Title
Essays in the Political Economy of Development

Author
Testa, Patrick Allen

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

ESSAYS IN THE POLITICAL ECONOMY OF DEVELOPMENT

A dissertation submitted in partial satisfaction of the
requirements for the degree
of
DOCTOR OF PHILOSOPHY
in
ECONOMICS

by

Patrick Allen Testa

Dissertation Committee:
Professor Stergios Skaperdas, Co-Chair
Professor Daniel Bogart, Co-Chair
Professor Jean-Paul Carvalho

2019

Dedication

To Amy and Arthur.

Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my wife, Amy, for her endless supply of encouragement and distractions, as well as my parents – the Flaadas, the Testas, and the Kopecs – for their love and support.

I would like to thank my three incredible advisors, Stergios Skaperdas, Jean-Paul Carvalho, and Dan Bogart, for their valuable lessons in patience, communication, and confidence.

I would like to thank Melissa Valdez, Adam Cook, and John Sommerhauser for their tireless administrative support.

I would like to thank the School of Social Sciences and the Department of Economics at the University of California, Irvine, as well as the Institute for Humane Studies for financial support.

I would like to thank Elsevier for agreeing to publish my third chapter and acknowledge their permission for me to reprint it here, as credited above.

I would like to thank all of the teachers who inspired me over the years and helped put me on this path.

Lastly, I would like to thank all of the friends I have made over the course of this journey, including but not limited to Mike Shin, Kara Dimitruk, Cole Williams, David Simpson, Andy Ferrara, Brian Albrecht, Alex Luttmann, Max Littlejohn, Tim Young, Brittany Bass, Cody Nehiba, Florian Madison, Mark Hup, Arpita Nepal, Zach Schaller, and Nishtha Sharma. I am much smarter for having known you all.

Curriculum Vitae

Education

University of California, Irvine, Irvine, CA
Doctor of Philosophy, Economics — June 2019
Distinction in Microeconomics
Master of Arts, Economics — December 2015

Loyola University, New Orleans, LA
Bachelor of Business Administration, Economics — May 2014
Minor in Mathematics

Awards

- A. Kimball Romney Outstanding Paper Award, *UCI School of Social Sciences* — 2019
  "The economic legacy of expulsion: Lessons from postwar Czechoslovakia"
- Sheen T. Kassouf Endowed Fellowship, *UCI Department of Economics* — 2018
- Christian Werner Fellowship, *UCI School of Social Sciences* — 2018
- Art DeVany Best Poster Award, *UCI Department of Economics* — 2016
- Economics Merit Fellowship, *UCI Department of Economics* — 2016, 2018
- Summer Research Fellowship, *UCI Department of Economics* — 2016 – 2018
- Humane Studies Fellowship, *Institute for Humane Studies* — 2014 – 2018
- Merit Fellowship in Economics, *UCI Department of Economics* — 2014 – 2019

ABSTRACT OF THE DISSERTATION

Essays in the

Political Economy of Development

by

Patrick A. Testa

Doctor of Philosophy in Economics

University of California, Irvine, 2019

Professor Stergios Skaperdas, Co-Chair

Professor Daniel Bogart, Co-Chair

This dissertation consists of three essays in the political economy of development. It uses a combination of empirical methods and microeconomic theory, utilizing both historical and contemporary data.

The first essay examines the long-run effects of forced migration on the origin economy, using Czechoslovakia's expulsion of 3 million Germans after WWII. For identification, I use the discontinuity at the border of the "Sudetenland" region where Germans lived, as made formal with the Munich Agreement in 1938. Since Germans had similar socioeconomic characteristics to Czechs, this bypasses factors that might drive effects elsewhere, such as differences in human capital and geography. The expulsion produced differences in population density, sectoral structure, and education between neighboring municipalities, which persist 70 years later. I trace effects to a selective resettlement of affected areas, generating de-urbanization and human capital decline. Empirical and historical evidence suggest agglomeration economies and extractive institutions as two forces driving this response.

The second essay examines how formal institutions influence local recovery to population shocks, using a model with multiple regions and increasing returns to economic activity within regions. Extractive institutions crowd out productive activity, making its spatial coordination more difficult in the aftermath of large, negative shocks. Given this, I show that when one region experiences such a shock, extractive institutions can hinder recovery, ensuring a redistribution of productive activity away from that region over the long-run.

The third essay considers the conditions under which nondemocratic regimes invest in public education. Nondemocratic regimes face a tradeoff when investing in public education. Education promotes human capital acquisition, expanding the tax base. Yet it also enhances political sophistication and participation, at a cost to nondemocratic regimes. To relax this tradeoff, a regime can disseminate propaganda through its education system. I show that even Bayesian citizens can be influenced by propaganda. By deterring political opposition, propaganda can induce nondemocracies to invest in education when they otherwise would not, improving social welfare. When propaganda is too strong, however, it can generate a backlash. Using cross-country and survey data, I find evidence consistent with the predictions.

Key words: political economy; development; human capital; migration; institutions; nation-building; European economic history

Chapter 1

The Economic Legacy of Expulsion: Lessons from Postwar Czechoslovakia

1.1 Introduction

Between 1945 and 1947, 3 million Germans were expelled from Czechoslovakia, in one of several forced migration events that transformed postwar Europe. Yet such is hardly unique. Jews, Greeks, Turks, and countless others were uprooted during the 20th century in the process of nation-building, and more recently, civil and ethnic conflict have driven large-scale population outflows in places like Uganda, Bosnia, Syria, and Myanmar. A large literature has documented the impact of such events, both on forced migrants, who often experience continued persecution abroad as refugees, and on their host economies (Ruiz and Vargas-Silva, 2013; Becker et al, 2019).

What, however, becomes of the places left behind? Comparably little work has been done to understand the effects of forced migration on the "origin economies" in which the displacement occurred. In this paper, I examine such effects over the long-run, using Czechoslovakia's expulsion of 3 million Germans after WWII. This event has several features which are well-suited for identifying the effects of a forced migration event in ways that prior literature has not. Following the rise of nationalism and the collapse of Austria-Hungary into nation-states in 1918, those still identifying as German within the Czech lands (i.e. the modern day Czech Republic) were concentrated in one region – the borderlands, or "Sudetenland"

– distinguished by a distinct "language border". Yet after centuries of coexistence and common rule, during which language had held little economic or social significance among the masses, Germans and Czechs were relatively similar in terms of their occupational and cultural characteristics (Zahra, 2008). This boundary was made formal in 1938 with the Munich Agreement, which enabled Germany's annexation of the majority-German borderlands, followed by its occupation of the remainder of the Czech lands shortly thereafter. It was in response to this that Czechoslovakia expelled nearly all of its German population in 1945.

My identification strategy exploits the discontinuity in exposure to the expulsion at this boundary to identify its local relative effects over the long-run, using a spatial regression discontinuity (RD) design.[1] To do this, I first construct a new dataset of municipal- and district-level data spanning 90 years. Using directories of Czech villages from during the war, I divide the Czech lands into a "borderlands" treatment region exposed to the expulsion and an "interior" region not exposed. I then examine economic outcomes in the borderlands, relative to interior areas only a few kilometers away. To the extent that ethnic differences were uncorrelated with relevant factors prior to the expulsion, any long-run differences must be driven by the expulsion and its subsequent channels, as opposed to differences in factors such as geography and human capital that often distinguish displaced populations and might therefore drive effects in other settings. Accordingly, long-run effects in this setting are *ex ante* unclear (i.e. path dependence versus steady state reversion).

I then combine data on a large set of socioeconomic variables from the interwar period with existing historical evidence to show that places with Germans and nearby places without them were indeed indistinguishable prior to the expulsion on a number of relevant dimensions, such as population density, literacy, and sector composition. This is true even if one omits segments of the language border around which the borderlands was more ethnically mixed.

Next, I document a large divergence in development today between neighboring mu-

[1] Use of ethnic boundaries in spatial RD follows a literature on ethnic conflict and institutions, including Michalopoulos and Papaioannou (2013), Grosfeld et al (2013), and Moscona et al (2018).

nicipalities on either side of the former German language border,[2] including differences in both the intensity and the composition of economic activity in a municipality. As of 2011, borderland municipalities had lower population density, higher unemployment, smaller skill-intensive sectors, and lower educational attainment relative to nearby interior municipalities. The magnitude of differences is robust to numerous specifications and sample restrictions.

Lastly, studying a historical expulsion allows me to shed light on the precise channels through which such differences emerged. I show that the borderlands underwent a *de-urbanization* after the expulsion, with a relative decline in population density and a shift in sectoral structure toward agriculture. This originated from a *selective initial migratory response* to the expulsion. Contrary to policymakers' expectations that former-German areas could be quickly and voluntarily resettled by Czechs from nearby interior areas, interior population outflows did not match overall losses in the borderlands by the time its initial resettlement had wound down in mid-1947. These differences were larger for more urban sectors, such as business and transportation. This in turn gave rise to *differences in human capital acquisition* between the regions. I document relatively lower levels of enrollment in advanced secondary, technical, and tertiary schooling in the borderlands in mid-1947, despite there being a similar or greater supply of education at these levels as measured in schools and teachers per pupil. I then discuss potential mechanisms though which the expulsion generated these short-run effects. Using data and historical evidence, I argue for (i) agglomeration economies and (ii) extractive institutions as two compelling forces driving these patterns. I also consider how other factors such as natural geography, central planning, and Cold War geopolitics may have mattered for long-run effects.

The borderlands' decline was initially not intended by Czechoslovak policymakers, for whom the region had great economic importance. Yet the expulsion of the Germans had a persistent impact on the places in which they had once lived, relative to non-German places

[2] Effects are relative due to spillovers across regions over time. Aggregate effects would likely be even larger, since the interior was negatively affected via out-migration, while the borderlands has likely benefited from market access spillovers in areas near the former German language border.

nearby. These findings provide valuable new insight into the economic effects of forced migration. While voluntary migration has been the subject of vast research and debate (Bell et al, 2013; Abramitzky et al, 2014), migration occurring as a result of expulsion as well as violence and disaster is increasingly relevant, with UN estimates placing the number of forcibly displaced people at nearly 70 million worldwide.[3] Moreover, forced migration is often followed by expropriation and conflict. Hence, it may have effects that differ from those of voluntary migration.

This paper contributes to the literature on forced migration in two ways. First, whereas existing research has focused largely on the effects of forced migration events for host countries (Hornung, 2014; Johnson and Koyama, 2017; Michalopoulos et al, 2019) or on migrants themselves (Bauer et al, 2013; Becker et al, 2019), less work has been done to study their effects on the origin countries overseeing such displacement (Becker and Ferrara, 2019). The findings in this paper suggest that forced migration may not only affect migrants and their host economies but contribute to persistent geographic inequality within the origin economy. This is most similar to Chaney and Hornbeck (2016), who find delayed convergence following the Spanish expulsion of the Moriscos between former Morisco and non-Morisco districts. As in their paper, studying a politically-motivated expulsion yields advantages for identification here, to the extent that it is less likely to be associated with loss of physical capital or selection within the targeted group, relative to war or natural disaster.[4]

Unlike this paper, however, existing research has focused on forced migrations involving relatively skilled or otherwise differentiated groups (Waldinger, 2010; Acemoglu et al, 2011; Akbulut-Yuksel and Yuksel, 2015; Pascali, 2016). To my knowledge, this paper provides the first evidence that a forced migration can have persistent local effects even when displaced populations are not compositionally distinct in relevant ways from those remaining. The existence of effects independent of relative composition suggests that expelling even relatively

[3]See https://www.unhcr.org/globaltrends2017.

[4]An estimated 7000 German civilians were murdered by Czechs during the expulsion, with expellees being allowed only 100 lbs of property to take with them (Gerlach, 2017).

low-skilled groups may have growth-inhibiting effects locally, raising concerns about economic arguments for mass deportation and forced emigration policies more generally. It also mirrors a large literature on the *benefits* of population *inflows*, of refugees and other immigrants (Foged and Peri, 2016; Rocha et al, 2017; Droller, 2017; Murard and Sakalli, 2018; Sequeira et al, 2019). Finally, it adds to a body of research examining the economic consequences of the postwar German expulsions specifically (Schumann, 2014; Semrad, 2015; Braun et al, 2017; Becker et al, 2019).

Lastly, this paper speaks to broader questions in development and urban economics regarding the importance of historical shocks for long-run development. Empirically, it contrasts with Davis and Weinstein (2002; 2008), who argue against the empirical relevance of multiple equilibria in economic activity, citing the relatively quick recovery of Japanese cities after WWII. Yet unlike many of the shocks studied in this literature, expulsion often occurs in weak or repressive institutional settings, and its effects may be contingent upon that context. In particular, an expulsion may be more likely to also generate spatial variation in fundamentals, such as local institutions and culture. This potentially suggests an alternative explanation for the differential effects of population shocks as noted in Acemoglu et al (2011), who attribute the Holocaust's persistent effects in Russia to structural changes stemming from compositional differences between expellees and non-expellees. Rather, the persistent effects of shocks may have more to do with how they interact with or shape institutions and culture, as in Chaney and Hornbeck (2016) and Nunn (2007; 2008).[5]

The remainder of the paper is structured as follows. Section 1.2 provides historical background. Section 1.3 describes the data. Section 1.4 describes identification and results. Section 1.5 considers possible channels. Section 1.6 concludes.

[5]See Acemoglu et al (2002), Tabellini (2010), Dell (2010), and Acemoglu and Dell (2010) for more on history, institutions, and development. Also see Maloney and Caicedo (2015), Jedwab et al (2019), Dell and Olken (2019), and Testa (2019) for examples of how agglomeration economies and institutions may interact to shape spatial development.

1.2 Historical background

The origins of the "Sudeten" Germans in the Czech borderlands can be traced to the 12th century, when early Bohemian kings opened them up to immigration by German-speaking artisans (de Zayas, 1989). They would pay taxes but trade relatively freely, diffusing their language and culture in the process (Agnew, 2004).

After the Thirty Years' War, the Czech lands underwent further "Germanization" under Habsburg rule. During this time, more German speakers moved into the borderlands, creating vast German language frontiers (Daněk, 1995). Yet despite growing German hegemony among the elite, German and Czech speakers coexisted peacefully at the local level, where identity depended more on local kinships than on language (King, 2002; Tampke, 2003). German industries attracted Czech speakers to German towns, creating bilingual economic centers where intermarriage was not uncommon and language choice was largely situational (Agnew, 2004; Zahra, 2008). By the early 1800s, language in the Czech lands was largely independent of economic, cultural, or even genetic factors among the masses (King, 2002; Zahra, 2008).

The origins of a language border

A series of events resulted in increased segregation of German and Czech speaking within the Czech lands. From the late 1800s, nationalist activists worked to build exclusively German or Czech societies by establishing linguistically segregated social associations and lobbying for reforms that limited bilingual education (Tampke, 2003; Judson, 2006; Zahra, 2008). These efforts often took on a geographic dimension. Because later Austrian censuses required citizens to select a single language of daily use, German nationalists developed a visual of the borderlands as a distinctly German region and aimed to "Germanize" its mixed elements. Czech nationalists, in contrast, sought to preserve the historic boundaries of the Lands of the Bohemian Crown (i.e. the modern day Czech Republic) and built exclusively Czech-language institutions to combat Germanization (Bryant, 2002; Zahra, 2008). As such,

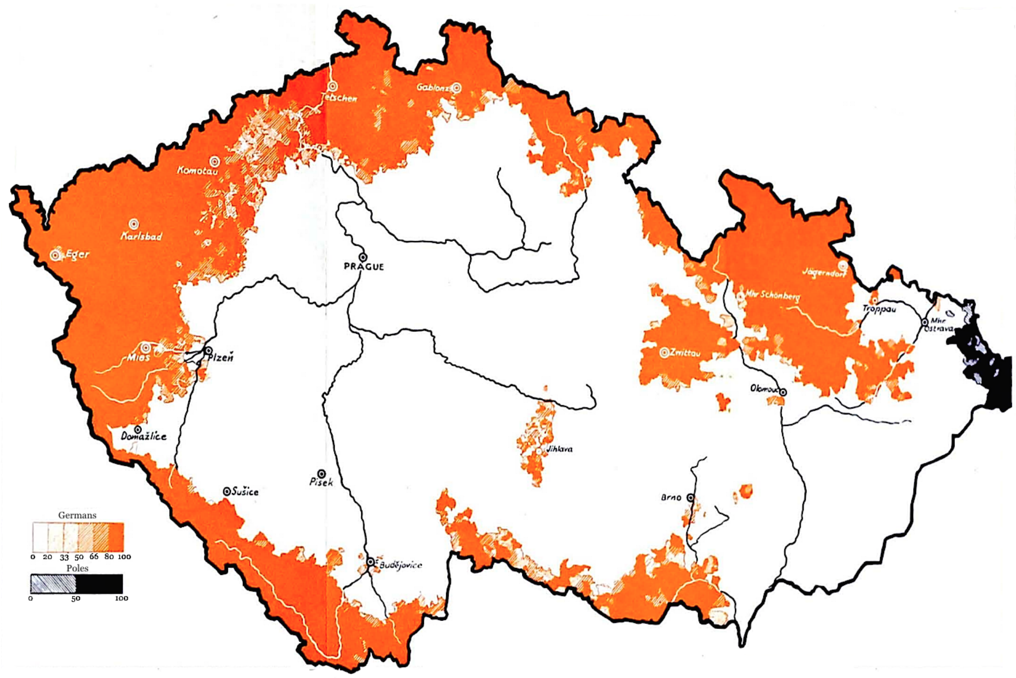


Figure 1.1: Germans in the Czech lands, post-1918 (Wiskemann, 1938)

much of activists' efforts focused on the more mixed areas where the borderlands met the Czech interior (Cornwall, 1994). And although the masses remained largely indifferent to national identification, the late 19th century indeed saw greater assimilation along lingual lines in these places. By the early 1900s, German "language islands" in the interior had shrunk dramatically and German-speaking populations in urban interior areas were rapidly declining (Agnew, 2004; Zahra, 2008).

This benefitted both groups, albeit in different ways. The notion of the borderlands as a formal, German region was further legitimized when, following the collapse of Austria-Hungary in 1918, it was included in German Imperial Council representatives' proposal for a new German-Austrian state – despite the several hundred thousand Czech speakers and countless nationally ambiguous and indifferent residing there (Agnew, 2004). When the Allies instead backed Czech efforts to keep historic boundaries intact, the inclusion of the over 3 million borderland German-speakers as a minority group in the new Czechoslovak state would serve as an important step toward unifying them around a cohesive "Sudeten" German identity, rather than local identities independent of language (Gerlach, 2017). Meanwhile, Czechoslovak policymakers now had the power and legitimacy of the state to influence the ethnolinguistic composition of the Czech lands. After 1918, a parent's nationality as it appeared on the census determined a child's language of instruction, with minority-language schooling and public services being provided only if a minority group exceeded 20% of the local population (Zahra, 2008; Agnew, 2004). Coinciding with this, census officials could now choose nationalities for citizens based on "objective" traits, and failure to comply became a punishable offense. In total, German population counts in mixed areas fell by 420,000 (Zahra, 2008). At least officially, over 90% of German-speakers in the Czech lands lived in the borderlands in 1930.

Despite this, issues of nationality did not dominate political discourse during the 1920s, and German and Czech societies functioned in relative political and economy harmony (Tampke, 2003). After centuries of coexistence and common rule, during which language differences had mattered little among the masses, such national assignment remained largely arbitrary in terms of economic and even cultural factors (Zahra, 2008). During the Great Depression, however, economic anxiety amplified German concerns about the Czechoslovak state. Export-based industries deep in the borderlands experienced some of the highest unemployment in Czechoslovakia, and many Germans blamed the Czechoslovak government. The nationalist Sudeten German Party (SdP) was founded in October of 1933, and before long demands for autonomy were part of its platform (de Zayas, 1989). Although most popular German political parties remained anti-separatist and coalesced with leading Czech parties on common socioeconomic issues during the 1920s, by 1938, 85% of Sudeten Germans supported the SdP (Glassheim, 2016).

