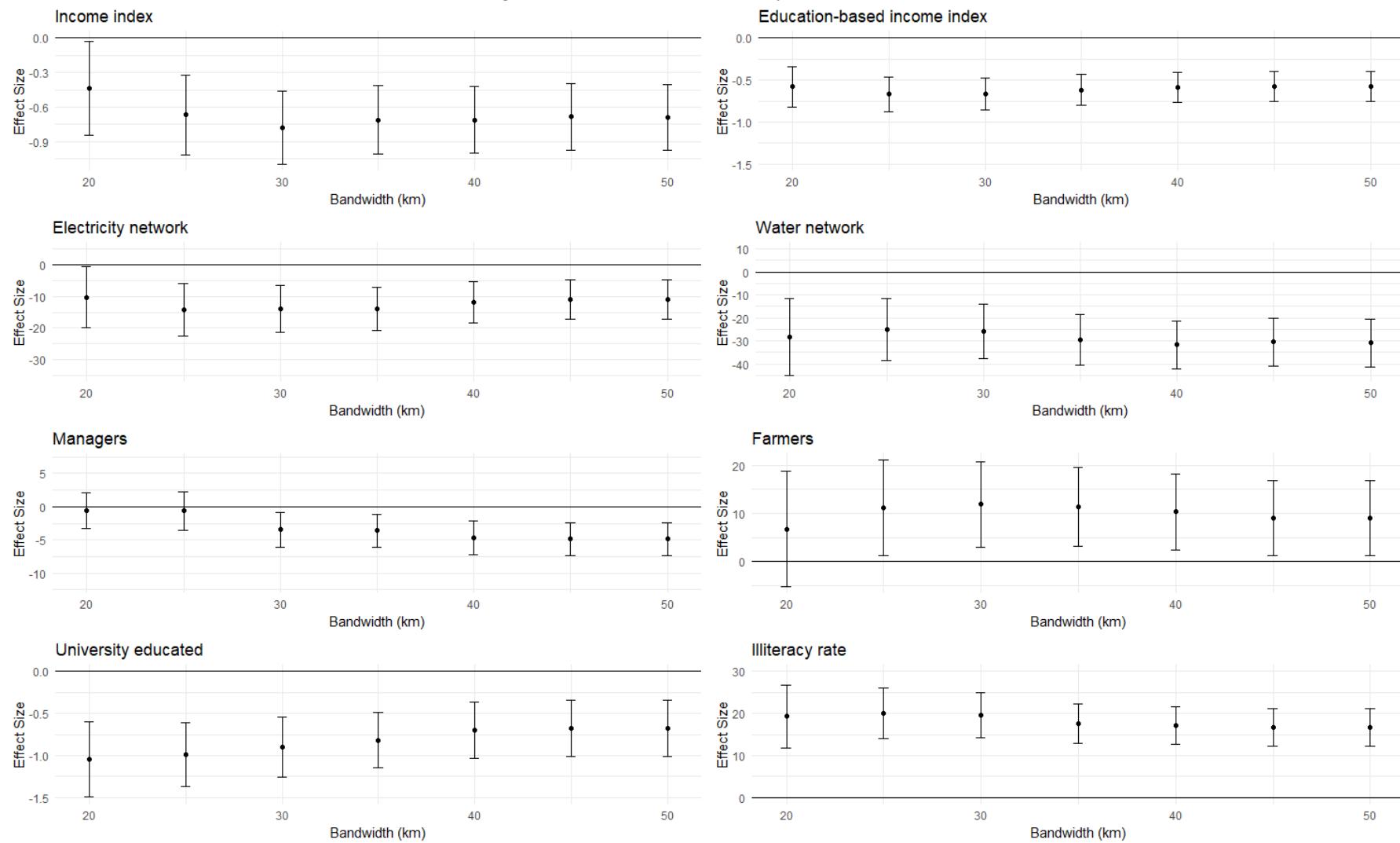


Table A1: Regression discontinuity effects using 25km bands

		<i>Dependent variable:</i>							
	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)
<i>Panel A: Linear model</i>									
Stateless	-0.321*** (0.064)	-0.333*** (0.040)	-3.706*** (1.398)	-15.950*** (2.210)	-17.691*** (2.665)	-3.984*** (0.464)	2.503 (1.962)	-0.417*** (0.087)	9.216*** (1.039)
<i>Panel B: Cubic polynomial model</i>									
Stateless	-0.563*** (0.143)	-0.548*** (0.080)	-10.448*** (3.232)	-5.398 (5.168)	-24.111*** (5.401)	-0.794 (0.933)	11.722*** (4.406)	-0.707*** (0.180)	17.231*** (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)
<i>Panel A: Linear model</i>									
Stateless	-1.382*** (0.224)	-0.718*** (0.123)	-7.427* (4.119)	-14.025* (7.704)	-39.782*** (9.287)	-3.009** (1.485)	35.223*** (6.493)	-1.306*** (0.236)	14.230*** (3.040)
<i>Panel B: Cubic polynomial model</i>									
Stateless	-1.121*** (0.418)	-0.600*** (0.230)	14.324 (9.319)	-3.535 (14.545)	4.686 (16.443)	-0.716 (2.349)	26.420** (11.860)	-1.146** (0.494)	6.927 (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*** (0.061)	-0.438*** (0.037)	-3.410** (1.326)	-21.942*** (2.047)	-22.491*** (2.529)	-5.010*** (0.407)	7.098*** (1.853)	-0.530*** (0.075)	9.216*** (1.039)