Germany's proximity to the borderlands made its invasion highly likely. In an attempt to avoid war, Allied leaders signed the Munich Agreement, formally annexing all majority-German areas to Nazi Germany (see Figure 1.2; Taylor, 1980; Goldstein, 1999), fully formal-

Figure 1.2: The occupied Czech lands within Central Europe, 1939

izing the "Sudetenland" as a region.[6] Meant to appease Germany, annexation severely weakened Czechoslovakia's military and industrial capacities (Agnew, 2004; Glassheim, 2016). Within a few months, Germany had occupied the remainder of the Czech lands, sending its government into exile.

The expulsion of the Sudeten Germans

The idea of expulsion arose soon after Nazi occupation. During the war, the exiled Czechoslovak government, led by former president Edvard Beněs, established the legal basis for the expulsion of Germans from Czechoslovakia through several decrees. Thousands of national committees were to be set up throughout the borderlands to manage the expulsion, including the confiscation of German farms, houses, and other property without compensation and their allocation to incoming settlers. In the end, 3 million Germans would be expelled, mostly to the Western Zones of occupied Germany, along with almost a million to the East Zone and 142,000 to Austria (*Odsun: Die Vertreibung der Sudetendeutschen*, 1995).

[6]That being said, because the entire region was annexed, some <50% German "islands" were inevitably included, especially in northern Moravia. Moreover, Zaolzie in Cieszyn Silesia, a largely Polish-speaking area, was annexed by Poland in 1938. I exclude this latter area from the analysis.

When the war ended in early 1945, Allied forces moved into the borderlands to liberate Czechoslovakia from Nazi Germany, resulting in the first expulsions. It was not until June, however, that the expulsions would gain momentum. By summer's end, as many as 800,000 Germans had been expelled (Gerlach, 2017).

The 33rd Beněs Decree, signed on August 2, 1945, followed the formal Allied approval of the expulsion at Potsdam and formally stripped all Sudeten Germans of their citizenship. Another 2.2 million Germans were expelled through mid-1947 (Gerlach, 2017). These transfers were more systematic in comparison to the earlier "wild transfers". All borderland residents were suspected of being German. When in doubt, the 1939 German or earlier Czechoslovak censuses could determine whether one was to be expelled. This meant that some Germans who had become Czechs by force prior to the war were not expelled, while some non-Germans who had "switched" to German following the borderlands' annexation were (Spurný, 2013). For others, having an ambiguous or mixed national identity meant being expelled, regardless of census identification. Once again, "objective" characteristics were to be used to determine someone's nationality when necessary (Zahra, 2008). Only a small number of Germans who were Czech by marriage, could prove their loyalty to the state, or were deemed economically vital were allowed to stay. By 1950, only 165,000 Germans remained, of which most would be re-granted citizenship (Cornwall, 1994).

The resettlement of the borderlands

The borderlands' resettlement was of central importance to the Czechoslovak government, which sought to maintain the region's great prewar output (Glassheim, 2016; Gerlach, 2017). In May 1945, the Czech borderlands contained upwards of 500,000 non-Germans, and the Czechoslovak government hoped that about 2.5 million more would arrive to resettle the region. Unlike those of other postwar expulsions, this process was to be voluntary, with property reallocation being managed by local national committees. The Czechoslovak government saw this as important for ensuring the elimination of perceived differences between the interior and borderlands, with the hope that confiscated land and property sold at low rates would be sufficient to spur a rapid and full resettlement (Gerlach, 2017). Settlers were to be made up solely of Czechs from nearby "interior" areas. However, as labor shortages ensued, policymakers recruited some Slovaks and others from abroad (Gerlach, 2017). Resettlement began in 1945, concurrently with the expulsion.

Early on, resettlement fed back into expulsion, with Germans being kept around for their labor until settlers began to arrive (Gerlach, 2017). Farmer-settlers were distributed 9-12 hectare lots at low rates (Korbel, 1959). Others were made property administrators of confiscated factories and other businesses. As a result, towns with the most appealing property, particularly those closest to the interior, were emptied and resettled most quickly (Daněk, 1995).

1.3 The data

This section provides an overview of the district- and municipal-level dataset compiled for this paper. It spans over 90 years and contains newly- and already-digitized data from historical censuses, statistical journals, and demographic yearbooks.

Treatment variable

The main treatment variable in Section 1.4 (i.e. located in the borderlands, or former "Sudetenland") is coded from two directories. As the primary source, I use *Amtliches Deutsches Ortsbuch für das Protektorat Böhmen und Mähren*, published in 1940, which lists villages not annexed by Nazi Germany in 1938 (i.e. in the Protectorate of Bohemia and Moravia, or interior) by German and Czech name, regional council district (*Oberlandratsbezirk*) and subdistrict (*Bezirk*).[7] As a supplementary source, I use *Sudetendeutsches ortsnamenverzeichnis: Amtliches Gemeinde- und Ortsnamenverzeichnis der nach dem Münchener Abkommen vom*

[7] This list can be accessed at http://www.hartau.de/PBM/Protektorat.html.

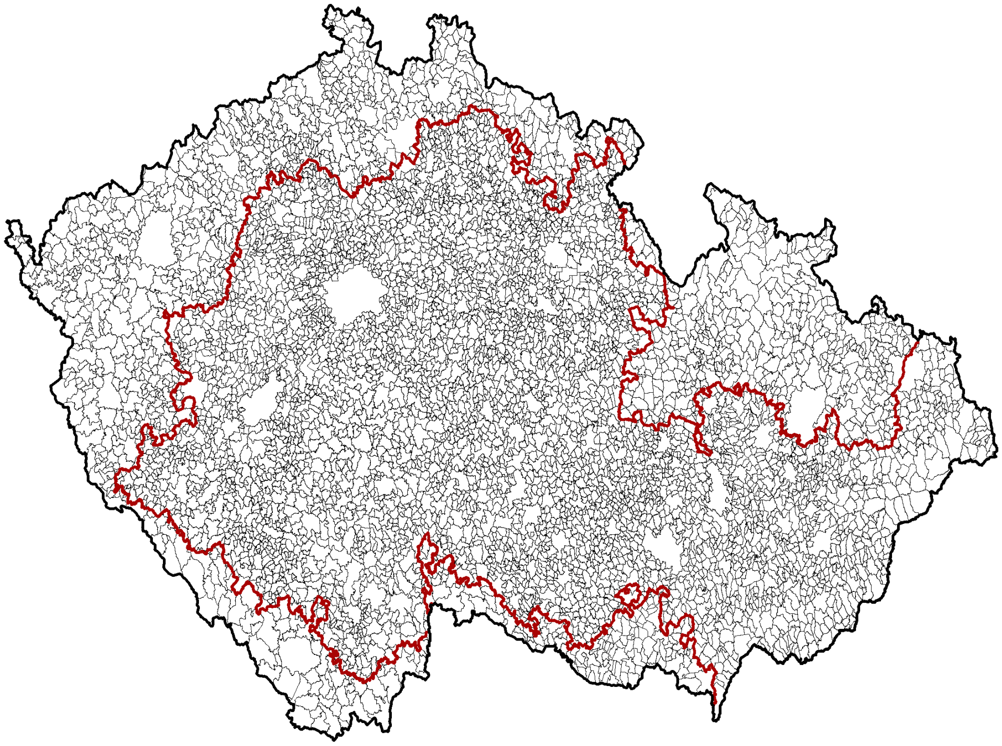


Figure 1.3: Munich Agreement line overlaying 2011 municipalities

29. 9. 1938, published in 1987, which lists villages in the annexed majority-German borderlands alphabetically by German name, along with their Czech name and government district (*Regierungsbezirk*).[8] With the aid of GIS maps of the Protectorate by Jelínek (2011) and 15,070 modern sub-municipal villages (*části obce*), provided by the Czech Statistical Office (CZSO) via its collaboration with the Czech Land Survey Office (ARCDATA PRAHA), I create a precise "Munich Agreement line" (MAL) to measure the German "language border" and sort modern villages into treatment or control groups (see Figure 1.3). I then aggregate this assignment for administrative units as necessary, while noting cases in which they overlap the MAL, and to what extent.

[8]This list can be accessed at http://www.sudeten-by.de/cms/userfiles/downloads/dokumente_nicht-loeschen/Ortsnamen.pdf.

Census data

The first Czechoslovak population census was taken in 1921. Decadal censuses have been held ever since except for during WWII. A smaller population index was also compiled for the Czech lands in May 1947. The 1930 census contains information for 330 judicial districts (*soudní okresy*) and 151 political districts (*politický okresy*) on ethnic composition, literacy, and employment by sector. The 1947 index and 1950 census contain data on employment but not literacy or education, with the number of political districts for the former increasing to 163, and judicial and political districts being consolidated in 1949 into 182 districts (*okresy*). The latter also provides post-WWII data on ethnic composition in the Czech lands (see Table 1.1). From 1961 to 1991, censuses contain information on ethnic composition, educational attainment, and employment by sector for 76 more-aggregated districts. For the 2001 and 2011 censuses, the number of districts increases to 77 and data are also provided for 6258 and 6251 municipalities (*obce*), respectively.

Most data from the 1930, 2001, and 2011 censuses are digitized or taken directly from census files provided by the CZSO. All other district-level census data are made available by the Urban and Regional Laboratory (URRlab) at Charles University in collaboration with the Czech Ministry of Culture, along with corresponding GIS district shapefiles (2017) for each census year. Other administrative boundary data are collected from ARCDATA PRAHA (2017). To construct panels of district-level census data used in supplementary analysis, I use ArcGIS to interpolate subpopulations for a common set of district boundaries. I use 1991 district boundaries, since districts were arguably at their highest level of aggregation that year, minimizing error.[9]

Non-census data

Non-census outcome data come from a variety of sources. I construct 1933 variables using data from two state statistical publications: *Nezaměstnanost a podpůrná péče v Československu*,

[9] See Table A.14 in Appendix A for a description of this process.

a social insurance report published in 1938, and *Statistika daně důchodové placené přímo, daně z vyššího služného, daně rentové placené přímo, všeobecné a zvláštní daně výdělkové podle předpisu za rok 1933*, a taxation report published in 1938. The former provides the number of registered unemployed in each political district as reported by the Minister of Social Affairs, while the latter lists income and the share of eligible taxpayers by political district. These data are combined with data on the size of the labor force and population from the 1930 census to estimate 1933 unemployment and income per capita, respectively.

I digitize non-census data from the post-expulsion period from several statistical reports published from the mid-1940s. In particular, district-level data for arable land in 1945 and school enrollment in 1947 are derived from the 1947 and 1948 editions of the state statistical report, *Zprávy státního úřadu statistického republiky Československé*, respectively. District-level data on migration come from a series of demographic yearbooks, *Pohyb obyvatelstva v republice Československé*, made available by the CZSO. Crime data are provided by URRlab. Data on the number of jobless by municipality in 2011 come from the Czech Ministry of Labor and Social Affairs.

1.4 The regional economic impact of expulsion

After WWII, around 95% of Germans living in the Czech lands (i.e. the modern day Czech Republic) were forced to permanently exit the country, leaving their homes and most property behind. This section examines how this impacted the relative development of former-German places over the long-run. Prior to the expulsion, those still identifying as German within the Czech lands were concentrated in one region, often called the "borderlands" and formerly the "Sudetenland". Although German-speaking had for centuries been more prominent in the corners of the Czech lands, historical developments associated with the rise of nationalism resulted in this region becoming semi-formal by the 1920s, defined by a sharp spatial discontinuity in German identification – a "language border" – in official statistics (see Figure 1.1). This boundary was formalized in 1938, when the Munich Agreement enabled the annexation of majority-German areas by Germany. This Munich Agreement line (MAL) in turn approximates a place's exposure to the expulsion in 1945 (see Table 1.1).

My identification strategy exploits the discontinuity in exposure to the expulsion at this boundary in order to identify the local relative effects of the expulsion over the long-run. In particular, I use a spatial regression discontinuity (RD) design, with the treatment variable being a discrete function of a place's location relative to the MAL. This strategy will help control for potentially confounding factors, such as natural geography and other historical shocks, to the extent that such factors vary smoothly through the MAL.

To do this, I assign a value of 1 to a municipality or district if it was on the majority-German side of the MAL and thus annexed by Germany in 1938. A value of 0 is assigned if it was located in the "interior" where few Germans lived. Consider the municipal-level specification,

$$y_{mdb} = \alpha + \beta InBorderlands_m + f(location_m) + \mathbf{X}'_m\Gamma + \Sigma_b + \Delta_d + \varepsilon_{mdb}, \tag{1.1}$$

where y_{mbd} is the outcome variable for municipality m in district d along segment b of the MAL; $InBorderlands_m$ is a dummy for if a municipality m lies in the borderlands where Germans lived;[10] $\mathbf{X}_m$ is a vector of geographic characteristics, including elevation, ruggedness, temperature, precipitation, and river density (km per km^2); and Δ_d captures district fixed effects. Σ_b gives the set of border segment fixed effects, denoting to which of the fifty segments, each roughly 50 km in length, a municipality is nearest. Finally, $f(location_m)$ is the running variable, capturing all other characteristics that vary smoothly through the MAL. For the main specifications, I use a municipality's centroid distance from the MAL, interacted with the treatment. I explain this choice below. For main 2011 specifications, I

[10] 94 municipalities for which only some parts were annexed are dropped. I define treated as $>$ 95% area annexed. All specifications and plots exclude those which otherwise overlap the MAL. For 1930-47 analyses, where units are larger such that more area is dropped, results are robust to including units that overlap the MAL but were nonetheless homogeneously German (i.e. treated in spite of overlap) or non-German – about half of units dropped – as shown in Tables A.6 and A.19.

Table 1.1: Exposure to expulsion

Region (subsample)	% German, 1930	% German, 1950
Borderlands (within 25 km of MAL, no overlap)	81.78	4.439
	(2.276)	(.597)
Interior (within 25 km of MAL, no overlap)	1.601	.495
	(.396)	(.047)
Borderlands (no bandwidth, no overlap)	86.67	6.821
	(1.646)	(1.184)
Interior (no bandwidth, no overlap)	1.646	.464
	(.418)	(.036)
Borderlands (full sample)	80.874	5.36
	(1.904)	(.808)
Interior (full sample)	3.545	.569
	(.544)	(.042)

Unit of measure for 1930 is 330 judicial districts. Unit of measure for 1950 is 182 districts. Standard errors are reported in parentheses. A district is considered to be in the borderlands (interior) if 50% or more of its area lies inside (outside) the lands annexed by Germany in 1938, as determined by the Munich Agreement line (MAL). The full sample includes districts that overlap the MAL. A district is classified as "no overlap" if >95% of its area lies on one side of the MAL or the other. A district is classified as "within 25 km" if its centroid lies within 25 km of the MAL. As Prague and Polish Zaolzie are always excluded from the analysis, I exclude them here as well.

cluster standard errors by the 71 districts in the final sample.

For 20th century outcomes, for which municipal-level data are not available, district fixed effects are omitted and geographic controls are calculated at the district-level. Depending on the number and size of districts, border segments are also lengthened and the number of them decreases. Most pre-1950 data are available for judicial districts. For those regressions, I control for 24 border segments, each about 100 km in length, and cluster standard errors by political districts, of which 98 remain in the final sample for main specifications. Some 1933 and 1947 data are available only for the larger political districts.[11] For those specifications, I control for 16 border segments, around 150 km in length, and report robust standard errors.

Balance testing

Historians have noted how, following centuries of admixture and common rule, there were few occupational or social differences between Germans and Czechs even in the early 1900s, and the historical developments that gave rise to the MAL and eventual expulsion were ultimately rooted in national rather than economic considerations (Zahra, 2008). Thus, adopting an

[11]Due to a higher level of aggregation, the sample underrepresents southern (i.e. non-Eastern Bloc-bordering) areas. I show in Section 1.5 that long-run effects are independent of this.

Table 1.2: Pre-expulsion economic differences between regions

	%German	Literacy	ln Pop. density	ln Labor force dens.	Labor force part.	Unemploy.
	(1a)	(1b)	(1c)	(1d)	(1e)	(1f)
In borderlands	67.968	-.222	-.318	-.329	-.400	-3.936
	(6.042)***	(.217)	(.208)	(.227)	(1.196)	(2.535)
R^2	.933	.533	.475	.482	.623	.69
Observations	165	165	165	165	165	104
Clusters	98	98	98	98	98	–
Border segments	24	24	24	24	24	16
Bandwidth	25 km	25 km	25 km	25 km	25 km	50 km
Year	1930	1930	1930	1930	1930	1933
	Income$_{pc}$	%Taxpayers	Convicts$_{pc}$ 1923-27	%Roma	%Jewish (ethnic)	%Jewish (faith)
	(2a)	(2b)	(2c)	(2d)	(2e)	(2f)
In borderlands	-1.866	.006	-.454	-.002	-.113	-.086
	(1.832)	(.604)	(.683)	(.002)	(.140)	(.274)
R^2	.409	.562	.361	.1	.324	.232
Observations	104	105	164	165	165	165
Clusters	–	–	98	98	98	98
Border segments	16	16	24	24	24	24
Geo. controls	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	50 km	50 km	25 km	25 km	25 km	25 km
Year	1933	1933	1923-27	1930	1930	1930

Robust standard errors are clustered by political district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. Standard errors in specifications with political districts are heteroskedasticity-robust. All regressions exclude Prague and Polish Zaolzie, include exogenous controls for elevation, ruggedness, precipitation, temperature, and river density (km per km^2), and utilize a local linear running variable of distance from the Munich Agreement line interacted with the treatment.

RD approach here should theoretically identify the local relative effects of the expulsion, absent differences in factors like geography and human capital, which commonly distinguish different ethnic groups and might therefore drive effects in other settings or under alternative empirical strategies. Nonetheless, it is important to check that the ethnolinguistic differences for which the MAL was drawn (and the drawing of the MAL itself) were not associated prior to 1945 with differences in factors relevant for economic performance.

To test whether ethnic differences were correlated with other distinctions locally, I estimate differences for 24 pre-treatment outcomes. Estimates can be found in Tables 1.2 and 1.3.[12] Column (1a) of Table 1.2 confirms that traversing the MAL was associated with an increase of about 68 percentage points in the German population in 1930. However, this

[12]See Appendix A for RD plots of all regressions, using a linear polynomial in distance from the Munich Agreement line and 50 km as the bandwidth, as well as geographic heatmaps for sectors (in addition to Figure 1.8).

Table 1.3: Pre-expulsion sectoral differences between regions

	Agricultural sector	Mining and extraction	Metals	Manufacturing	Glass	Textiles
	(1a)	(1b)	(1c)	(1d)	(1e)	(1f)
In borderlands	2.915	-1.150	.455	-.421	1.065	-3.311
	(3.649)	(1.698)	(1.458)	(.561)	(1.681)	(2.534)
R^2	.526	.376	.313	.298	.338	.634
	Other industry	Con-struction	Transport sector	Finance and insurance	Trade	Other service
	(3a)	(3b)	(3c)	(3d)	(3e)	(3f)
In borderlands	.102	.500	-.457	-.110	-.534	-.355
	(1.429)	(.719)	(.665)	(.094)	(.796)	(.927)
R^2	.315	.332	.318	.258	.377	.209
Observations	165	165	165	165	165	165
Clusters	98	98	98	98	98	98
Border segments	24	24	24	24	24	24
Geo. controls	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	25 km	25 km	25 km	25 km	25 km	25 km
Year	1930	1930	1930	1930	1930	1930

Robust standard errors are clustered by political district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie, include exogenous controls for elevation, ruggedness, precipitation, temperature, and river density (km per km^2), and utilize a local linear running variable of distance from the Munich Agreement line interacted with the treatment.

does not coincide with statistically significant discontinuities in other factors at the MAL. This is true even for mining, glassworks, and textile manufacturing, for which German labor was considered to be highly important *on the whole*. These findings are robust to changes in RD bandwidth and running polynomial, as shown in Table A.4 in Appendix A. In Figure 1.4, I show similar patterns for elevation, ruggedness, temperature, and precipitation.[13] With the exception of ruggedness, geographic differences appear smooth through the MAL. To check smoothness for ruggedness, I report sample means at various bandwidths in Table A.2. Long-run results are also robust to excluding mountainous regions entirely.

The historical literature also downplays local differences in institutions. Even as national identification became more common and geographically salient after WWI, individual Czechoslovak citizens retained equal rights in the eyes of the state, including education and basic healthcare (Bryant, 2002; Tampke, 2003). Minority group rights were weaker, yet still liberal relative to the country's peers, with facilities provided for German-language educational and legal activities in sufficiently German (i.e. $> 20\%$) areas and a role for German

[13] See Figure A.3 for the same plots for river density and arable land in 1945.

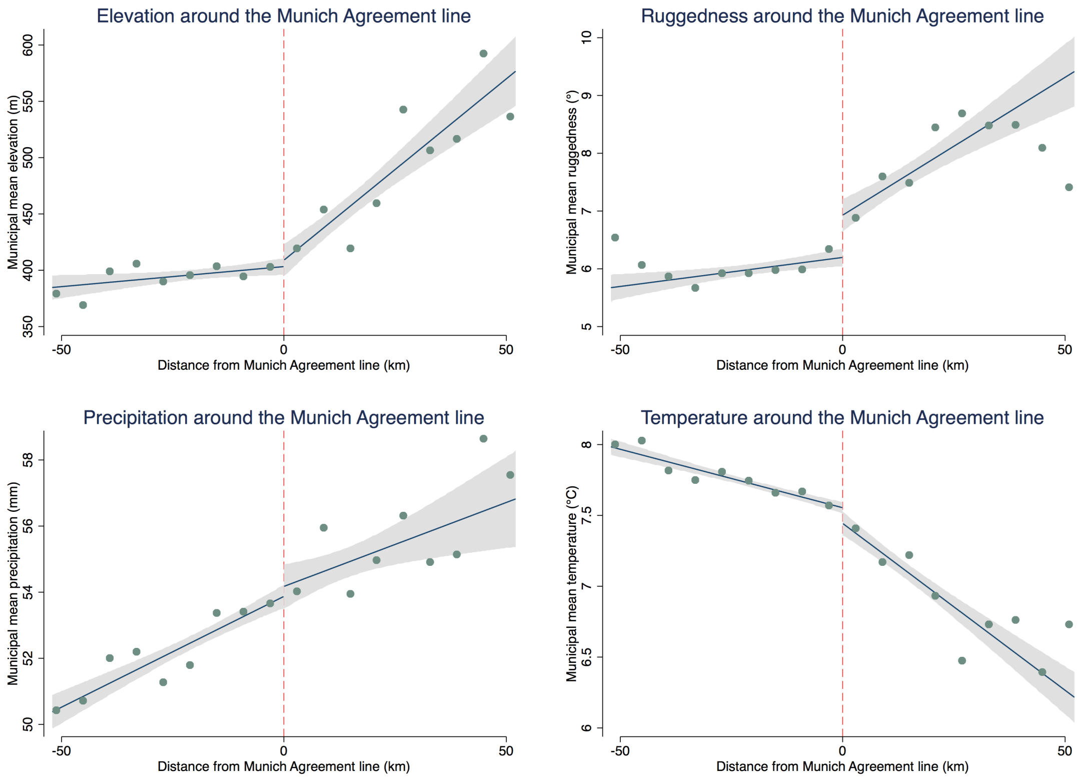


Figure 1.4: Smoothness of geography around Munich Agreement line

politicians in parliamentary politics (Tampke, 2003; Zahra, 2008). This relative egalitarianism is apparent in the borderlands' and interior's similar economic and taxation statuses, as shown above. Overall, historians recognize the widespread similarities of Germans and Czechs in the Czech lands after centuries of coexistence – as did the Nazis (Bryant, 2007).

WWII, borderland Czechs, and pre-treatment sorting

The same nationalism that motivated the expulsion of the Germans had also inspired the annexation of the borderlands and the occupation of Czechoslovakia by Nazi Germany years prior. Yet one concern is that such events and others associated with them may have differentially affected borderland areas near the MAL. Indeed, a major limitation to studying this era is the lack of data from during World War II. I address these concerns here.

Following annexation, the remainder of the Czech lands was quickly occupied by Nazi Germany until 1945. The Czech government was exiled and preferential treatment of the

Czech regions by the Nazis during the war would have favored the borderlands, biasing estimates toward zero. That being said, the Czech lands experienced little physical destruction during the war.[14] There were also few acts of resistance by Czechs. The Czech lands held a large and important industrial workforce, and Nazi officials came to see much of the Czech masses as "Germanizable" due to a cultural and genetic closeness forged by centuries of coexistence. Thus, life in the Czech interior continued largely as normal, avoiding much of the violence experienced by Yugoslavia and Poland (Agnew, 2004; Bryant, 2007; Glassheim, 2016). Economic life in the borderlands also changed little, at the displeasure of some borderland Germans, who had sought greater integration into German economy and society. In all, "Czechoslovakia emerged from the war with much of its industrial base intact" (Gerlach, 2017, 208). Differences in wartime deaths are more difficult to discern. Historical accounts suggest the borderlands suffered somewhat more in terms of war casualties than interior areas, due to conscription of some Sudeten Germans. However, these estimates would have been driven up by casualty counts from the violent liberation of Czechoslovakia in May of 1945, which also marked the beginning of the expulsions. And although Jews and Roma were also expelled from and murdered in the Czech lands during this time, these groups were distributed uniformly through the MAL, as shown in Table 1.2. Thus, pre-expulsion differences in casualty rates between the regions were likely small on net.[15]

To further quell concerns that the effects presented below reflect anything other than the expulsion, I also estimate pre-trends in Table A.5 in Appendix A using data from the 1921 census. With the exception of literacy, which increased slightly in the borderlands relative to the interior between 1921 and 1930, there are no statistically significant differences in pre-trends in my data. To the extent that this implies stability in the relative development trajectories and population patterns of both regions leading up to the war, this also helps further minimize concerns that they might have diverged significantly during the war.

That being said, one thing that could have disturbed pre-trends involves the presence

[14]See Figure A.2 for a map of confirmed Allied bombings during WWII.
[15]See Table A.3 for more details on estimates.

of Czechs in various parts of the borderlands. Unlike the interior, which by 1930 was strikingly homogeneous (at least officially), some parts of the borderlands were more ethnically mixed, especially near the MAL. Related to this was *pre-treatment sorting* among Czechs, which according to historians occurred following Czechoslovak independence in 1918 – first by interior Czechs immigrating to some borderland cities; then by 300,000 of the 730,000 borderland Czechs into the interior during the war; and finally by nearly all of those Czechs back into the borderlands after the war (Cornwall, 1994; Agnew, 2004; Glassheim, 2016).

Even absent pre-treatment sorting concerns, the presence of borderland Czechs renders the cross-sectional balance tests above potentially problematic, as cross-region differences are not as such perfectly informative of cross-ethnicity differences. For example, if Czechs near the MAL were in fact less skilled than Germans, then it could bias balance tests toward zero. In this case there would be relatively smooth estimates, while in actuality the expulsion of the borderlands' Germans would have interacted with crucial, preexisting distinctions between the regions and their residents. In such a scenario, persistent effects would be less surprising (Acemoglu et al, 2011).

Since I cannot compare Czechs and Germans directly, I must rely on heterogeneity in the ethnic composition of the borderlands prior to the expulsion to better compare Czechs and Germans across regions. In particular, I reexamine the balance tests above using a sample of "concrete" stretches of the MAL, in which I compare only homogeneous parts of the interior with nearby, homogeneous parts of the borderlands with few borderland Czechs (as well as less pre-treatment sorting, in turn). In Table A.9 in Appendix A, I show that while this increases the size of the ethnic discontinuity to 86 percentage points, estimates regarding relative literacy, population density, and sector composition change little. Only eligible taxpayers shows a statistically significant change associated with crossing the MAL, with an increase of about 1.6 eligible taxpayers per 100 persons from entering the borderlands.[16] This exercise helps reaffirm the historical narrative and assumption that spatial variation in

[16]Relative pre-trends also remain largely unchanged if I examine only the concrete sample, as do long-run outcomes if I adopt the same concrete stretches. These results are available upon request.

ethnic composition was not associated with differences in relevant factors in the pre-expulsion Czech lands, at least near the MAL.[17]

In contrast, post-treatment sorting is a natural channel of persistence in many spatial RD settings (Dell, 2010). Indeed, it will be an important channel here, since expulsion in the borderlands necessitated resettlement from interior areas.

Controlling for smooth differences

Despite the smoothness of relevant variables through the MAL, identification is still threatened if the running variable $f(location_m)$ does not adequately control for these smooth differences. To deal with this, many spatial RD designs limit the sample to a narrow bandwidth around the border of interest. If imposing a narrow bandwidth is feasible given the sample, then a linear polynomial is likely to be a reasonable control (Gelman and Imbens, 2018). Fortunately, the region of study in this case is quite small, and all samples are easily limited to within 50 km or less of the MAL. For most specifications I adopt the standard bandwidth of 25 km.[18]

A related concern is choice of running variable. A common choice in spatial RD settings is a two dimensional linear polynomial in longitude and latitude. This would vary smoothly through the MAL while controlling for local characteristics. However, it would not allow such differences to vary with treatment status, which could bias estimates. A second common choice is distance from the border of interest. Here, an interaction term lets smooth differences vary in slope on either side of the MAL, with the treatment effect being evaluated at the MAL. Although being x km from the MAL could mean something different in southern Bohemia versus northern Moravia, border segment dummies and other controls should deal with any heterogeneity. I thus opt for the latter in all main specifications. In Appendix A, however, I test alternative specifications, including using local linear polynomials in latitude

[17] And in addition to the kinds of pre-treatment sorting discussed above, this exercise also reaffirms that the kind in which Germans "switched" to Czech also would not have mattered.

[18] Increments of 25 km are standard in the prior literature. Output from optimal bandwidth algorithms varies with outcome but tends near this for municipal- or judicial district-level specifications.

Table 1.4: Long-run differences in economic activity, 2011

	ln Population density	ln Labor force density	Unemployment
	(1a)	(1b)	(1c)
In borderlands	-.312	-.317	2.729
	(.095)***	(.097)***	(.546)***
R^2	.398	.399	.404
Observations	4049	4049	4049
Clusters	71	71	71
Border segments	50	50	50
Geographic controls	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes
Bandwidth	25 km	25 km	25 km
Year	2011	2011	2011

Robust standard errors are clustered by district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie, include exogenous controls for elevation, ruggedness, precipitation, temperature, and river density (km per km^2), and utilize a local linear running variable of distance from the Munich Agreement line interacted with the treatment.

and longitude. These estimates vary somewhat in size but are substantively similar.

1.4.1 Long-run effects

I will now provide evidence that the expulsion of the Germans from the Czech borderlands (i.e. the former "Sudetenland") had long-lasting effects on the spatial distribution of economic activity within the Czech lands. I will focus on two types of outcomes: (i) the intensity of economic activity, as measured by population and employment, and (ii) the composition of economic activity, as measured by variables like education and occupation.

Prior to the expulsion of the Germans, borderland areas were as educated and prosperous as their adjacent "interior" neighbors. If the determinants of economic activity are invariant to the expulsion, then we should see interior Czechs spilling over into relatively similar, nearby borderland areas in smooth ways thereafter. Instead, borderland municipalities today show discontinuous signs of marked economic decline relative to interior towns just a few miles away. As of 2011, population and employment were relatively more concentrated in the interior, as shown in columns (1a) and (1b) of Table 1.4.

Moreover, borderland towns have seen some of the highest unemployment rates since transition within the Czech Republic, which has relatively low unemployment overall. Col-

Table 1.5: Long-run differences in sectoral composition, 2011

	Agricultural sector	Industry	Construction	Transport sector	Finance and insurance	Hospitality
	(1a)	(1b)	(1c)	(1d)	(1e)	(1f)
In borderlands	-.573	-.621	-.447	-.117	-.369	.365
	(.433)	(.860)	(.267)*	(.248)	(.080)***	(.280)
R^2	.303	.34	.155	.19	.134	.339
	Auto trade and repair	Public	Communications	Education	Healthcare	Other service
	(2a)	(2b)	(2c)	(2d)	(2e)	(2f)
In borderlands	-.864	.001	-.297	-.864	-.993	-.219
	(.282)***	(.274)	(.089)***	(.176)***	(.230)***	(.166)
R^2	.201	.125	.204	.085	.139	.214
Observations	4049	4049	4049	4049	4049	4049
Clusters	71	71	71	71	71	71
Border segments	50	50	50	50	50	50
Geo. controls	Yes	Yes	Yes	Yes	Yes	Yes
District F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	25 km	25 km	25 km	25 km	25 km	25 km
Year	2011	2011	2011	2011	2011	2011

Robust standard errors are clustered by district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie, include exogenous controls for elevation, ruggedness, precipitation, temperature, and river density (km per km^2), and utilize a local linear running variable of distance from the Munich Agreement line interacted with the treatment.

umn (1c) shows that simply traversing the MAL is associated with almost a 3% absolute increase in the municipal unemployment rate as of 2011. This illustrates the extent to which production remains less likely to locate in the borderlands, seven decades after the expulsion and over two decades since transition, with low internal migration helping keep rates high (Fidrmuc, 2004; Fidrmuc and Huber, 2007).[19]

Table 1.5 suggests that this is due to an inability of borderland towns to develop skill-intensive service sectors. The borderlands shows significantly less employment on average in finance, education, healthcare, communications, and the auto sector relative to nearby interior areas.

This also suggests a lower concentration of skill in borderland areas in 2011 relative to neighboring interior towns. If true, one would expect borderland municipalities around the MAL to have lower levels of educational attainment than those in the nearby interior. The next set of estimates in this section, in Table 1.6, show this to be the case. The results are

[19] For more on migration, see Table A.21 in Appendix A.

Table 1.6: Long-run differences in educational attainment, 2011

	% Primary education or less	% Secondary education	% Tertiary education
	(1a)	(1b)	(1c)
In borderlands	4.883	-3.923	-1.936
	(.634)***	(.516)***	(.391)***
R^2	.297	.199	.271
Observations	4049	4049	4049
Clusters	71	71	71
Border segments	50	50	50
Geographic controls	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes
Bandwidth	25 km	25 km	25 km
Year	2011	2011	2011

Robust standard errors are clustered by district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie, include exogenous controls for elevation, ruggedness, precipitation, temperature, and river density (km per km^2), and utilize a local linear running variable of distance from the Munich Agreement line interacted with the treatment.

striking: even with a sample bandwidth of just 25 kilometers, there is an absolute increase of over 4.8% of the adult (i.e. age 15 or over) population claiming to have completed no more than a primary education from an average interior municipality to a borderland one. This is matched by an absolute decrease of over 3.9% in secondary schooling completion from crossing the MAL and a 1.9% decrease in tertiary education, such as a college degree.

One concern is that despite the smoothness of geographic characteristics through the MAL, borderland and interior municipalities in the sample still have too many geographic differences for which to adequately control, as Figure 1.4 might imply. One solution is to limit the sample to areas around the MAL that are more geographically homogeneous (i.e. less mountainous). As it turns out, the results do not change if one excludes the mountainous sections of the borderlands from the sample, suggesting geography is not driving the results in the main specification. Results are also robust to numerous changes to bandwidth and RD polynomial (see Tables A.10-12).

I have used this section to show that the expulsion of the Germans generated a persistent, discontinuous spatial divergence in local development within the Czech lands. Whereas the spatial distribution of economic activity was relatively smooth through the Munich Agreement line prior to WWII, productive economic activity is now relatively more concentrated

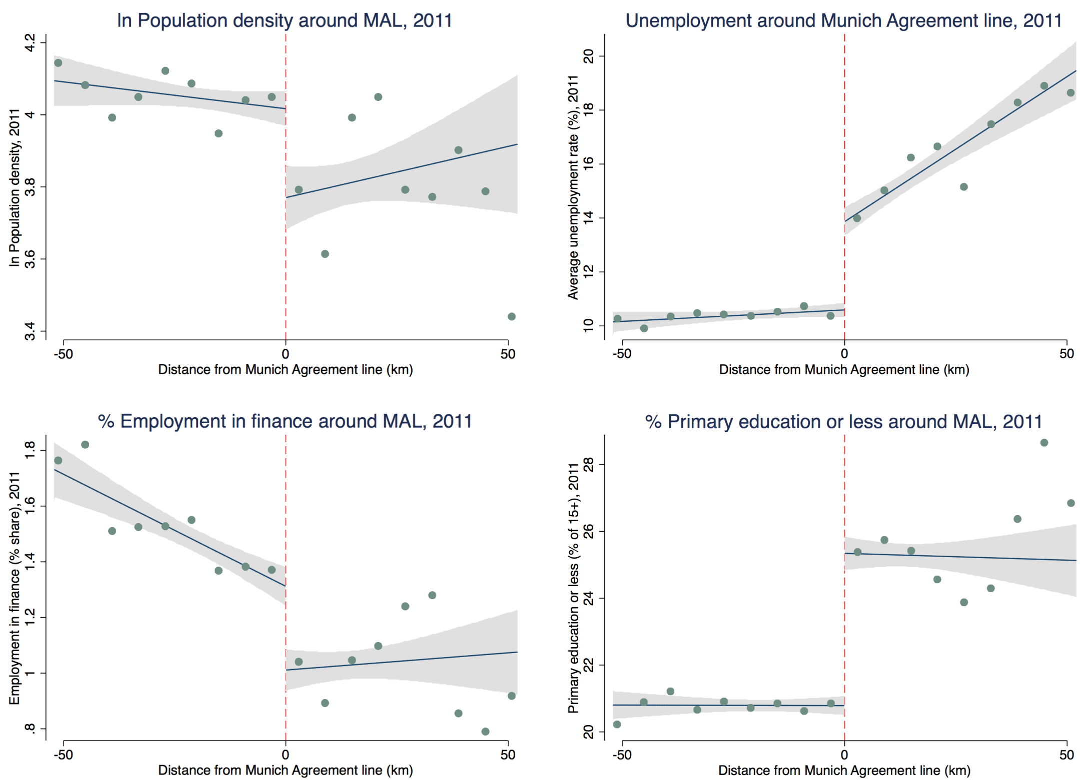


Figure 1.5: Economic activity around the Munich Agreement line, 2011

on the interior side. In the next section, I will explore the precise channels through which the expulsion resulted in this divergence.

1.5 Short-run effects and mechanisms

The previous section documented an unevenness in economic activity present within the Czech lands, nearly 70 years after the German population was expelled from its borderlands, i.e. the former "Sudetenland". This section investigates the origins of these patterns. Since no such unevenness existed prior to the expulsion, I begin by examining the intensity and composition of economic activity of those who voluntarily *resettled* the borderlands from nearby interior areas immediately following the commencement of the expulsion, relative to those who did not. Indeed, the expulsion's impact on the borderlands' relative development depended on it being resettled in a quick and convergent manner. Because the expulsion was near-absolute (i.e about 95% of Germans), I can use data from after the expulsion and the borderlands' initial resettlement had largely concluded in mid-1947 to compare these populations.[20] This will also allow me to examine whether or not key structural changes had already occurred prior to the communist coup of 1948.

In this section, I show that the expulsion was instead followed by a (i) selective initial resettlement, with relatively few workers arriving from sectors like transportation and business relative to agriculture. This culminated in a relative (ii) de-urbanization of former-German areas and the emergence of (iii) human capital inequalities between the regions thereafter. Using a combination of data and historical evidence, I then propose and discuss two mechanisms underlying these short-run effects. First, I argue that large population losses interacted with agglomeration economies, *decreasing* the expected relative return to moving to affected borderland towns for workers in more urban sectors. Second, the expulsion engendered an extractive institutional environment in the borderlands, which helped crowd out production there in the short-run. This analysis focuses on the immediate post-expulsion period, during which the Czechoslovak economy remained generally unplanned, although I also occasionally discuss patterns of local development during the communist period and following transition to a market economy in 1989.

1.5.1 Selective resettlement

The economic future of the borderlands depended foremost on the German population being replaced by settlers of similar count and composition. The stated goal of resettlement was that it be geographically convergent: since the expulsion "would reduce the Czechoslovakia population by 25%, the borderlands would only be resettled up to 75% of [its] original population," preserving prewar *relative* densities while creating one homogeneous nation (Radvanovský, 2001, 203). This was to be done quickly and concurrently with the expulsion,

[20]The historical literature indicates that the 5% or so remaining were disproportionately skilled workers initially kept out of necessary and later dispersed throughout the country in the 1950s. This may downward bias 1947 estimates slightly.

with the goal of maintaining the region's output (Glassheim, 2016). It was also to be carried out on a voluntary basis. Policymakers utilized propaganda and incentives, with local committees transferring confiscated farmland, businesses, and other assets to settlers at low rates. If settlers selected into more appealing locales first, committees would direct surplus settlers elsewhere (Radvanovský, 2001). Besides these interventions, worker and firm-level decisions remained unplanned until the communist coup of 1948 (Bernàšek, 1970).