log(to_Aleppo)	0.064 (0.051)	0.263*** (0.025)	1.897 (1.153)	-4.460** (1.963)	4.832** (2.105)	2.511*** (0.237)	1.820 (1.658)	0.291*** (0.044)	0.559 (1.005)
Stateless:long	0.608*** (0.146)	0.097 (0.089)	2.143 (2.506)	14.725*** (4.333)	19.644*** (6.004)	4.324*** (0.960)	-16.727*** (4.240)	-0.082 (0.133)	-3.884 (2.430)
Stateless:lat	-0.781*** (0.135)	-0.504*** (0.078)	-0.152 (2.133)	-22.324*** (3.823)	-17.730*** (4.688)	-7.478*** (1.080)	6.335* (3.551)	-0.323*** (0.115)	13.796*** (1.853)
long:Control	-0.517*** (0.097)	-0.341*** (0.062)	-0.849 (1.837)	2.815 (3.456)	-45.193*** (3.145)	-2.290*** (0.837)	9.300*** (2.633)	-0.223 (0.148)	4.748*** (1.304)
lat:Control	-0.357*** (0.094)	-0.320*** (0.062)	1.638 (1.791)	-21.362*** (3.303)	8.828*** (3.069)	-5.116*** (0.888)	6.159** (2.453)	-0.579*** (0.156)	6.489*** (1.234)
Constant	-1.186*** (0.230)	-1.936*** (0.114)	85.048*** (5.110)	55.066*** (8.812)	35.440*** (9.353)	-3.640*** (1.085)	33.210*** (7.375)	-0.100 (0.215)	23.611*** (4.457)
Observations	1,908	1,908	1,908	1,908	1,908	1,908	1,908	1,908	1,908
R ²	0.200	0.409	0.007	0.124	0.250	0.310	0.062	0.199	0.209
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
F Statistic	78.981***	219.409***	2.260**	44.746***	105.553***	142.336***	20.765***	78.576***	83.555***

Note:

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

	(0.980)	(0.781)	(15.680)	(35.423)	(33.707)	(10.670)	(28.749)	(2.446)	(14.450)
I(lat2):Control	0.694 (0.469)	0.803* (0.415)	-3.888 (7.754)	45.299*** (17.504)	7.173 (16.662)	5.527 (5.431)	6.523 (12.305)	2.383* (1.353)	-4.571 (6.102)
I(long3):Control	0.686 (0.921)	1.284* (0.682)	23.837 (16.348)	45.496 (34.803)	56.780* (31.143)	0.755 (8.970)	-10.192 (26.977)	4.842** (2.053)	-11.681 (13.388)
Control:I(lat3)	1.862*** (0.477)	0.703* (0.359)	-7.802 (8.996)	45.668*** (17.267)	39.235*** (14.567)	20.397*** (4.747)	-20.401* (11.868)	0.409 (1.071)	-0.834 (5.660)
Stateless:long:lat	-0.648 (1.115)	0.256 (0.693)	-4.485 (20.732)	78.115** (36.527)	-106.757*** (40.930)	-2.458 (8.245)	33.977 (29.094)	1.926* (0.998)	10.024 (15.637)
Stateless:lat:I(long2)	-2.476 (2.890)	-1.295 (1.880)	-70.232 (49.050)	-156.345* (84.778)	-7.889 (104.594)	-3.646 (19.450)	-48.651 (75.564)	-2.935 (2.451)	1.445 (48.397)
Stateless:long:I(lat2)	2.321 (1.451)	0.927 (0.921)	22.072 (23.895)	43.891 (42.777)	-41.618 (51.037)	8.830 (11.162)	-6.409 (36.929)	0.738 (1.326)	-0.976 (22.467)
long:lat:Control	0.170 (1.310)	-1.821 (1.122)	-8.093 (22.552)	-46.805 (48.078)	-4.616 (45.988)	-6.889 (14.862)	-62.188* (36.884)	-6.849* (3.612)	8.908 (18.924)
lat:I(long2):Control	0.017 (2.126)	-3.424** (1.736)	-29.093 (38.033)	-28.718 (77.756)	-4.773 (66.692)	0.033 (24.159)	-34.567 (56.631)	-11.416** (5.527)	29.946 (27.615)
long:I(lat2):Control	-1.759 (1.508)	0.902 (1.309)	4.612 (26.844)	-60.545 (53.739)	-27.705 (43.422)	-20.505 (18.363)	46.328 (36.137)	5.004 (4.313)	-12.402 (16.195)
Constant	0.569 (0.454)	-0.835*** (0.245)	68.959*** (10.468)	92.621*** (18.701)	68.507*** (18.851)	16.131*** (2.409)	15.063 (14.329)	0.922** (0.420)	21.981** (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
F Statistic	37.008***	83.303***	3.119***	20.661***	41.500***	54.143***	15.519***	29.898***	33.664***

Note:

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

Stateless:lat	-0.007 (0.346)	-0.110 (0.202)	26.177*** (7.438)	-23.109* (12.814)	20.637 (14.537)	-5.846*** (2.229)	-12.982 (10.125)	0.249 (0.366)	-6.712 (5.686)
Stateless:I(long2)	4.089*** (0.993)	1.785*** (0.550)	12.068 (23.437)	-37.138 (31.591)	172.891*** (40.980)	10.542** (4.690)	-103.059*** (31.119)	0.298 (0.885)	-40.040** (18.322)
Stateless:I(lat2)	1.252*** (0.358)	0.558*** (0.206)	18.982*** (7.068)	9.473 (10.736)	43.458*** (12.832)	4.270 (2.771)	-26.260*** (9.339)	0.514 (0.314)	-22.104*** (4.663)
Stateless:I(long3)	-0.943 (1.529)	-0.303 (0.931)	37.241 (32.728)	87.194* (47.702)	-41.753 (61.267)	1.046 (8.142)	61.276 (46.004)	1.364 (1.349)	21.677 (28.378)
long:Control	-1.127*** (0.389)	-0.164 (0.261)	3.941 (8.253)	-11.906 (15.368)	-70.013*** (15.625)	1.432 (2.959)	44.910*** (11.940)	0.511 (0.747)	2.803 (7.307)
lat:Control	-0.616 (0.382)	-0.514** (0.246)	14.839** (7.537)	-28.179* (14.733)	-7.312 (13.724)	- (2.904)	-15.054 (10.422)	-1.152* (0.600)	0.712 (5.119)
I(long2):Control	-1.097 (0.941)	0.753 (0.721)	-5.134 (16.563)	35.502 (35.549)	-22.775 (34.067)	6.729 (10.151)	68.851** (27.657)	3.451 (2.313)	8.594 (14.466)
I(lat2):Control	0.306 (0.464)	0.401 (0.380)	-10.575 (8.330)	38.115** (17.439)	-1.785 (16.870)	3.209 (5.202)	13.678 (12.252)	1.603 (1.275)	1.862 (6.290)
I(long3):Control	0.458 (0.909)	0.984 (0.653)	15.380 (17.369)	35.523 (34.670)	47.777 (31.475)	0.975 (8.705)	-5.278 (26.442)	4.104** (2.013)	-5.099 (13.650)
Control:I(lat3)	1.273*** (0.467)	0.127 (0.348)	-10.770 (9.403)	42.466** (17.262)	30.516** (14.813)	15.028*** (4.492)	-11.666 (11.741)	-0.420 (1.019)	5.748 (5.471)
Stateless:long:lat	-0.723 (1.103)	0.276 (0.669)	-15.534 (20.893)	68.430* (37.247)	-118.121*** (42.357)	1.651 (8.317)	39.843 (29.381)	1.732* (0.992)	15.930 (16.054)
Stateless:lat:I(long2)	-2.976 (2.843)	-1.858 (1.827)	-54.139 (49.207)	-143.214* (85.862)	5.727 (106.240)	-17.365 (19.563)	-48.190 (76.424)	-3.310 (2.413)	-3.323 (48.582)
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
R ²	0.326	0.515	0.041	0.191	0.304	0.422	0.172	0.314	0.272
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
F Statistic	34.984***	76.688***	3.129***	17.087***	31.642***	52.889***	15.071***	33.117***	26.965***