Under such circumstances, the theoretical migratory response to the expulsion is ambiguous, at least in the neighborhood of the MAL, around which differences in fundamentals and economic activity were smooth *ex ante*. Indeed, one might expect this to have made the borderlands an attractive place in which to settle and invest, inducing Czechs from nearby interior areas to move to the borderlands in a relatively smooth manner. And as it turns out, the expulsion was indeed met with a large-scale migratory response, beginning in mid-1945 through the first half of 1947, with the vast majority of settlers arriving from nearby interior areas (Radvanovský, 2001; Gerlach, 2017; see Figure 1.6).

However, this initial resettlement was *selective* in nature. One way to examine this is to compare population *losses* endured by the interior (i.e. to the voluntary resettling of the borderlands) with those of nearby, similar borderland areas (i.e. from the expulsion net of resettlement) within each sector. Since contiguous borderland and interior districts had similar densities, compositions, and relevant pre-trends on average prior to WWII, with few shocks of permanence or scale during the war, the convergent migratory response envisioned by policymakers would have required that these losses be similar for each sector by the time expulsion and the borderlands' initial resettlement had wound down in mid-1947, relative to pre-expulsion levels. If instead there was differential resettlement across sectors, it could in turn have generated the emergence of the patterns we see today.

I begin by constructing an outcome variable to measure population losses in each judicial district d for each sector s between the 1930 census and mid-1947 index:

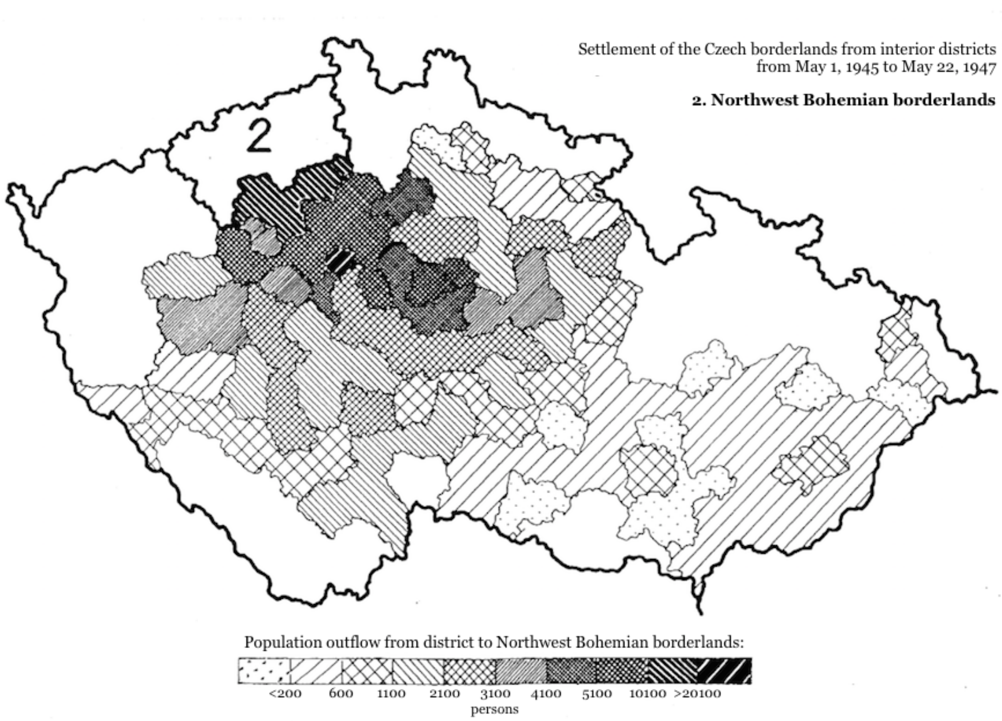


Figure 1.6: Sources of immigration to Northern Bohemia, 1945-7 (von Arburg, 2001)

$$PopLoss_{d,s} = \frac{Pop_{d,s,1947} - Pop_{d,s,1930}}{Pop_{d,s,1930}}.$$

I then examine population dynamics visually by plotting $PopLoss_{d,s}$ overall and by sector (with observations binned for every 6 km) by distance from the MAL. Figure 1.7 shows that while interior areas *did* endure significant population losses to nearby borderland areas during the main period of resettlement (see Figure 1.6), borderland districts near the MAL on average lost much more than the anticipated 25% of their populations on net, while interior districts on average lost less. Net population losses in the borderlands continue to increase as one moves away from the MAL, whereas interior losses decrease.

This lack of short-run convergence has its origins in the sectoral distribution of settlers. Namely, differences in worker losses around the MAL were relatively small for agriculture and large for the other, more "urban" sectors.[21] This is consistent with the historical narrative

[21] Although it is worth noting that the secondary sector includes the relatively small construction sector, which was more prevalent in less populated, developing areas in 1930, as discussed below.

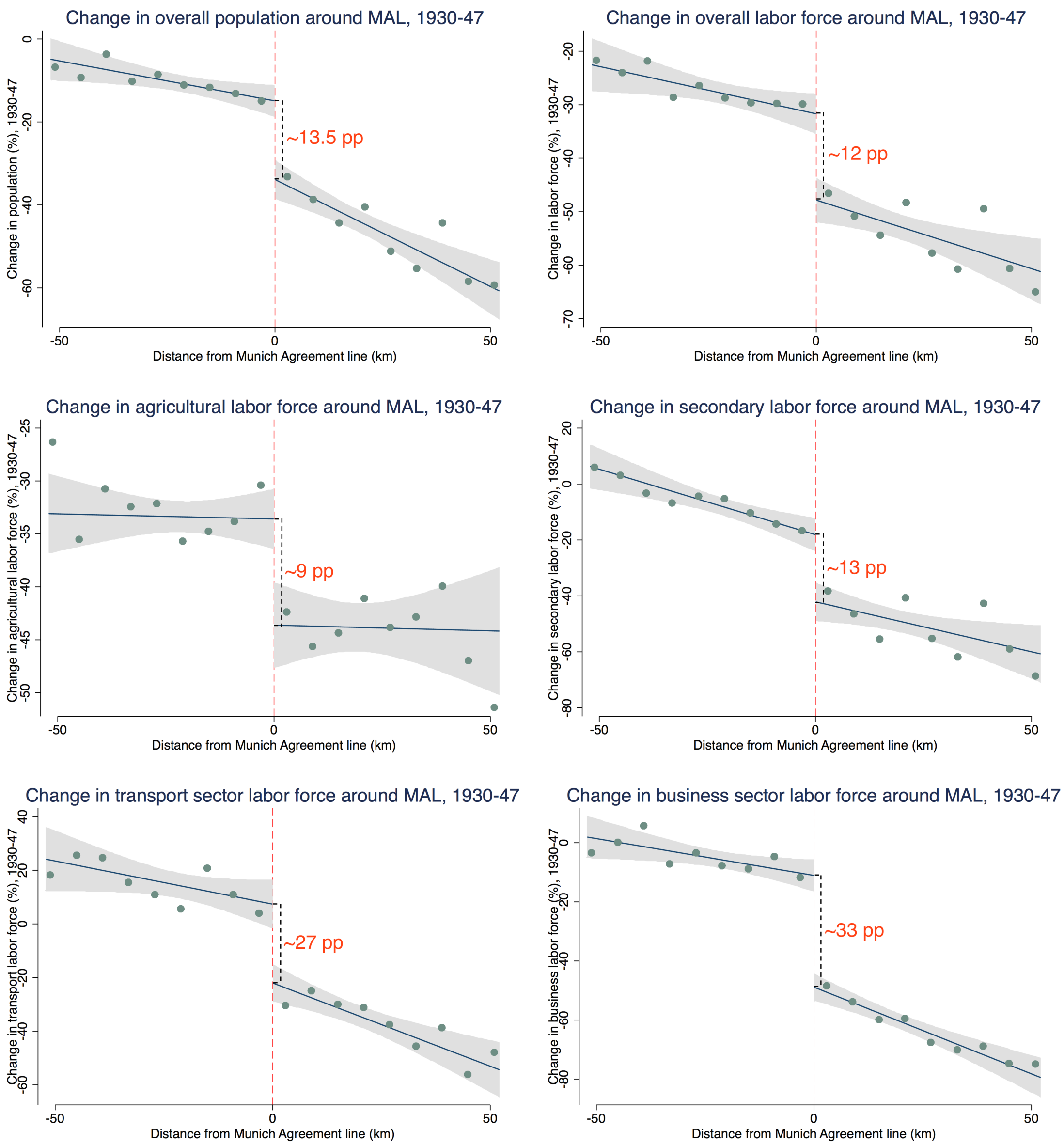


Figure 1.7: Sectoral change around the Munich Agreement line, 1930-47

that settlers were disproportionately unskilled farmers, with settlement initiatives struggling to attract skilled workers while interior areas suffered shortages in agricultural labor (Radvanovský, 2001; Gerlach, 2017). Nevertheless, there tends to be fewer differences in net losses between regions (i.e. greater convergence) as one approaches the MAL, around which most migration from the interior to the borderlands had occurred as of mid-1947.[22] Overall, this exercise provides insight into the origins of the divergence found above.[23] I will now show how these migratory patterns altered the sectoral structure of the borderlands as early as May 1947, culminating in the de-urbanization of the region we still see today.

1.5.2 De-urbanization

To see how selection in the borderlands' initial resettlement altered its relative urban development in the post-expulsion period, I apply the RD approach above to examining differences between outcomes in 1930, prior to the expulsion, and outcomes in mid-1947, after the near-total expulsion and the borderlands' initial resettlement had wound down. In particular, I adopt a "local" difference-in-differences (DD) model,

$$y_{dbt} = \alpha + \beta InBorderlands_d \times \mathbb{Y}_{1947} + f(location_{dt}) + \mathbf{X}_d' \Gamma_t + \mathbb{Y}_t + \Sigma_{bt} + \Delta_d + \varepsilon_{dbt},$$

which is equivalent to differencing (1.1) across years t, indicated by the dummy $\mathbb{Y}_t$, using time-invariant district fixed effects Δ_d while allowing the effects of the running variable $f(location_{dt})$, border segments Σ_{bt}, and natural geography $\mathbf{X}_d$ to vary over time. This estimates the effect of being *just* in the borderlands in mid-1947 *relative* to the same effect in 1930. As before, I adopt a linear interacted polynomial in distance from the MAL and a default bandwidth of 25 km, with robustness for all short-run outcomes in Table A.18 in Appendix A.

[22] While it is possible that more settlers coming from abroad or Slovakia were of those sectors, these made up only a very small fraction of settlers and would only have made the borderlands' relative declines appear smaller.

[23] For the RD estimates, see Table A.13 in Appendix A.

The first estimate of interest can be found in column (1) of Table 1.7 and shows that whereas differences in population density between the regions were statistically smooth and stable over time prior to the expulsion, population density became relatively lower in the borderlands after the expulsion and remained as such following the large-scale resettlement discussed above – even in the most easily accessible areas near the MAL. This is consistent with the historical literature, in which only 1.3 million settlers had arrived and remained by mid-1947, largely from nearby interior areas (Daněk, 1995; Gerlach, 2017). In combination with Czechs living in the borderlands pre-1938, its population was still below two thirds of its prior size. A more "global" DD analysis covering the entire period from 1921 to 2011 shows that such differences persisted on the whole throughout the communist period (see Table A.14). Moreover, they have actually grown since liberalization in 1989, suggesting that state investments under central planning may have cushioned the borderlands somewhat.

This suggests that the expulsion of the Germans triggered a de-urbanization of the borderlands. To explore this further, I consider a second important dimension of urban development: sectoral structure. Columns (2a-d) in Table 1.7 show a structural shift toward agriculture occurring alongside declines in density by the time expulsion and resettlement had concluded in mid-1947, in spite of a previous lack of differences and parallel pre-trends. And though all major sectors shrunk, the sector which bears the largest and only statistically significant relative loss is the business sector – i.e. finance, insurance, trade, and other commerce – which was highly important in northern Bohemia in 1930, including in the borderlands (see Figure 1.8).

Once again, the same patterns can also be seen "globally" in a panel covering 1921 to 2011 (see Table A.15). This shows an immediate and persistent shift toward agriculture relative to 1930 differences, with differences growing alongside a continued relative loss of industrial workers after 1950. This reflects how, after resettlement wound down in the late 1940s, migration *out* of the borderlands occurred by many skilled workers as "administrative conflicts, a lack of suitable settlers, labor shortages, and property squabbles" beset them

Table 1.7: De-urbanization around the Munich Agreement line, 1930-47

	ln Population density	Agricultural sector	Secondary sector	Transport sector	Business sector
	(1)	(2a)	(2b)	(2c)	(2d)
$\Delta_{`47-`30}$ln borderlands	-.190 (.048)***	4.046 (1.868)**	-.311 (1.810)	-.221 (.400)	-1.306 (.515)**
R^2	.919	.59	.76	.673	.635
Observations	330	330	330	330	330
Clusters	98	98	98	98	98
Border segments×1947	24	24	24	24	24
Geographic controls	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Bandwidth	25 km	25 km	25 km	25 km	25 km
Year	1930-47	1930-47	1930-47	1930-47	1930-47

Robust standard errors are clustered by political district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie, include exogenous controls for elevation, ruggedness, precipitation, temperature, and river density (km per km^2) interacted with year, and utilize a local linear running variable of distance from the Munich Agreement line interacted with the treatment and year.

(Gerlach, 2017, 14).[24] Meanwhile, service sector differences have widened even more since transition to a market economy, when state investments in steel, coal, heavy weapons, and other industries were reduced and their labor forces transitioned into sectors like banking and the auto industry (Illner and Andrle, 1994). This mirrors the findings in Table 1.5, in which borderland municipalities in 2011 had lower employment shares in skill-intensive service-based sectors.[25] Hence, the data show that the expulsion of the Germans generated an immediate urban divergence within in the Czech lands, at the expense of the former Sudetenland, which persists long after the conclusion of intervening shocks.

1.5.3 Regional human capital inequalities

Recall that there was little difference on average in literacy between districts around the MAL prior to the expulsion of the Germans. However, compositional changes associated with weak and selective migration into the post-expulsion borderlands as well as its subsequent de-urbanization may have generated differences in human capital between the regions,

[24]See Figure A.6 for a heatmap of out-migration by 1950 districts.

[25]In contrast, agriculture has tended to remain dispersed smoothly through the MAL over time, despite still being relatively more prominent in the borderlands "globally" as shown in Table A.15.

Table 1.8: Short-run educational effects, mid-1947

	Primary & lower secondary		Upper secondary		Tertiary
	Enrollment, $\text{general}_{5\text{-}14}$	Enrollment, $\text{civic}_{10\text{-}14}$	Enrollment, $\text{agricultural}_{15\text{-}19}$	Enrollment, $\text{vocational}_{15\text{-}19}$	Enrollment, $\text{college}_{15\text{-}24}$
	(1a)	(1b)	(2a)	(2b)	(3)
In borderlands	5.651	-8.204	7.174	-12.980	-2.926
	(.967)***	(2.661)***	(1.817)***	(5.15)**	(.449)***
R^2	.867	.589	.485	.241	.734
Observations	115	115	115	115	115
Border segs.	16	16	16	16	16
Geo. controls	Yes	Yes	Yes	Yes	Yes
Bandwidth	50 km	50 km	50 km	50 km	50 km
Year	1947	1947	1947	1947	1947

Robust standard errors reported in brackets, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie, include exogenous controls for elevation, ruggedness, precipitation, temperature, and river density (km per km^2), and utilize a local linear running variable of distance from the Munich Agreement line interacted with the treatment.

making it relatively more concentrated in the interior. This in turn could have fed back into de-urbanization (Moretti, 2004), generating further structural change as well as persistent human capital inequalities between the regions. Such a pattern would be consistent with the historical narrative that settlers were relatively unskilled (Radvanovský, 2001; Glassheim, 2016), as well as the finding that the borderlands is a less educated region today.

While no data exist on educational attainment for the Czech lands between 1930 and 1961, some evidence that the borderlands had lower human capital post-resettlement can be found by examining regional school enrollment patterns in mid-1947, after the main waves of expulsion and resettlement had wound down. Until the communist coup in 1948, Czechoslovakia had 8 years of compulsory education. For primary education, children attended a general school (*obecná škola*) for 5 years. Following this, one could either complete his or her education with 3 more years at a general school or pursue a more advanced lower secondary education. Commonly, this entailed attending a civic school (*měšťanská škola*) for 3 years. Civic schooling in turn was a prerequisite for subsequent upper secondary (e.g. vocational) education.[26]

Upon examining differences in school enrollment in mid-1947 between the post-expulsion, post-resettlement borderlands and nearby interior areas, I find that rates of enrollment in

[26] For more, see Greger et al (2012) and https://is.cuni.cz/webapps/zzp/download/120238971.

general schools among the relevant age group were significantly *higher* in the former relative to the latter. Given that all primary school-age children were enrolled in general schools regardless of region, this had to have been driven by differences at the lower secondary level – namely, a greater share of 10-14 year olds in the borderlands forgoing civic schooling in favor of the terminal general school track. Consistent with this, I find that civil school enrollment rates among 10-14 year olds were significantly *lower* in the borderlands in mid-1947 relative to those of nearby interior areas. These RD results can be found in Table 1.8.

This also means that there would have been relatively fewer students in the borderlands going on to pursue upper secondary and tertiary education. Indeed, column (3) confirms lower rates of college enrollment in the borderlands relative to the interior in mid-1947. Yet there may also have been differences in the *types* of schooling being pursued among those who did. To test this, I examine data from mid-1947 on enrollment for two common types of upper secondary schools: basic vocational schools (*základní odborná škola*) and agricultural folk schools (*lidová škola zemědělská*). The former provided education specific to a variety of trades, while the latter were meant specifically for those going into agriculture and horticulture.

In examining these, I observe another striking trend, which mirrors the structural changes observed above. Namely, while basic vocational schooling was significantly less popular among those living in the borderlands in mid-1947 relative to those in nearby interior areas, enrollment in agricultural folk schools was actually somewhat higher. This suggests that even among those in the borderlands who *did* go on to pursue advanced education, there was a tendency to invest in agricultural skills over other, more technical skills.

Overall, these findings are consistent with a demand-side explanation, in which those who selected into the borderlands after the expulsion of 3 million Germans had less demand for advanced and technical education, relative to those who stayed in nearby interior areas. Further evidence for this can be found in Appendix A. Table A.16 shows that, contrary to a supply-side argument, lower rates of civic and basic vocational schooling in the borderlands did not correspond to relative shortages in the number of schools or teachers per pupil. On the contrary, there appears to be a *greater* teacher-pupil ratio in the borderlands for basic vocational schools.

Hence, despite Germans and Czechs previously having similar levels of education, the expulsion of the Germans generated a loss of human capital in former-German areas beginning with compositional changes they endured during their resettlement. That being said, the exact source of this decline in demand is harder to discern. In particular, it is difficult to differentiate between this being driven by (i) a lack of skilled workers selecting into the borderlands and (ii) relatively fewer incentives to become skilled in the post-expulsion borderlands due to the relative lack of skill-intensive sectors there. More than likely there is a circular feedback between the two, with de-urbanization feeding into both and vice versa.[27]

1.5.4 Discussion of mechanisms

This section has established that the long-run differences observed around the Munich Agreement line (MAL) today can be traced back to the initial migratory responses to the expulsion, which culminated in the de-urbanization and subsequent human capital decline of the former "Sudetenland". But why did the expulsion produce this response rather than the convergent one anticipated by policymakers? The setting here rules out some of the typical explanations, such as that Germans were too difficult to replace from a labor or skill standpoint; that they lived in geographically distinct places; that there was excessive destruction of physical capital; or that there was selection as to the types of Germans being expelled.

Indeed, the answer to this is not obvious in theory, at least in the neighborhood of the MAL around which differences in fundamentals and economic activity were historically smooth. For instance, one might expect the cheap land and capital made available by the expulsion to have made the borderlands an attractive place in which to settle and invest, inducing Czechs from nearby interior areas to spill over into borderland areas in relatively

[27] Long-run trends show that these enrollment differences indeed correspond to a structural change which persists to today. See Table A.17.

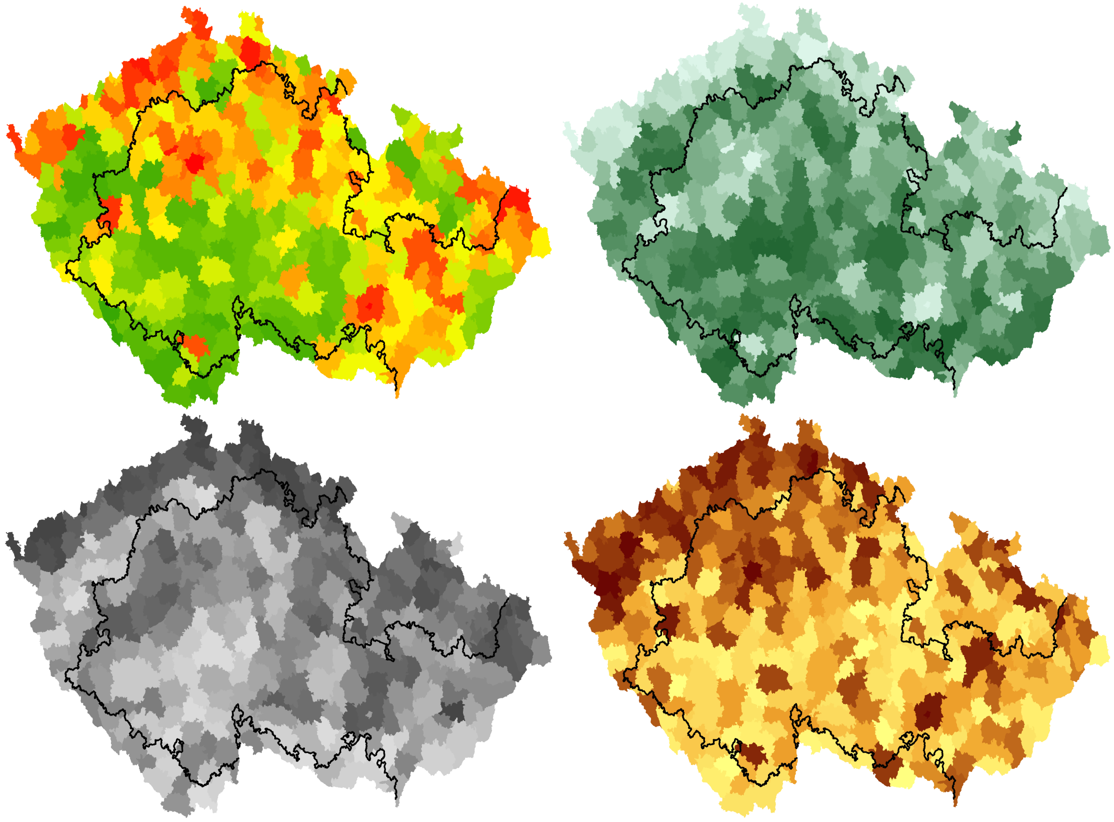


Figure 1.8: Population density, agriculture, business, industry heatmaps (clockwise), 1930

smooth ways. Instead, fewer settlers arrived to resettle the borderlands than remained in unaffected interior areas. Some settlers even returned to the interior, having not found the economic opportunity they sought. And those who did stick around were seemingly less skilled than those who did not.

I now discuss two possible forces through which the expulsion resulted in differences in economic activity around the MAL, rather than the spatial convergence initially envisioned by policymakers: (i) agglomeration economies and (ii) extractive institutions. These forces may or may not be complementary, and there may also be others, although I will argue that they are sufficient.[28]

First, the expulsion of the Germans served as a very large, negative shock to the borderlands' population, relative to nearby interior areas in which few Germans lived. In the

[28] See Testa (2019) for theory and discussion of these forces and their interactions.

presence of agglomeration economies, this would have generated a massive spatial coordination problem among the remaining population, potentially *decreasing* the expected relative return to moving to affected borderland towns for workers in urban sectors in the short-run (Bosker et al, 2007; Murard and Sakalli, 2018). This would in turn have inhibited the convergent resettlement envisioned by policymakers, helping lock in the de-urbanization of borderland areas over the longer-run (Krugman, 1991; Schumann, 2014).

Indeed, data from the 1930s suggest the presence of strong agglomeration economies throughout the Czech lands prior to the expulsion. Table A.20 in Appendix A shows that population density was strongly and positively correlated with income as well as employment in overall industry, transport, and business, and negatively with employment in agriculture and construction, with magnitudes independent of region. Figure 1.8 also shows a significant degree of localization along sector lines in 1930, with industry featuring prominently in Northern Moravia and business in Northern Bohemia, which would have been severely disrupted by the expulsion.[29]

As it turns out, the sectors most associated with density were the ones most negatively selected against during the borderlands' initial resettlement, as shown above. Importantly, these patterns first emerged prior to the communist coup of 1948. And while it is difficult to know how they might have evolved over the subsequent four decades in the absence of central planning, there exists evidence that agglomeration economies continue to matter in the post-transition period. Since 1991, net-migration into the borderlands on the whole has declined, causing overall differences in population density between the borderlands and interior to increase (see Table A.14 and A.21). And around the MAL, differences in relevant variables remain remarkably stable since transition, as discussed later.

That being said, the political economy of forced migration cannot be ignored. Indeed, the expulsion also had important implications for the institutional environment within the borderlands: when the Czechoslovak government formally sanctioned the expulsions, it ex-

[29]See Figure A.8 for more heatmaps and discussion.

propriated nearly all of Germans' property. While confiscated property was intended to be sold and transferred to incoming settlers, in the interim this vacuum of property rights was often exploited by local authorities, as well as appropriative Czechs called "gold-diggers" (*zlatokopové*), who saw the borderlands not as a place in which to settle but as a free lunch. Historians have written extensively on how this helped crowd out the reemergence of production in the borderlands and left it with a "Wild West"-like character. As a result, the borderlands often failed to attract or retain productive workers, which in turn fed into the concentration of economic activity in interior areas (Radvanovský, 2001; Glassheim, 2016; Gerlach, 2017). Hence, the extractive institutions that enabled the expulsion also afflicted the borderlands thereafter. That historically extractive institutions can impact future local development mirrors patterns observed in countless other settings (Nunn, 2007; Becker et al, 2014; Acemoglu et al, 2019).

It did not end here. The expulsion also enabled state officials to expand their political and economic powers. Prior to the 1948 coup, communists oversaw the structure of the property allocation process, which lent them popularity among some settlers. The party won a plurality of seats for the first time in the 1946 elections, due to its new base of support in the borderlands, where it won three fourths of the vote (Radvanovský, 2001). Hence, the expulsion served as a source of patronage and legitimacy for the communists. In fact, historians consider the expulsion to have been key, if not necessary, for the communist takeover (Glassheim, 2001; Tampke, 2003). From there, historians have documented how the communist regime extracted from the borderlands. The liquidation and relocation of mills, machinery, and other physical capital from under-settled borderland areas to other parts of the country aided in the decline of thousands of towns, in spite of local resistance by borderland workers (Radvanovský, 2001; Gerlach, 2017). Others were destroyed or uprooted to extract raw materials beneath them (Glassheim, 2016). Before long, "empty factory buildings [dotted] the borderlands like gravestones" (Gerlach, 2017, 258).[30]

[30]In Figure A.7, I use data from *Zaniklé obce a objekty* (zanikleobce.cz), an ongoing open-source project to document abandoned and destroyed sites in the Czech lands predominantly from the postwar period, to

Hence, the expulsion shocked the borderlands' formerly productive equilibrium in two key ways: by removing the bulk of its workforce in the presence of (i) agglomeration economies, and by engendering an (ii) extractive institutional environment there. And although gold-diggers disappeared and the Czech lands have since liberalized throughout, these two forces together can explain how the borderlands ended up on a lower development path as early as 1947, on which it still remains today.

1.5.5 Alternative channels

This paper shows a persistent spatial divergence in local development within the Czech lands following the 1945 expulsion of 3 million Germans from its borderlands. It has also discussed the origins of these patterns, which emerged immediately following the expulsion and persist decades after. That being said, there may be other forces which contributed to the persistence of or variation in such differences. In this section, I consider three additional channels.

Natural geography

The analysis thus far controls for locational fundamentals, with physical geography being smooth through the MAL. Yet some questions remain. First, some stretches of the MAL lie close to edges of the Sudete, Sumava, and Ore mountain ranges. If natural geography is too different in the borderlands relative to the interior, a spatial RD approach may poorly approximate them. As a robustness check, I consider only stretches of the MAL that do not border mountain ranges. Tables A.11 and A.12 in Appendix A show geographic balance tests and 2011 outcomes that exclude these mountainous stretches. Even when only geographically similar borderland and interior areas are compared, long-run effects remain practically unchanged.

However, locational fundamentals may nonetheless have helped mitigate the long-run

illustrate this phenomenon relative to the MAL.

effects of the expulsion *within* the borderlands. Since access to water is considered an important determinant of economic geography (Rappaport and Sachs, 2003), I expand the analysis by interacting the treatment variable with a mean-normalized river density control, which indicates kilometers (km) of waterways per square km in each municipality. Although the long-run effects on unemployment, population density, and educational attainment remain strong and significant, I find evidence that sufficiently high local river density can attenuate them. This can be found in Table A.22. However, note that this does not eliminate the treatment effect, as 95% of municipalities in the main specification have fewer than 0.9 km per square km of river greater over the mean. I find no such effect for ruggedness.

Central planning

The expulsion and the borderlands' resettlement occurred prior to the communist coup of 1948, when Czechoslovakia's labor economy still operated to a large extent via the market mechanism (Bernàšek, 1970). However, it is important to consider to what extent central-planning institutions were responsible for *preserving* this divergence post-1947. If signs of convergence are observed as having occurred around the MAL since transition, then it might mean that this persistence had more to do with active central-planning decisions, whereas e.g. losses in density in the borderlands may not have persisted to the same extent locally in a market economy.

To test this, I consider in Table A.23 trends in relevant outcomes between 2001 and 2011, after the end of communism and the restructuring of the 1990s. Differences in density and unemployment remain remarkably stable near the MAL. I find similar stability for sector shares, with the exception of the construction sector, which declined somewhat *more* as a share of the labor force in the borderlands relative to the interior during the period.

Thus, despite large-scale structural change since transition, the overall *magnitude* of the spatial divergence remains remarkably stable. This downplays the importance of central-planning institutions. That being said, the institutional environment is an important part of expulsion and plays a key role here, as discussed above.

Geopolitics

Economic geography may interact with political geography. After WWII, the Czechoslovak government intentionally kept parts of the borderlands near the West German border emptied for use as a military buffer zone (Illner and Andrle, 1994). Although these areas were closer to international borders than to the MAL, it is possible that once it was clear the borderlands would be under-settled, geopolitical concerns would have influenced other central-planning decisions, harmed trade networks, or altered foreign investment in former-German areas more broadly. It is therefore of interest to consider how stretches of the borderlands near the MAL that were closer the Eastern Bloc fared relative to stretches near West Germany and Austria. Table A.24 shows several of the baseline long-run regressions with an interaction term for if a municipality was closer to the Eastern Bloc prior to 1989. Interestingly, the interaction generates no significant effect.

1.6 Conclusion

This paper considers the long-run implications of Czechoslovakia's postwar German expulsions for the relative development of former-German areas. Using a spatial regression discontinuity design on the border of the "Sudetenland" region where Germans lived – today known as the Czech borderlands – I show that the expulsion had immediate and long-lasting consequences for these areas relative to nearby places in the Czech interior. I then show that these patterns first emerged from a selective resettlement following the expulsion, which culminated in the region's de-urbanization and relative human capital decline. Using data and historical evidence, I propose agglomeration economies and extractive institutions as two compelling forces through which the expulsion generated this outcome.

Much of this analysis has focused on how effects *emerged* out of the borderlands' resettlement after the expulsion. In contrast, more work is needing to connect the short- and long-run outcomes documented in the paper and address the numerous questions that remain. For example, what role did communism play in driving persistence relative to what would have persisted regardless? On one hand, the relative stability of trends around the former German "language border" today suggests that central planning was unnecessary *ex post* for persistence. On the other hand, historians have documented how the communists extracted capital and raw materials from the borderlands following its initial de-urbanization, as discussed in Section 1.5. As such, it is unclear how effects might have differed had Czechoslovakia been more like Austria or Finland during the Cold War. Future work should examine how the communists influenced the first- and second-nature geography of the borderlands and on what margins this mattered for long-run persistence (i.e. post-transition). This would tell us more about both the channels of persistence in this particular setting, as well as the extent to which policy can influence long-run development more generally. Future research should also examine and compare the long-run effects of forced migration in a variety of institutional settings, including forced migration as a result of non-political factors such as climate change. Finally, household-level data collection and analysis is needed. A next step should entail examining narrow areas around the former language border, as in Karaja and Rubin (2017), to better understand the choice to resettle the borderlands as well as vertical and horizontal transmission mechanisms at play at a more micro level.

This paper provides several important lessons for understanding forced migration, at a time when the number of forcibly displaced worldwide is approaching 70 million. First, it illustrates how such events may affect not only targeted groups but also have long-term implications for development and regional inequality within the origin economy. Moreover, it shows that forced migration can have strong and persistent effects even when displaced populations are compositionally similar to those remaining. Other factors, such as agglomeration and institutional forces, may be important for explaining why affected areas do not recover their former relative statuses. This suggests that expelling even relatively disadvantaged minorities may have negative effects. This is no small finding. While previous research has focused on expulsions involving high-skilled minorities, forced migration has plagued groups from many backgrounds throughout history. This should give both dictators and legislators pause when considering the exclusion of even relatively low-skilled minorities, as doing so may leave a lasting mark on the places left behind.

Chapter 2

Shocks and the Spatial Distribution of Economic Activity: The Role of Institutions

2.1 Introduction

What factors determine the distribution of economic activity across space? And why do some shocks to local development permanently alter this distribution, while others do not? These questions are of central importance in urban and development economics for understanding differences in economic performance within countries, as well as the potential role of policy. Despite this, their answers remain subject to active debate.

One theory is that there exists the potential for multiple equilibria in spatial development but conditions set by both nature and history select among them. In this view, the location of economic activity is driven in part by incentives for humans to locate near each other, such as in production (i.e. agglomeration spillovers). Such increasing returns can generate path dependence, while also implying a potential for policy to induce or transplant economic activity in self-reinforcing ways (Kline and Moretti, 2014; Jedwab and Moradi, 2016). However, some empirical research has cast doubt upon the empirical relevance of multiple equilibria. Davis and Weinstein (2002, 2008) and Miguel and Roland (2011), among others, have shown how even massive shocks may only temporarily redistribute economic activity across space. This literature supports a more deterministic view, in which individuals co-gravitate toward strong fundamentals over the long-run, while returns to scale matter more for determining spatial dispersion (e.g. of cities). Efforts to reconcile these findings have varied considerably, with selection in shock exposure, focal points, and heterogeneity in physical geography all being proposed as potential sources of differential effects (Redding, 2010; Acemoglu et al, 2011; Bleakley and Lin, 2012; Nunn, 2009, 2014; Jedwab et al, 2019).

This paper provides an alternative approach to understanding this empirical puzzle, by considering the interaction of increasing returns with another important force for long-run development: formal institutions (Acemoglu et al, 2001; Dell, 2010; Acemoglu and Dell, 2010). Using a two-region model with migration between regions, I explore the role of institutions in explaining the differential impact of temporary shocks on the long-run spatial distribution of economic activity. In the model, more extractive institutions decrease the return on production relative to "unproductive" activities that do not contribute to the productive process, thus utilizing resources at the expense of it (Nunn, 2007). In the presence of increasing returns to productive activity within regions, a large negative shock to a region's population can temporarily reduce productive spillovers. When institutions are sufficiently extractive, this can induce substitution among workers from productive into unproductive activities. Now absent productive spillovers, relatively fewer workers will prefer to live in the affected region, while those who do migrate there will also prefer to engage in relatively unproductive activities, locking in asymmetries in both population and production between regions.

Hence, the model exhibits multiple equilibria: one with two similarly productive and populated regions, and one with a single highly populated, productive region neighboring a less populated and relatively unproductive region. Moreover, these asymmetries can arise even when there are no differences in natural advantages, local institutions, or endowments *ex ante* between regions. Then, as institutions become stronger and less extractive, spatial equilibria become more robust to large shocks. This illustrates how relatively low levels of economic activity may persist in a region following a negative population shock, causing population and productive inputs such as human capital to become concentrated in select regions over the long-run – even if formal institutions eventually improve.

The notion that historical institutions can have long-lasting effects is not new. A large theoretical and empirical literature exists documenting numerous cases throughout history in which extraction negatively impacted long-run economic development. Human capital (Acemoglu et al, 2014), culture (Tabellini, 2010), and public goods provision (Dell, 2010) have all been cited as important channels through which historical institutions continue to matter. Most similar to this paper is Nunn (2007), which assumes a similar tradeoff between productive and unproductive activities in explaining the importance of historical extraction. This paper goes a step further, exploring how national institutions influence the persistence of shocks, and therefore the distribution of economic activity, *within* countries. In particular, it argues that in places that feature less economic activity, extractive institutions promote comparative advantages in unproductive activities that, as such, do not attract productive workers. In the context of a large shock, such as a natural disaster, this means that activities such as corruption and property theft made more attractive by weak institutions are present to reinforce the effects of the initial shock. It also means that, by weakening incentives underlying urban recovery in the short-run, weak central institutions may produce greater variation in development within countries.

Nevertheless, a link between the institutional environment experienced by a country or region and the persistence of population shocks therein has often been alluded to in existing empirical research on war, expulsion, and natural disaster. Mirroring Davis and Weinstein's (2002, 2008) finding that Japanese city size and composition were robust to the bombings of WWII, returning to their prewar distributions within decades, Miguel and Roland (2011) observe similar convergence in Vietnam. At the same time, they argue that differential convergence would be unsurprising in a larger sample of studies. In particular, the authors note that while postwar Japan was a market democracy and Vietnam a socialist regime, both had relatively strong political institutions, which would have aided in catch-up in both places. Similar points about the importance of preexisting institutions are made by Brakman et al (2004), who find swift convergence after WWII in West but not East Germany, as well as in surveys of the empirical literature by Redding (2010) and Nunn (2014).

Given this, it is perhaps unsurprising that much of the work on forced migration has shown, in contrast, strong persistence in the origin economy. For instance, Chaney and Hornbeck (2016) find delayed convergence following the expulsion of Moriscos from Spain in 1609, citing preexisting extractive institutions in Morisco areas as a potential source. Testa (2019) similarly finds Czech municipalities affected by expulsions of Germans after WWII to be worse off today relative to unaffected areas nearby, attributing these differences in part to the widespread property exploitation that took place of affected areas by settlers and local officials. Meanwhile, Nunn (2008) finds a negative relationship between exports of slaves and future economic performance in African countries, characterizing the slave trade as an extractive regime that gave rise to raiding and internal warfare in origin economies.

Such can also be found in the relatively smaller literature on non-political population shocks, such as natural disaster. In the case of earthquakes, Barone and Mocetti (2014) cite preexisting institutions as a source of differential effects, with corruption and declining social capital impeding recovery in poorly institutionalized places, while Anbarci et al (2005) similarly show poor collective action to exacerbate earthquake fatalities in places with greater inequality, and Belloc et al (2016) observe local institutional stagnation following earthquakes in medieval Italy in places where separation of relevant powers had previously been weak. Meanwhile, Acemoglu et al (2019) find that severe droughts paved the way for the Sicilian Mafia where institutions were weak, at the expense of subsequent local economic development.[1] That being said, this is but a subset of papers in a very large literature, and many other factors may explain differential persistence, such as selection in the initial shock (Acemoglu et al, 2011) as well as various focal points in nature and history aiding in

[1] Also see Maloney and Caicedo (2015) and Jedwab et al (2016) for examples of institutions as a source of heterogeneous effects in the persistence of pre-colonial American agglomerations and the Black Death, respectively, as well as Dell and Olken (2019) for evidence that within the extractive colonial Dutch regime in Java, agglomeration economies from sugar factories gave rise to countervailing long-run effects locally.

post-shock recovery (Jedwab et al, 2019). I therefore go now to the model, before concluding in Section 2.3.