Note:

* 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*** (0.112)	-0.409*** (0.067)	-7.338*** (2.673)	-0.903 (4.080)	-26.805*** (4.313)	-3.589*** (0.806)	7.743** (3.379)	-0.421*** (0.157)	13.541*** (1.959)
log(to_Aleppo)	-0.490*** (0.125)	-0.055 (0.068)	4.725* (2.809)	-23.100*** (5.009)	-3.729 (5.101)	-1.779** (0.728)	9.849** (3.860)	-0.136 (0.131)	0.165 (2.367)
Slope	-0.018*** (0.005)	-0.009*** (0.003)	0.104 (0.087)	-0.688*** (0.178)	-0.375** (0.172)	-0.078 (0.052)	0.475*** (0.128)	-0.001 (0.008)	0.325*** (0.093)
Elevation	0.001* (0.0003)	0.0003* (0.0002)	0.004 (0.006)	-0.013 (0.009)	-0.003 (0.009)	0.001 (0.003)	-0.009 (0.007)	0.001*** (0.0004)	0.002 (0.004)
Temperature	-0.354*** (0.078)	-0.052 (0.049)	-0.393 (1.889)	-1.274 (2.834)	-0.192 (2.680)	-1.707** (0.690)	11.347*** (1.917)	0.017 (0.098)	0.591 (1.115)
Precipitation	0.001 (0.001)	0.001 (0.001)	-0.073*** (0.026)	-0.218*** (0.045)	-0.035 (0.040)	0.061*** (0.014)	0.044 (0.032)	-0.002 (0.002)	-0.027* (0.015)
Flow accumulation	-0.00001** (0.00000)	0.00000 (0.00000)	0.00003 (0.0001)	-0.00003 (0.0002)	0.0001 (0.0002)	-0.00005 (0.00004)	0.0002** (0.0001)	-0.00000 (0.00001)	-0.00002 (0.0001)
Stateless:long_d	-0.751*** (0.289)	-0.547*** (0.185)	-31.095*** (6.651)	-10.987 (10.124)	-43.756*** (12.152)	7.412*** (1.831)	22.488*** (8.531)	-0.990*** (0.294)	17.227*** (5.114)
Stateless:lat_d	-0.909*** (0.336)	-0.680*** (0.202)	20.556*** (7.527)	-23.125* (12.531)	11.294 (13.996)	-14.023*** (2.177)	4.788 (9.808)	-0.398 (0.339)	2.667 (5.579)
Stateless:I(long_d2)	1.921* (1.039)	1.335** (0.615)	10.368 (23.421)	-12.633 (34.136)	173.235*** (42.726)	-1.267 (5.466)	-42.243 (31.991)	-0.196 (1.063)	-34.851* (18.364)