2.2 The model

The economy in the model is composed of a share of non-atomic workers M_r in each region $r \in \{1, 2\}$. Workers are long-lived but myopic. I therefore focus for now on a single time period. Each worker begins a period with some endowment, which she may choose to transform into a *labor input*, h (e.g. human capital). If she does, then her labor input is combined with a firm's resources to produce goods, and she is compensated at the regional market wage rate w_r. In this scenario, she is said to be *engaged in production.* At any given time, the share of all workers living in region r *and* engaged in production is $m_r \leq M_r$.[2] Each region also has a fixed stock of *resources*, K, which are divided amongst a mass $\lambda * m_r$ of identically-producing firms, indexed by ω.[3] In a given period, each region r firm has some amount $k_r \leq K/\lambda m_r$ of resources for use in production, to be defined shortly.

However, a worker may also choose to forgo engagement in production. In this scenario, she simply consumes resources directly – resources which might otherwise be used by firms as inputs in production. I refer to such behavior as *unproductive*, to the extent that it does not contribute to the local production process and as such comes at its expense. The relative payoffs from unproductive activities as compared to productive ones crucially depends on the formal institutional environment. The assumption that extractive or weak institutions decrease the relative payoffs from productive activities and give rise to unproductive behavior is long-standing in the political economy literature (Skaperdas, 1992; Nunn, 2007). In this model, the *quality of institutions* is exogenous to local economic activity and represented simply by the parameter β.

[2]Thus the share of the world's workforce that is engaged in production is $m_r + m_{-r} \leq 1$.
[3]For simplicity, K is immobile and regenerates each period.

The productive environment

Production is subject to constant returns to scale within firms in resources and labor inputs. However, the model allows for *external* increasing returns (i.e. within regions) in regional labor inputs $H_r \equiv m_r h$. In this case, as the relative amount of productive workers in region r increases, its firms can produce relatively more given the same labor inputs. This *agglomeration spillover* is represented by H_r^γ, where $\gamma \geq 0$ gives its magnitude.[4]

Besides agglomeration, heterogeneity across regions in firm-level productivity can also be attributed to differences in *natural advantages*. These locational benefits are given by the parameter a_r. Thus the overall productivity level for region r is given by:

$$A_r = a_r H_r^\gamma.$$

Altogether, this yields a firm-level CRS production function of:

$$f_r(\omega) = A_r k_r h_r(\omega),$$

where $h_r(\omega)$ gives a region r firm's demand for labor inputs. Hence, a firm's profit maximization problem can be given by:

$$\max_{h_r(\omega)} p a_r H_r^\gamma k_r h_r(\omega) - w_r h_r(\omega), \tag{2.1}$$

where prices p are set collectively by all regions with workers engaged in production.[5] Assuming zero profit, this implies a real income and thus consumption payoffs for productive

[4]This term is common in the economic geography literature (Allen and Donaldson, 2018). For microfoundations, see Marshall (1920) and Duranton and Puga (2004). For empirical estimation, see Rosenthal and Strange (2004), Moretti (2004), Greenstone et al (2010), and Ellison et al (2010).

[5]Since regions are in close proximity, I assume no trade costs or differences in market access.

workers in region r of:

$$
\begin{aligned}
V_r(h) &= a_r H_r^{\gamma} k_r * h \\
&= a_r k_r h^{1+\gamma} m_r^{\gamma}.
\end{aligned}
\tag{2.2}
$$

How does the institutional environment matter?

Unlike typical two-region models of economic geography, this framework incorporates a strategic component in which workers may prefer to engage in unproductive activities. In contrast with production, which involves combining worker endowments with resources to create value for consumers, unproductive activities involve acquiring and consuming resources directly (i.e. every man for himself), which does not entail external economies of scale, while coming at the expense of the local productive sector, which does. This distinction between productive and unproductive activities is common in the literature on institutions and conflict (Acemoglu, 1995; Nunn, 2007). Real world examples might include corruption and rent-seeking, looting and property crime, and free-riding on public goods.

I model the acquisition of resources using a variant of the contest success function (Skaperdas, 1996). Unproductive workers consume resources that would otherwise be used in production, where the total amount of resources acquired by unproductive workers in region r is proportional to the relative prevalence of unproductive behavior in the regional economy:

$$
\frac{M_r - m_r}{M_r} K,
$$

where $M_r - m_r$ equals the share of all workers living in region r and engaged in unproductive activities. This leaves each region r firm with a final resource endowment of:

$$
\begin{aligned}
k_r^* &= \left(1 - \frac{M_r - m_r}{M_r}\right) \frac{K}{\lambda m_r} \\
&= \frac{K}{\lambda M_r}.
\end{aligned}
\tag{2.3}
$$

Unproductive payoffs follow from this. Adopting the assumption that relative value from unproductive activities is derived inversely from the quality of institutions, consider the following payoffs from engaging in unproductive activities in region r:

$$\begin{aligned} V_r(u) &= \frac{1}{\beta}\frac{1}{M_r - m_r}\left(\frac{M_r - m_r}{M_r}\right)K \\ &= \frac{K}{\beta M_r}. \end{aligned} \tag{2.4}$$

Recall that β describes the quality of institutions, which I consider to be a deep parameter that is the same in both regions. In spite of this, unproductive behaviors may become widespread in one region and not the other, as one will see shortly. At the same time, because resources are fixed in each region and of use in both production and unproductive activities, they will serve as a relative congestion force in each region that prevents "black hole" equilibria, in which all workers locate in one region, from arising in the long-run.

2.2.1 Short-run equilibria

I assume that in the short-run, workers cannot move between regions but can move between productive and unproductive activities. For analytical simplicity, this choice is modeled as a simple binary decision. That is, an agent prefers to transform her endowment into a productive labor input if and only if

$$a_r \frac{K}{\lambda M_r} h^{1+\gamma} m_r^{\gamma} \geq \frac{K}{\beta M_r}. \tag{2.5}$$

For now, let $\gamma > 0$. Since agents are non-atomic, each takes m_r as given when deciding whether to deviate. Hence, worker behavior exhibits strategic complementaries around some critical threshold, $\widehat{m_r}$, above which the optimal $m_r^* = M_r$. To describe this outcome, I define the following:

Definition 2.2.1. A *high production short-run equilibrium* [HPSE] consists of all workers in a region r specializing in production ($m_r^* = M_r$).

Definition 2.2.2. A *low production short-run equilibrium* [LPSE] consists of all workers in a region r specializing in unproductive activities ($m_r^* = 0$).

Now consider the first result:[6]

Proposition 1. *There exists a high production short-run equilibrium [HPSE] for each region r. In every HPSE:*

(i) There is a threshold relative amount of productive workers $\widehat{m_r}$ when $\gamma > 0$, above which all workers in region r, M_r, prefer to specialize in production, $m_r^* = M_r$.

(ii) The share of all workers located in r must be sufficiently large, $M_r \geq \widehat{m_r}$, where this equilibrium is locally stable in M_r whenever this inequality is strict.

(iii) $\widehat{m_r}$ is decreasing in a_r, h, and β and increasing in λ.

The remaining space below $\widehat{m_r}$ is characterized by a LPSE, in which all agents in region r forgo production and instead simply acquire and consume local resources (i.e. $m_r^* = 0$). Importantly, since the productive population of a region is constrained by the size of its total worker population, a temporary decrease (i.e. shock) in the share of the population in r has the potential to permanently shift a region from a HPSE to a LPSE (i.e. $M_r < \widehat{m_r}$ implies $m_r < \widehat{m_r}$). However, this depends on the quality of formal institutions. When the quality of institutions β is sufficiently high, even large shocks will not generate incentives for workers to substitute toward unproductive activities in the affected region. This is important, as a population shock which cannot induce a shift from one short-run equilibrium to another within a region will also have no effect on the long-run equilibrium population distribution *across* regions, as we will see shortly.

Now let $\gamma = 0$, such that there are no agglomeration spillovers. This is relevant for understanding how sectors such as agriculture respond to population shocks in the short-

[6] All proofs are in Appendix B.

run. As it turns out, population shocks cannot shift a region from one short-run equilibrium to another in the absence of agglomeration spillovers:

Remark 1. *In the absence of agglomeration spillovers, $\gamma = 0$, if a HPSE exists in region r for some M_r, then it exists for all M_r'.*

Hence, the propensity for a population shock to shift a regional economy from a HPSE to a LPSE depends not only on the quality of formal institutions but also on the presence of external increasing returns (i.e. $\gamma > 0$), which generate strategic complementarities in production choices within regions. In fact, absent agglomeration spillovers, economic activity will always tend toward its initial distribution as determined by fundamentals *regardless of institutions.* To show this, however, I must first introduce population dynamics, in the form of migration between regions over time.

2.2.2 Long-run equilibria

In the long-run, agents can move between productive and unproductive activities within regions as well as migrate between regions. Population dynamics are modeled as in the economic geography literature,[7] using a standard replicator dynamic:

$$\dot{M}_r = M_r(V_r - \overline{V}), \tag{2.6}$$

where $\dot{M}_r$ gives the change over time in the share of the population in region r, which depends on the relative size of the short-run payoffs in region r, and where $\overline{V} \equiv M_1 V_1 + M_2 V_2$ gives the national average payoffs. There is no cost to migration. However, since agents are non-atomic, short- and long-run incentives can interact to generate coordination problems which in turn constrain migration.

[7] In this tradition, population dynamics are often framed in relative terms (i.e. regional shares), such that the size of the total population does not matter for the long-run analysis. For instance, see Krugman (1991), Davis and Weinstein (2008), and Allen and Donaldson (2018).

Suppose, for instance, that $m_r \geq \widehat{m_r}$ initially in each region r, such that both regions specialize in production (i.e. $m_r^* = M_r$ for all r). Then there exists some steady state $M_r \equiv M_r^*$ at which $\dot{M}_r = 0$ as long as $M_r^* \geq \widehat{m_r}$ for each region r. That is, when enough agents are coordinating on productive behavior in each region, $m_r \geq \widehat{m_r}$, there is some population distribution M_r at which both regions have high levels of production and no worker prefers to deviate from one region to other. From (2.2) and (2.6), this is the solution to:

$$\frac{a_1 K h^{1+\gamma}}{\lambda} M_1^{\gamma-1} = \frac{a_2 K h^{1+\gamma}}{\lambda} M_2^{\gamma-1}, \tag{2.7}$$

which implies for each region r:

$$M_r^* = \frac{a_r^{\frac{1}{1-\gamma}}}{a_r^{\frac{1}{1-\gamma}} + a_{-r}^{\frac{1}{1-\gamma}}}.$$

However, the local stability of this as a long-run equilibrium depends on γ.

Assume for now that $\gamma \in (0, 1)$. When $\gamma \in (0, 1)$, the lefthand side of (2.7) is strictly decreasing in M_r while the righthand side is strictly increasing. Then small changes in M_r will have only temporary effects, holding short-run equilibria fixed. However, by Proposition 1.2, the stability of this state also depends on the size of M_r relative to the threshold $\widehat{m_r}$ for each region, i.e. local stability in the short-run. I thus define the following:

Definition 2.2.3. A *symmetric high production long-run equilibrium* [HPLE] consists of (i) each region being in a HPSE ($m_r^* = M_r$ for all r), with (ii) a steady state share of workers located in each region, M_r^*, which is locally stable in M_r in both the short- and long-run.[8]

Now suppose that M_r and $\widehat{m_r}$ are sufficiently close. Then a large, negative shock to M_r in the short-run (e.g. from death or displacement) may result in a shift to a LPSE in region r, such that the steady state population distribution is no longer determined by (2.7).

This brings me to a second case, in which a large, negative shock to e.g. M_2 occurs,

[8]Formally, local stability requires that small population shocks around M_r^* are temporary ($\frac{\partial \dot{M}_r}{\partial M_r}|_{M_r = M_r^*} < 0$), as well as that short-run equilibria be locally stable in M_r for all r ($M_r^* > \widehat{m_r}$).

shifting region 2 from a HPSE to a LPSE. In other words, in depleting region 2 of its productive workforce relative to region 1, a large negative population shock reduces its productive spillovers, thus making it relatively more appealing for those living in region 2 to engage in unproductive behavior, so long as institutions are sufficiently extractive. Furthermore, conditional upon engaging in unproductive activities, it also increases the consumable amount of resources *per capita* in region 2. In the long-run, this will trigger migration into region 2 by those who see opportunity in unproductive activities, but *not productive ones.* Assuming that such a shock did not also occur in region 1, the new steady states M_r will be the solution to:

$$\frac{a_1 K h^{1+\gamma}}{\lambda} M_1^{\gamma-1} = \frac{K}{\beta M_2}. \tag{2.8}$$

It can be shown that such a shock should leave region 2 at a permanently lower relative population level when $\gamma \in (0,1)$, as in Figure 2.1. To do this, I first define the following:

Definition 2.2.4. An *asymmetric long-run equilibrium* [ALE] consists of (i) one region r being in a HPSE ($m_r^* = M_r$) and (ii) the other region $-r$ being in a LPSE ($m_{-r}^* = 0$), with (iii) a steady state share of workers located in each region, M_r^{**}, which is locally stable in M_r in both the short- and long-run.

Altogether, these results can now be summarized by the following proposition:

Proposition 2. *(i) There exists a locally stable symmetric high production long-run equilibrium [HPLE], with a unique interior steady state population $M_r^* \in (0,1)$ when agglomeration spillovers are moderately strong, specifically $\gamma \in (0,1)$, where $M_r^* = \frac{1}{2}$ if and only if $a_2 = a_1$, and where M_r^* increasing in a_r and decreasing in a_{-r}.*

(ii) There also exists a locally stable asymmetric long-run equilibrium [ALE], with a steady state population share in the productive (unproductive) region of $M_r^{**} > M_r^*$ ($M_{-r}^{**} < M_{-r}^*$), with M_r^{**} increasing in a_r, h, and β and decreasing in λ.

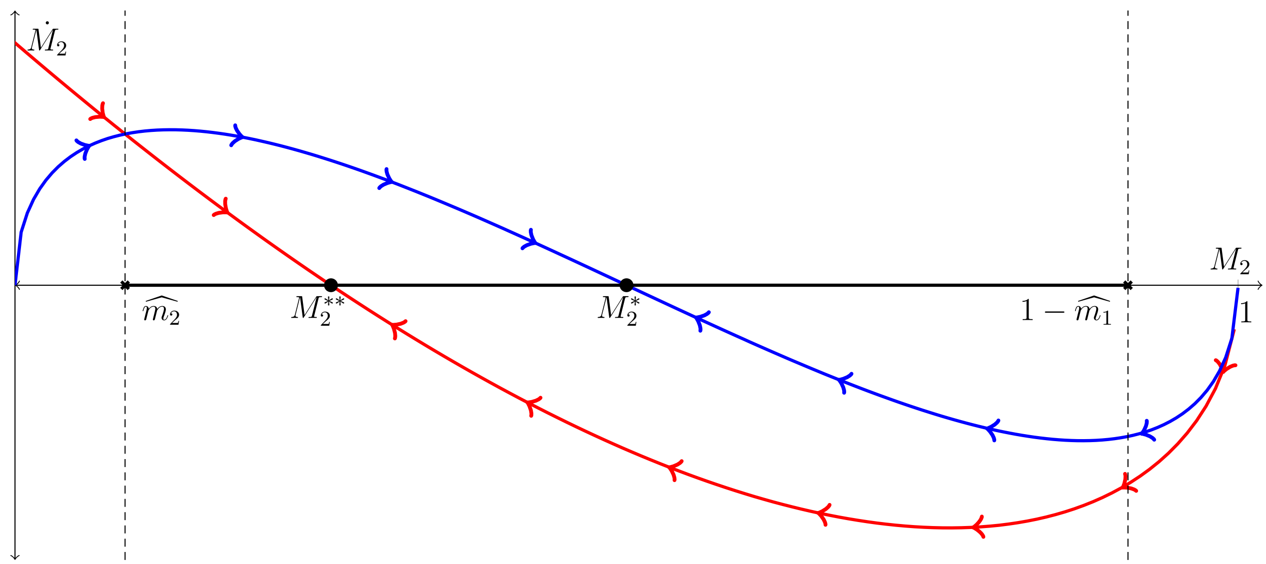


Figure 2.1: Symmetric high production versus asymmetric long-run equilibria

Thus, a sufficiently large, negative population shock in one region (i.e. such that $M_r < \widehat{m_r}$) can permanently (i) shift its local economy from production to unproductive activity, (ii) lowering its relative population due to now relatively larger productive spillovers in the other region, (iii) leaving a population that is nonetheless positive to the extent that its resources may still be utilized in relatively unproductive ways, with such payoffs being determined by the quality of institutions (i.e. $\frac{\partial \widehat{m_r}}{\partial \beta} < 0$).[9]

In other words, given more extractive institutions, large-scale population loss tends to induce a shift toward unproductive activities in the affected region by those remaining as well as incoming migrants (e.g. property exploitation, corruption), rendering it less productive and populated over the long-run. Stronger institutions, meanwhile, limit the extent to which being production becomes relatively unappealing following large shocks, such that agents are more likely to coordinate back to pre-shock patterns.[10]

Finally, let $\gamma \notin (0, 1)$. We know that when $\gamma = 0$, short-run shocks of any size should have

[9]Hence, if a shock also negatively affects K over the long-run in that region, even fewer would reside there.

[10]As a corollary to Propositions 1 and 2, note that if institutions become too extractive, then unproductive activities will become too appealing relative to productive ones, such that no HPLE or ALE can survive and both regions will be in LPSE as part of a globally stable symmetric low production long-run equilibrium [LPLE]. To see this, note that the derivative on $\widehat{m_r}$ and those on both steady states M_r^* and M_r^{**} are opposing in β. For sufficiently low β, $M_r \geq \widehat{m_r}$ can never occur, regardless of shocks.

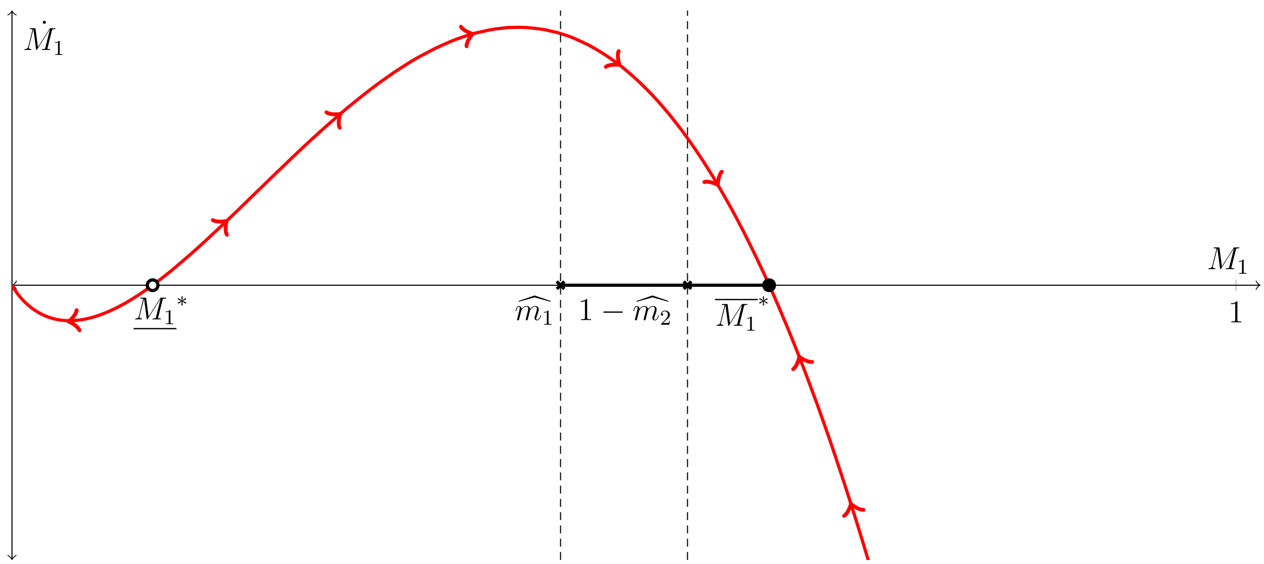


Figure 2.2: Asymmetric long-run equilibria with strong agglomeration spillovers

no bearing on short-run equilibria. Hence, sectors lacking agglomeration spillovers should see their workers return to their pre-shock distribution, as determined by either (2.7) or (2.8).

Remark 2. *In the absence of agglomeration spillovers, $\gamma = 0$, population shocks have no long-run effect on the spatial distribution of productive activity.*

In other words, in sectors like agriculture which lack external economies of scale, shocks have no permanent effects, regardless of institutions. Rather, the distribution of productive activity is determined solely by the fundamentals. If fundamentals are sufficiently pro-production, then activity will tend to locate more where they are stronger. If they are symmetric across regions (i.e. $a_2 = a_1$), then so will be the distribution of e.g. farmers, both before and after population shocks, assuming a HPLE to begin with.

What about when agglomeration spillovers are very strong, i.e. $\gamma > 1$? As it turns out, when productive spillovers are sufficiently great, all HPLE become unstable. This means all of the productive population will be located in one region or another. Whichever region has its worker population drained by such forces will eventually specialize in unproductive activities. This generates a single locally stable long-run equilibrium with one productive region and one unproductive region, as before:

Proposition 3. *When agglomeration spillovers are sufficiently strong, specifically $\gamma > 1$, then:*

(i) There are no locally stable HPLE.

(ii) There are at most two interior ALE, with steady state populations in the productive region r of $\underline{M_r}^* \in (0, \frac{\gamma-1}{\gamma})$ and $\overline{M_r}^* \in (\frac{\gamma-1}{\gamma}, 1)$.

(iii) Only the second, more populated steady state $\overline{M_r}^*$ can be locally stable.

Hence, the existence of a locally stable HPLE is sufficient but not necessary for uneven spatial patterns of development to arise. When agglomeration spillovers are sufficiently strong, unevenness will always arise from initial conditions (or historical conditions from prior to the emergence of spillovers) and remain as such forever.

The role of local development policy

Finally, note that in contrast with negative shocks, the model implies that the effects of temporary local development policies and positive population shocks associated with them (assuming an unproductive equilibrium to begin with) may actually be *more* likely to persist in the long-run in places with strong institutions. This is because while it takes a negative shock to m_r larger than $M_r - \widehat{m_r}$ to move a region from a productive equilibrium to an unproductive one, it takes a positive policy shock larger than $\widehat{m_r}$ to do the opposite. Hence, when β is large, only a small-scale coordinated effort (e.g. by some "city corporation") is needed to move a region from an unproductive equilibrium to a productive one: an investment need only attract a few to the region simultaneously before complementarities take over.

2.3 Conclusion

Why do some shocks to local development permanently impact the spatial distribution of economic activity, while others do not? This paper considers the role of formal institutions in explaining these differential effects. In the model, extractive institutions decrease the return on productive relative to unproductive activities. In the presence of increasing returns to productive activity within regions, a sufficiently large, negative shock to one region's labor force can serve as a tipping point, inducing its workers to substitute into unproductive activities. Thereafter, new migrants will also prefer to engage in unproductive activities, resulting in regional asymmetries in population and production levels over the long-run. This suggests that extractive institutions may magnify the importance of increasing returns for long-run local development, while long-run equilibria may be more robust to large negative shocks where there are strong institutions to help coordinate the reemergence of production.

I end with a few remarks. First, it is important to note that the choice between productive and unproductive activity is, in reality, not binary. Rather, the prevalence of the latter will depend on its relative return as determined by the institutional environment (i.e. β). As institutions improve, it is likely that even unproductive regions would become more developed. Nor is this stylized model sufficient to explain all differences in development observed within countries. A richer model is needed to capture the nuanced interactions between institutions, agglomeration spillovers, natural geography, infrastructure, etc. That being said, this model illustrates in simple terms how formal institutions can influence the importance of increasing returns in the face of large, negative population shocks. In particular, it suggests that spatial equilibria may be more subject to influence by negative shocks in places and times in which institutions are weak, becoming more robust as they improve.

Chapter 3

Education and Propaganda: Tradeoffs to Public Education Provision in Nondemocracies

3.1 Introduction

Why do nondemocracies invest in public education? Doing so promotes human capital acquisition and social cohesion, favoring economic development and expanding the tax base.[1] Yet the same forces may also promote political development, enhancing political sophistication and participation.[2] Indeed, education has historically been restricted to certain groups in expectation of this. Many U.S. states banned the education of slaves prior to the Civil War, while Colonial powers in Africa and East Asia often adopted similar measures against indigenous peoples, even though the resulting productivity gains would have benefitted economic elites in both cases (Woodson, 1915; Bjork, 2005). Given this, the extent to which many modern nondemocracies invest in education is puzzling.[3]

Nevertheless, another view holds that nondemocratic regimes use public education to

[1]See Gradstein and Justman (2000, 2002), Glaeser et al (2004), and Galor and Moav (2006).

[2]See Lipset (1959), Dee (2004), Milligan et al (2004), Yanagizawa-Drott (2014), Chong and Gradstein (2015), and Friedman et al (2016). Some argue that education leads to democracy. For this debate, see Acemoglu et al (2005) and Castelló-Climent (2008).

[3]Lott (1999) and Bursztyn (2016) note positive associations between nondemocracy and various measures of public education provision among lower income countries. Aghion et al (2019) find this trend more generally, although Acemoglu et al (2015) suggest a reversal at the secondary level. De la Croix and Doepke (2009) do not observe this trend, but find variance in education spending to be increasing in non-freedom, with non-free countries dominating the right tail.

shape public opinion and control civil society. Early compulsory education systems in Europe and North America were designed principally to unify historically disparate groups around new national identities in the process of nation-building (Alesina et al, 2019; Bandiera et al, 2017). Likewise, 20th century authoritarian states often used public schools to promote compliance with autocratic power structures and state-sanctioned ideologies.[4]

This paper brings these two views together, treating them as separate dimensions of public education that interact in important ways. In doing so, it generates new insight into why many nondemocracies choose to invest in public education when doing so can promote political development. In particular, this paper explores how a ruler can alter the incentives underlying public education provision by manipulating the content of education. To do this, I draw from a vast literature from across the social sciences, which conceives of political engagement as something learned and habituated early in life.[5] In this view, public education promotes political participation to the extent that it tends to facilitate involvement in political and other social activities. However, this effect varies with the style and content of education.

With that in mind, I develop a model examining a ruler's decision to invest in public education. With some probability the ruler has predatory objectives, seeking only to maximize his own rents. Indeed, providing education may increase his taxation revenues. Yet education also enables citizens to become active in political "clubs." In the model, this promotes acquisition of inputs needed for citizens to credibly oppose high taxes in the future (i.e. political participation).

A ruler is not naive to this: if he invests in education, then he decides whether to adopt a curriculum embedded with propaganda. In the model, the content of education provides a signal of the political environment,[6] influencing citizens' beliefs about the importance of

[4] See Cantoni et al (2017), Alesina and Fuchs-Schundeln (2007) and Voigtlander and Voth (2015).

[5] See Nie et al (1996), Nie and Hillygus (2001), Plutzer (2002), McFarland and Thomas (2006), Campbell (2008), Quintelier (2008), Gardner et al (2008), and Kahne et al (2013).

[6] Bowles and Gintis (1976) argue that public schooling is informative of political and economic institutions not only in terms of its content but also in its organizational structures.

developing inputs useful for political participation relative to labor. By distorting this signal, propaganda can decrease a citizen's expected payoffs from political uses of human capital relative to productive ones, making joining a political club seem more costly. As a result, it can induce suboptimal levels of political participation and higher taxes in an equilibrium with public education, making its provision more appealing to a predatory ruler.

However, citizens are also not naive to a ruler's incentives to disseminate propaganda. Although propaganda increases the likelihood that a citizen is exposed to a "ruler-favorable" curriculum, it also decreases the value of public schooling for learning about the political environment. Too much propaganda can generate excessive uncertainty, such that joining a political club becomes social insurance against future expropriation. In this case, propaganda *encourages* political participation and in turn lower taxes whenever a predatory ruler invests in public education.[7] Moreover, as society becomes more productive, the potential costs of not becoming politically active increase, such that this outcome is more likely to occur.

Propaganda can therefore make investments in education more appealing to a predatory ruler when initial productivity growth from education is relatively modest, letting him promote economic development while extracting greater rents. Nevertheless, by making education provision desirable, propaganda can actually make citizens *better off*. Further development will then correspond to a shift away from propaganda to more neutral educational content, and in sufficiently high productivity settings, educational content will forgo propaganda entirely. In this latter scenario, a predatory ruler always benefits from investing in public education, inducing optimal political participation levels and low taxes. I provide evidence consistent with these predictions in Section 3.3.

As previously noted, much of the existing literature treats public education provision as either a means for a ruler to increase his rents, often at the risk of uprising or democratization, or a means of indoctrinating citizens in order to *reduce* the risk of insurrection. This paper adds to this discussion by considering an interaction of these two views. On one hand,

[7]For discussions of "backlash" and other disutilities from the character of public schooling, see Fouka (2019), Carvalho and Koyama (2016), Gradstein and Justman (2005), and Swee (2015).

it follows Bourguignon and Verdier (2000), in which education promotes economic growth but is also politically costly, as in this paper. Both models generate equilibria in which sufficient income growth from education induces investment, even if political development cannot be constrained. In their setup, however, education serves as a *de facto* franchise extension, generating intermediate equilibria in which elites retain political power by providing education only to some. This paper considers an alternative mechanism by which political externalities of public education can be reduced without varying its provision. Namely, it considers how changes within the education system, particularly with regard to its content, can relax political barriers to public education provision. Also related is Glaeser et al (2007), who discuss a model of political transition in which human capital benefits more democratic regimes, and Campante and Chor (2012), who examine empirically the conditions under which education promotes political participation.

This article is also related to Alesina et al (2019). They model public education as useful for its homogenizing effect, bringing citizens' preferences closer to elites'. In contrast, this paper considers the incentives to invest in public education when the homogenizing nature of education may actually be costly to a ruler, by facilitating political participation. This is consistent with empirical evidence that education favors political development. As in Alesina et al, however, public education may also serve as a political instrument. By using propaganda, a ruler can provide schooling while simultaneously minimizing the risk of uprising. A connected literature explores how the ideological content of education influences educational outcomes (Clots-Figueras and Masella, 2013; Meyersson, 2014; Carvalho et al, 2017). Also related is Fuchs-Schundeln and Masella (2016), who show that propaganda-based education in East Germany negatively impacted outcomes after reunification.

This paper relates in a more general way to the literature on nondemocratic institutions. Like many in that tradition, this paper concerns the expansion of public goods in settings with self-interested elites (Myerson, 2008; Gehlbach et al, 2016). Finally, the model itself is inspired by other models of strategic information transmission (Crawford and Sobel, 1982;

Kamenica and Gentzkow, 2011, Edmond, 2013).

The remainder of the paper is structured as follows. Section 3.2 presents the model and extensions. Section 3.3 presents the empirical evidence. Section 3.4 concludes.

3.2 The model

The model has two players: a ruler and a citizen. The ruler's type is $\theta \in \{0, 1\}$, where $\theta = 1$ denotes a "benevolent" ruler who will choose policies that maximize social welfare, while $\theta = 0$ denotes a "predatory" ruler who will maximize his own rents. A ruler's type is private knowledge. The citizen's prior belief, $Pr(\theta = 1) = \pi \in (0, 1)$, indicates the extent to which she has been socialized prior to attending school to *trust* the ruler.

After learning his type, a ruler first decides whether to invest in public education, $g_\theta \in \{0, 1\}$. If he invests, the citizen then decides whether or not to enroll in a public school, $e \in \{0, 1\}$. Let $y(e)$ be her education-contingent labor income. If she enrolls, she acquires human capital and her labor income becomes $y(1) = w + \alpha h$, where $h > 1$ is the *productivity effect* of education, and $\alpha > 0$.[8] Otherwise, she earns $y(0) = w \geq 0$. Human capital may also benefit the ruler down the line: once income is accrued, he imposes upon the citizen a lump sum tax, $\tau_\theta \geq 0$.

Public education provision also has political consequences. For instance, public schooling may expose the citizen to a larger peer network, enhancing her political efficacy via new social capital.[9] In the model, the citizen can become active in a *political club* while enrolled in school, a choice represented by $c \in \{0, 1\}$. If she does, then she acquires the capacity to challenge a predatory ruler's choice of τ_0 after entering the labor force. However, joining a political club comes at the cost of time and effort that might otherwise have gone toward developing more productive inputs: if $c = 1$, then one unit of productivity is allocated toward *political participation* once she enters the labor force, inducing a proportionate decrease of

[8]For algebraic simplicity, I treat αh as absorbing any direct private costs of attending school.

[9]See Gradstein and Justman (2000, 2002) and Helliwell and Putnam (2007). Also see Appendix C for a discussion of a two-citizen set up.

α in labor income. Thus, the citizen's payoff function is

$$U = y(e) - \alpha c - \tau_\theta.$$

Here, political participation means having the capacity to challenge τ_0. Let $r \in \{0,1\}$ denote a politically active citizen's decision to mount a *political challenge* against a predatory ruler.[10] A challenge might arise nonviolently as a campaign or violently as a revolt. If the citizen mounts a challenge, with probability $q \in (0,1)$ she succeeds in recovering τ_0, but a fraction $1-\sigma \in (0,1)$ of income is destroyed.

Although nondemocracies often display kleptocratic qualities (i.e. $\theta = 0$), they vary with respect to the means and ease with which a citizen can participate in the political process and thus constrain those tendencies. q and σ characterize the institutional environment underlying political participation. I interpret q as reflecting the citizen's *political power* relative to the ruler. For example, q may reflect the prevalence of explicit repression (e.g. banning of parties and ideologies). σ gives the extent to which challenges are resolved nonviolently (e.g. via trials or elections), independent of who is favored to win a challenge. I interpret this as a measure of *rule of law*, or the availability of formal institutions for conflict resolution. Altogether, I take higher q and σ to imply the prevalence of more *inclusive* political institutions (i.e. greater initial political development).

Given these institutions and the possible outcomes following public education provision, a predatory ruler may or may not want to restrict public education. Depending on θ, a ruler prefers to invest in public education if and only if his payoff,

$$V(\theta) = \theta U + (1-\theta) R_\theta,$$

is greater with public education provision than in its absence. Here, $R_\theta \geq 0$ denotes the

[10] Since a benevolent ruler only sets a tax to balance his budget constraint (i.e. pay off debts from education provision), I do not discuss the possibility that the citizen challenges a benevolent ruler.

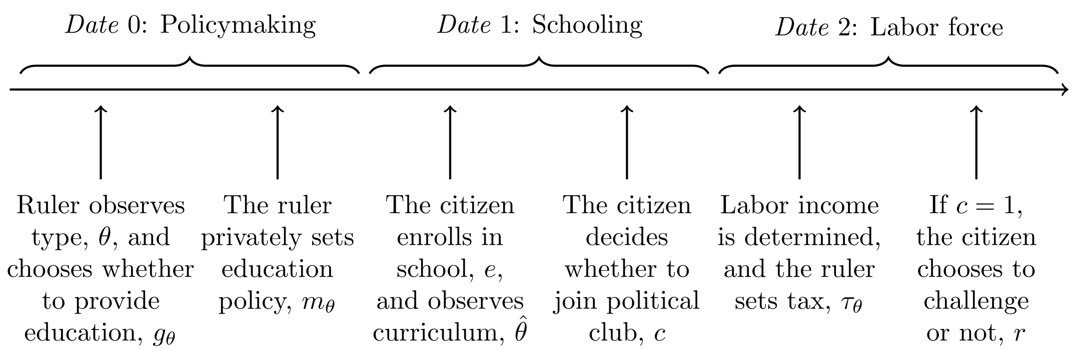


Figure 3.1: Timing of actions and events

ruler's rents at the end of the game. Investment in public education also entails a cost, normalized to 1 if the citizen enrolls in public school, and is feasible only when the ruler's expected tax revenues are sufficiently large, i.e. satisfy the budget constraint

$$R_\theta + eg_\theta \leq R^I + \mathbb{E}(\tau_\theta|\theta), \tag{3.1}$$

where R^I denotes his initial rents. For analytical simplicity, I restrict $R^I \geq \max\{1 - (1-q)\sigma(w + \alpha h - \alpha), 0\}$ so that any equilibrium in which education provision is incentive-compatible satisfies (3.1). Finally, taxes are bounded by the feasibility constraint $\tau_\theta \leq y - \alpha c$, with no other limits on fiscal capacity.

If a ruler does invest in public schooling, he can influence the content of education. In the model, the curriculum is informative of the political environment, i.e. provides a signal of the ruler's type, $\hat{\theta} \in \{0, 1\}$. If the citizen enrolls in public school, she observes $\hat{\theta}$ and updates her beliefs about the ruler's type, before deciding whether to join a political club. As such, a ruler may want to manipulate the curriculum to disseminate propaganda, a choice given by m_θ. I refer to this as his *education policy*, where $m_\theta = 1$ indicates a *biased* education policy with propaganda, while $m_\theta = 0$ indicates a *neutral* education policy without. Critically, the citizen does not observe m_θ; she can only infer it given her information about the ruler's type, which itself is endogenous to m_θ. Without loss of generality, I limit $m_1 = 0$ to focus on the importance of propaganda for public education provision by a predatory ruler.[11]

By default, education policy is neutral. I assume that it remains neutral if a ruler is indifferent between policies. Under a neutral education policy, the curriculum provides a precise signal of the ruler's type (i.e. $Pr(\hat{\theta} = 0|\theta = 0, m_0 = 0) = 1$). If a predatory ruler sets a biased education policy, with probability $\gamma \in (0, 1]$ propaganda is successfully transmitted to the citizen, who then observes a *ruler-favorable* curriculum $\hat{\theta} = 1$.[12] Hence, $\hat{\theta} = 1$ may originate with a benevolent ruler, or a predatory ruler who has successfully shaped the education system through propaganda. In this scenario, the citizen's posterior belief is

$$\hat{\pi}(\theta = 1|\hat{\theta} = 1, m_0 = 1) = \frac{\pi}{\pi + \gamma(1 - \pi)} \in [\pi, 1).$$

In other words, since education policy is privately chosen, the citizen cannot tell apart a dishonest predatory ruler from a truly benevolent one whenever both would transmit $\hat{\theta} = 1$. In turn, her process of evaluating the ruler's preferences will be distorted, skewing her expectations about what τ_θ will be once income is accrued.

Although propaganda often appeals outwardly to a ruler's benevolence, the actual channel through which it influences a citizen's behavior can vary. One potential channel is optimism, wherein propaganda convinces the citizen that a predatory ruler is truly benevolent and that she has less to lose in trusting him. Another is fear, wherein propaganda makes her believe that she has more to lose in *not* trusting him, i.e. by becoming politically active. The better interpretation will depend on the ruler. Some dictators, like Suharto or Muammar Gaddafi, have gained real popularity through development initiatives while simultaneously enriching themselves.[13] Others, like Ferdinand Marcos, have relied largely on intimidation. The interpretation adopted will not matter, however, so long as propaganda affect the citizen's belief about the opportunity cost of joining a political club.

[11] See the proof of Proposition 2 in Appendix C.

[12] I refer to all curricula $\hat{\theta} = 1$ as ruler-favorable, regardless of θ, as it induces a higher payoff.

[13] Winters (2011) notes how the perception of Suharto's Indonesia as a genuinely modernizing state, as opposed to an extractive oligarchy, was shared even by many prominent Western scholars.

Finally, there is one caveat to using propaganda. Though education policy is privately chosen, γ gives the exogenous and commonly-known probability that the citizen will nonetheless correctly infer a predatory ruler's type when he sets a biased education policy. I conceive of γ as the state of the ruler's *propaganda technology*. For example, a strong propaganda technology ($\gamma \to 1$) might reflect a highly centralized education system, while a weak technology ($\gamma \to 0$) might imply a greater prevalence of substitutes for information transmission among teachers and students.