Stateless:I(lat_d2)	0.202 (0.388)	0.222 (0.235)	15.520** (7.655)	14.597 (11.714)	41.007*** (13.527)	-1.586 (3.037)	2.193 (9.824)	0.072 (0.349)	-17.876*** (4.923)
Stateless:I(long_d3)	0.987 (1.559)	-0.138 (0.981)	47.430 (32.642)	75.874 (48.635)	-39.012 (61.571)	4.029 (8.612)	-0.072 (46.025)	1.900 (1.480)	18.839 (28.336)
long_d:Control	-0.427 (0.411)	0.408 (0.295)	-4.559 (8.055)	-23.642 (15.817)	-73.894*** (15.539)	10.151*** (3.427)	32.243*** (12.219)	1.330* (0.804)	1.093 (7.472)
lat_d:Control	-1.814*** (0.398)	-1.250*** (0.276)	12.085 (7.785)	-29.157** (14.873)	-13.984 (13.923)	-22.635*** (3.309)	10.179 (10.553)	-2.237*** (0.645)	7.832 (5.436)
I(long_d2):Control	-1.243 (1.193)	1.449 (0.891)	33.451* (19.601)	152.109*** (41.577)	10.985 (39.731)	-17.403 (12.167)	63.976* (33.986)	6.913*** (2.638)	8.188 (17.231)
I(lat_d2):Control	-0.015 (0.528)	0.556 (0.445)	5.721 (9.151)	78.501*** (19.496)	13.698 (18.878)	-6.697 (5.822)	20.087 (13.825)	2.693* (1.401)	-0.600 (7.127)
I(long_d3):Control	-0.487 (0.984)	0.810 (0.707)	38.228** (17.179)	92.847** (36.612)	64.121* (33.133)	-17.573* (9.303)	11.999 (29.064)	5.076** (2.085)	-4.182 (14.543)
Control:I(lat_d3)	2.714*** (0.479)	1.009*** (0.368)	-11.762 (9.618)	36.207** (17.655)	39.776*** (15.376)	27.821*** (4.983)	-41.251*** (11.794)	0.522 (1.084)	-5.653 (6.227)
Stateless:long_d:lat_d	0.922 (1.120)	0.686 (0.700)	-1.337 (21.574)	73.452* (37.493)	-107.873*** (41.795)	4.354 (8.321)	-9.536 (29.451)	2.667*** (1.026)	8.469 (15.911)
Stateless:lat_d:I(long_d2)	-4.370 (2.869)	-1.777 (1.879)	-79.273 (49.795)	-158.128* (85.222)	-6.621 (105.066)	-8.285 (19.548)	6.184 (75.038)	-4.139* (2.465)	0.232 (48.470)
Stateless:long_d:I(lat_d2)	2.882** (1.438)	1.019 (0.919)	28.647 (24.360)	57.094 (43.019)	-39.893 (51.325)	6.978 (11.194)	-25.715 (36.469)	1.212 (1.338)	0.519 (22.506)
long_d:lat_d:Control	1.375 (1.449)	-1.340 (1.193)	-33.811 (25.392)	-125.850** (52.107)	-18.904 (50.663)	21.178 (15.684)	-80.308** (40.550)	-7.892** (3.731)	-1.483 (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** (1.738)	-0.219 (1.416)	44.061 (31.061)	53.082 (61.185)	-11.693 (52.032)	-66.670*** (19.828)	94.641** (42.741)	6.009 (4.489)	9.388 (21.026)
Constant	7.009*** (1.592)	-0.135 (0.992)	90.924** (40.893)	163.611*** (58.041)	79.725 (56.421)	35.155** (13.712)	-204.279*** (40.404)	0.896 (1.893)	16.027 (24.634)
Observations	1,907	1,907	1,907	1,907	1,907	1,907	1,907	1,907	1,907
R ²	0.297	0.461	0.036	0.186	0.296	0.369	0.166	0.239	0.258
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
F Statistic	33.109***	67.075***	2.898***	17.951***	32.933***	45.953***	15.626***	24.627***	27.302***

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

Table A7: RD effects on informality in property ownership

	<i>Dependent variable:</i>					
	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

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*** (0.001)	0.002*** (0.0005)	0.150*** (0.015)	0.329*** (0.026)	0.215*** (0.027)	0.007 (0.007)	-0.067*** (0.020)	0.005*** (0.001)	-0.031*** (0.011)
Agricultural deed	0.002** (0.001)	-0.0001 (0.0004)	0.117*** (0.015)	0.015 (0.020)	-0.029 (0.027)	0.002 (0.006)	-0.022 (0.021)	-0.0001 (0.001)	0.009 (0.012)
Public notary	0.002 (0.003)	0.009*** (0.002)	0.188*** (0.038)	0.835*** (0.172)	0.619*** (0.114)	0.027 (0.018)	-0.028 (0.104)	0.008** (0.004)	-0.252*** (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 in the male population and the sex ratio according to age (Sunni Muslim sample)

<i>Dependent variable:</i>						
	Males under 15 (1)	Males 15 to 64 (2)	Males 65 and over (3)	Males/females under 15 (4)	Males/females 15 to 64 (5)	Males/females 65 and over (6)
Stateless	0.013*** (0.004)	-0.059*** (0.022)	0.009 (0.017)	0.025** (0.012)	-0.046** (0.018)	0.068 (0.075)
Mean Y (Control)	1.012	1.156	1.110	1.079	1.028	1.486
SD	0.050	0.338	0.224	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 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. The sample excludes non-Sunni Muslim groups and adds a dummy variable for areas with a tribal majority.