3.2.1 Public education without propaganda

The purpose of this section is to establish the basic incentives that underlie public education provision, particularly by a predatory ruler. To begin, I define:

Definition 3.2.1. A *schooling equilibrium* [SE] consists of (i) public education provision by both types of ruler ($g_0^* = g_1^* = 1$); (ii) a predatory ruler's education policy choice m_0^*; (iii) enrollment in public schooling by the citizen ($e^* = 1$); (iv) the citizen's political club choice c^*; (v) a pair of tax choices (τ_0^*, τ_1^*); and (vi) the citizen's political challenge choice r^* whenever $c^* = 1$.

Since I am interested mainly in the strategic behavior of a predatory ruler, I will refer to all other equilibria, in which a predatory ruler refrains from providing schooling ($g_0^* = 0$), as *non-schooling equilibria* [NE]. In and of itself, a NE is rather uninteresting. In the absence of education, a predatory ruler just expropriates all income w from the citizen. If a benevolent ruler also refrains from providing schooling, he sets a social welfare maximizing tax of 0. If the citizen observes public education provision, however, there are three possible taxes τ_θ^* that she may face, conditional upon the ruler's type θ and whether she joins a political club while in school:

1. *Benevolent taxation given education provision* (BE): any surplus rents R_1 would come at the expense of the citizen, such that $\tau_1^* = \max\{1 - R^I, 0\} \equiv \tau_1^{BE}$ and $R_1^* = R^I +$

$\tau_1^{BE} - 1 \equiv R_1^{BE}$ under a benevolent ruler when $g_1 = 1$.

2. *Predatory taxation given education provision and a politically inactive citizen* (PI): a predatory ruler risks nothing by expropriating everything, such that $\tau_0^* = w + \alpha h \equiv \tau_0^{PI}$ and $R_0^* = R^I + \tau_0^{PI} - 1 \equiv R_0^{PI}$.

3. *Predatory taxation given education provision and a politically active citizen* (PA): a tax τ_0 exists that is neither high enough to induce a political challenge, nor too low that he would rather risk a challenge than just commit to the tax, such that $\tau_0^* = (1 - q\sigma)(w + \alpha h - \alpha) \equiv \tau_0^{PA}$ and $R_0^* = R^I + \tau_0^{PA} - 1 \equiv R_0^{PA}$.[14]

Altogether, these imply:

Lemma 1. $\tau_0^{PA} < \tau_0^{PI}$*, where:*

(i) For any pair of political power levels (q, q') *such that* $q' > q$, $\tau_0^{PA}(q') < \tau_0^{PA}(q)$.

(ii) For any pair of rule of law levels (σ, σ') *such that* $\sigma' > \sigma$, $\tau_0^{PA}(\sigma') < \tau_0^{PA}(\sigma)$.

Thus, insofar as it socializes the citizen to be politically active, participation in a political club while in school induces greater accountability later on under a predatory ruler in the form of strictly lower taxes. Furthermore, the magnitude of this decrease in taxes is larger when initial political institutions are more inclusive.

Given these potential taxes, I start by solving a version of the model in which education policy is fixed at its default neutral state (i.e. is always $m_\theta = 0$). As such, the political knowledge generated by the curriculum will be accurate with regard to the ruler's type (i.e. $Pr(\hat{\theta} = 0 | \theta = 0) = 1$). Consider the following equilibrium:

Definition 3.2.2. A *developmental schooling equilibrium* [DSE] is a SE wherein (i) education policy remains neutral under a predatory ruler ($m_0^* = 0$), such that ruler types separate over $\hat{\theta}$; hence (ii) the citizen joins a political club ($c^* = 1$) if and only if the ruler is predatory ($\theta = 0$), and (iii) a predatory ruler always sets $\tau_0^* = \tau_0^{PA}$.

[14]See the proof of Lemma 1. All proofs can be found in Appendix C.

It is not difficult to imagine conditions under which a DSE arises. Whereas a benevolent ruler only invests in schooling if the citizen will benefit from it, subject to the budget constraint (1), a predatory ruler has other concerns – namely, whether promoting a more educated society will yield him sufficient rents. Moreover, when education policy is neutral, the citizen can perfectly infer the threat of future expropriation and hence always joins a political club when the ruler is predatory. Thus, stronger economic incentives (i.e. higher h) are needed to induce education provision by a predatory ruler relative to a benevolent one. Formally,

Proposition 4. *There exists a developmental schooling equilibrium [DSE] when* $m_0 = 0$. *In every DSE:*

(i) $\overline{h}$ *is the minimum* h *such that a predatory ruler invests in public education, while* $\underline{h}$ *is the minimum* h *such that a benevolent ruler would invest in public education. In a DSE,* $h \geq \overline{h} > \underline{h}$.

(ii) $\overline{h}$ *is increasing in* q, σ, *and* w *and decreasing in* α, *and* $\underline{h}$ *is weakly decreasing in* α *and* R^I.

By definition, this outcome always constitutes a socially efficient arrangement *ex post*. Thus, DSE are characterized by *optimal political participation*, where the citizen becomes politically active if and only if the threat of high taxes (i.e. $\tau_0 = \tau_0^{PI}$) is credible. Absent sufficient economic incentives for a predatory ruler, however:

Definition 3.2.3. A *repressive non-schooling equilibrium* [RNE] is a NE in which (i) a predatory ruler cannot suppress political club participation ($c^* = 1$), such that (ii) he does not invest in public education ($g_0^* = 0$) and (iii) expropriates $\tau_0^* = w$, even though (iv) the citizen prefers to enroll in public school ($e^* = 1$); meanwhile (v) a benevolent ruler invests in public education ($g_1^* = 1$) and (vi) sets $\tau_1^* = \tau_1^{BE}$.

From Proposition 1 can be inferred conditions under which a predatory ruler will not invest in schooling when $m_0 = 0$, whereas a benevolent ruler will (i.e. $h \in [\underline{h}, \overline{h})$), inducing a RNE in which education provision alone accurately signals ruler type.

Hence, the interval $[\underline{h}, \overline{h})$ characterizes the space in which a citizen can always expect to be made better off by education ($e^* = 1$) but a predatory ruler does not provide it ($g_0^* = 0$).[15] In turn, when a predatory ruler *does* in invest in education, the citizen will always enroll.[16] When this happens, however, the citizen will initially face uncertainty about the ruler's true objective in providing education. It is this initial uncertainty which a predatory ruler aims to exploit with propaganda. To see how this works, I will now solve the full model, in which setting a biased education policy is a feasible option for a predatory ruler.

3.2.2 Public education, propaganda, and productivity

In the baseline model just discussed, any uncertainty that arose when conditions for education provision by both types were satisfied disappeared as the citizen became educated, inducing optimal political participation. In this section, I explore how a predatory ruler can exploit this uncertainty via education policy (i.e. when $m_0 \in \{0, 1\}$) to influence the citizen's political socialization, in turn altering his incentives to invest in public education. I begin with the case in which the citizen sometimes does not join a political club when the ruler is a predatory type.

Suboptimal political participation

In the model, public education is necessary but not sufficient for the citizen to become politically active. Rather, it is the *content* of education that drives political socialization. Lemma 1 illustrates how a predatory ruler benefits from a less politically active citizenry. I now explore the conditions under which a predatory ruler will provide public education but embed the curriculum with propaganda.

Solving the model yields conditions under which a biased education policy ($m_0^* = 1$) generates a ruler-favorable curriculum ($\hat{\theta} = 1$) under a predatory ruler ($\theta = 0$), yet where

[15] When $h < \underline{h}$, there arises a NE in which neither ruler type provides education. For the sake of space, I will henceforth discuss only equilibria in which at least one type invests.

[16] To see why this always holds, see the proofs of Propositions 1 and 2 in Appendix C.

the citizen nonetheless behaves as in a developmental schooling equilibrium [DSE] with optimal political participation (i.e. $c^* = 1$ only if $\hat{\theta} = 0$). As such, she sometimes does not join a political club when the ruler is predatory. In such cases, all human capital is allocated toward labor, rendering political challenges infeasible. In turn, political participation levels will be weakly *suboptimal* for her *ex post*, and a predatory ruler can expect his rents to be greater than in a DSE, *ceteris paribus.*

Definition 3.2.4. A *biased schooling equilibrium* [BSE] is a SE in which (i) a predatory ruler chooses a biased education policy ($m_0^* = 1$); (ii) the citizen joins a political club ($c^* = 1$) if and only if the curriculum is not ruler-favorable ($\hat{\theta} = 0$); and (iii) a predatory ruler chooses $\tau_0^* = \tau_0^{PA}$ if and only if $c^* = 1$.

Proposition 5. *There exists a biased schooling equilibrium [BSE]. In every BSE:*

(i) There exists a threshold level of propaganda $\overline{\gamma} > 0$ below which the citizen joins a political club if and only if the curriculum is not ruler-favorable ($\hat{\theta} = 0$).

(ii) There exists a threshold $\underline{\gamma}$ above which a predatory ruler invests in public education and sets a biased education policy if $\gamma < \overline{\gamma}$, where $h \geq \underline{h}$ if $\gamma \geq \underline{\gamma}$.

(iii) When $\overline{\gamma} > \underline{\gamma}$ and $\gamma \in [\underline{\gamma}, \overline{\gamma})$,[17] *with probability $(1-\pi)\gamma$ the citizen does not join a political club even though the ruler is predatory, inducing suboptimal political participation, and he sets the maximal tax $\tau_0^* = \tau_0^{PI}$.*

(iv) $\overline{\gamma}$ is decreasing in q, σ, h, and w and increasing in α and π, while $\underline{\gamma}$ is decreasing in h and α and increasing in q, σ, and w when $h > \alpha^{-1}$.

The intuition behind this equilibrium is simple. A populace of politically engaged citizens is needed to prevent expropriation, and citizens that undergo less political socialization while in school make less politically sophisticated and engaged adults. In the model, the citizen's exposure to some propaganda prompts her to forgo political club participation. Hence,

[17] Note that $h \geq \overline{h}$ is sufficient but not necessary for $\overline{\gamma} > \underline{\gamma}$.

a moderately strong propaganda technology can induce a more apolitical citizenry. Since this helps to reinforce a balance of power otherwise reduced by schooling, public education provision becomes more desirable to a predatory ruler.

This outcome is more common ($\uparrow \overline{\gamma} - \underline{\gamma}$) when the citizen is more trusting in the ruler prior to schooling ($\uparrow \pi$), when political institutions are less inclusive ($\downarrow q, \sigma$), and when human capital is relatively more valuable ($\uparrow \alpha$, $\downarrow w$). Of these comparative statics, the first suggests a complementary role for other types of propaganda outside of the education system, while the other two complement the existing literature. For example, Campante and Chor (2012) provide evidence in support of a model in which the relationship between education and political participation is weaker when relative payoffs from productive activities are greater. In the present model, less inclusive institutions and relatively more productive human capital both increase the opportunity cost of engaging in political activities. Since propaganda works by increasing that perceived opportunity cost, weaker institutions and more productive human capital increase the marginal effectiveness of propaganda, making BSE more common.

Finally, the relationship between a BSE and productivity levels h is more ambiguous, and I will revisit it shortly. First, however, consider the case in which a biased education policy induces suboptimal political participation, but insufficiently so. That is, the propaganda technology is too weak (i.e. when $\gamma \leq \underline{\gamma}$). Formally,

Definition 3.2.5. A *weak non-schooling equilibrium* [WNE] is a non-schooling equilibrium [NE] in which (i) a biased education policy ($m_0^* = 1$) induces political club participation ($c^* = 1$) if and only if the curriculum is not ruler-favorable ($\hat{\theta} = 0$), but (ii) a predatory ruler forgoes education provision ($g_0^* = 0$) and (iii) expropriates $\tau_0^* = w$; (iv) a benevolent ruler invests in schooling ($g_1^* = 1$) and (v) sets $\tau_1^* = \tau_1^{BE}$.

Remark 3. *There exists a weak non-schooling equilibrium [WNE]. In every WNE, (i) $h \geq \underline{h}$ and $\gamma < \overline{\gamma}$ but $\gamma < \underline{\gamma}$, such that the propaganda technology is too weak to sufficiently suppress political participation, where (ii) $\gamma \not< \underline{\gamma}$ for all $h \geq \overline{h}$.*

Hence, a WNE arises when the propaganda technology cannot adequately suppress political development associated with schooling and the economic gains from education are insufficiently great to induce education provision under a predatory ruler without propaganda (i.e. $h \in [\underline{h}, \overline{h})$). This mirrors the last subsection, wherein a predatory ruler's education policy choice set was limited to $m_0 = 0$. However, now:

Remark 4. *When* $m_0 \in \{0, 1\}$ *and* $h \in [\underline{h}, \overline{h})$*: if* $\gamma < \underline{\gamma} < \overline{\gamma}$*, then there exists a stronger propaganda technology* $\hat{\gamma} \in [\underline{\gamma}, \overline{\gamma})$*, which would strictly improve the citizen's expected payoffs by inducing public education provision under a predatory ruler.*

In other words, when setting a biased education policy is feasible and the propaganda technology is sufficiently strong, public education provision becomes attractive to predatory rulers in lower productivity settings (i.e. $h \in [\underline{h}, \overline{h})$). Since the citizen prefers to become educated ($e^* = 1$) when $h \geq \underline{h}$, propaganda can make her *better off* by making education provision possible. But can propaganda be too strong?

Overparticipation

The citizen knows $Pr(\hat{\theta} = 1 | \theta = 0, m_0 = 1) = \gamma$, such that her decision to join a political club after observing $\hat{\theta} = 1$ depends on the state of the propaganda technology. However, Proposition 2 alludes to an interesting caveat: the relationship between the level of propaganda and the incentive to become politically active is non-monotonic.

On one hand, if the citizen knows that a predatory ruler prefers a biased education policy ($m_0^* = 1$) but propaganda is sufficiently moderated, then apathy in the face of $\hat{\theta} = 1$ is rational. However, as the propaganda technology strengthens, the probability that she is wrongly exposed to a ruler-favorable curriculum ($\hat{\theta} = 1$) increases. By generating excessive uncertainty about the ruler's type, too much propaganda ($\gamma \geq \overline{\gamma}$) lowers the value of political knowledge generated by the curriculum. When this happens, propaganda actually *heightens* political engagement, inducing the citizen to join a political club under all curricula as social insurance against future expropriation, and producing an activistic citizenry that can always

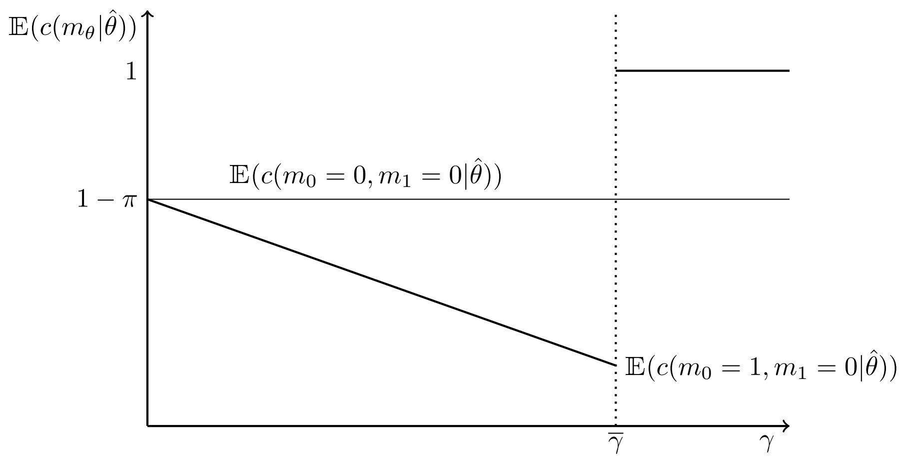


Figure 3.2: Average political participation of the educated, with and without propaganda

demand greater accountability over taxes. As before, this "overparticipation" outcome is suboptimal for the citizen, insofar as it involves her joining a political club even when it will induce an unnecessary loss (i.e. $c^* = 1$ when $\theta = 1$).

However, even more interesting are the welfare implications for a predatory ruler. In this event, he will either choose to invest in schooling but maintain a neutral education policy, inducing a developmental schooling equilibrium [DSE], or forgo education altogether, inducing a repressive schooling equilibrium [RNE]. Formally,

Proposition 6. *If setting a biased education policy is feasible for a predatory ruler,* $m_0 \in \{0,1\}$*, but the propaganda technology is sufficiently strong,* $\gamma \geq \overline{\gamma}$*:*

(i) There exists a developmental schooling equilibrium [DSE], with optimal political participation and less-than-feasible taxes, when $h \geq \overline{h}$*.*

(ii) There exists a repressive non-schooling equilibrium [RNE], with no public education and no political participation, when $h \in [\underline{h}, \overline{h})$*.*

(iii) Propaganda disappears with productivity growth: $\overline{\gamma}|_{h=\overline{h}} > 0$*, but* $\lim_{h\to\infty} \overline{\gamma} = 0$*.*

As it turns out, a predatory ruler could have done better in both cases had he been able to set a *lower* value of γ *ex ante*, inducing a biased schooling equilibrium [BSE]:

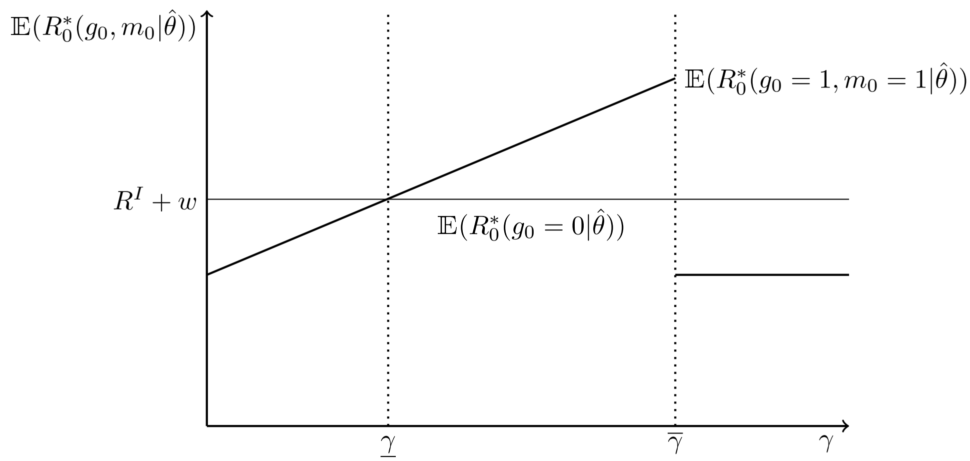


Figure 3.3: A predatory ruler's expected rents, with and without education provision, $h < \overline{h}$

Remark 5. *When* $m_0 \in \{0, 1\}$ *but* $\gamma \geq \overline{\gamma}$*: if* $\overline{\gamma} > \underline{\gamma}$*, there exists a weaker propaganda technology* $\tilde{\gamma} \in (\max\{\underline{\gamma}, 0\}, \overline{\gamma})$*, which would:*

(i) Strictly improve a predatory ruler's expected payoffs, with a payoff-maximizing $\tilde{\gamma} \equiv \tilde{\gamma}^* = \lim_{\overline{\varepsilon} \to 0} \overline{\gamma} - \overline{\varepsilon}$.

(ii) Strictly improve the citizen's expected payoffs when $h \in [\underline{h}, \overline{h})$*, by inducing public education provision by a predatory ruler.*

Hence, a predatory ruler can sometimes improve his payoffs by *weakening* propaganda. In lower productivity settings ($h \in [\underline{h}, \overline{h})$), for instance, a somewhat weak propaganda technology can help him avoid overparticipation and induce him to invest in public education.

This result also has dynamic implications. Any subsequent rise in productivity ($\uparrow h$) that might occur with development would lower the citizen's opportunity cost of becoming politically active. Insofar as this would make her more sensitive to propaganda ($\downarrow \overline{\gamma}$), rising productivity should further weaken a predatory ruler's incentives to use propaganda. Moreover, this effect would be amplified if political institutions were to also evolve to become more inclusive ($\uparrow q, \sigma$). Eventually, $\overline{\gamma} \to 0$ or $h \geq \overline{h}$, such that propaganda would cease to be useful, while productivity levels would be sufficiently high to sustain education provision without propaganda.

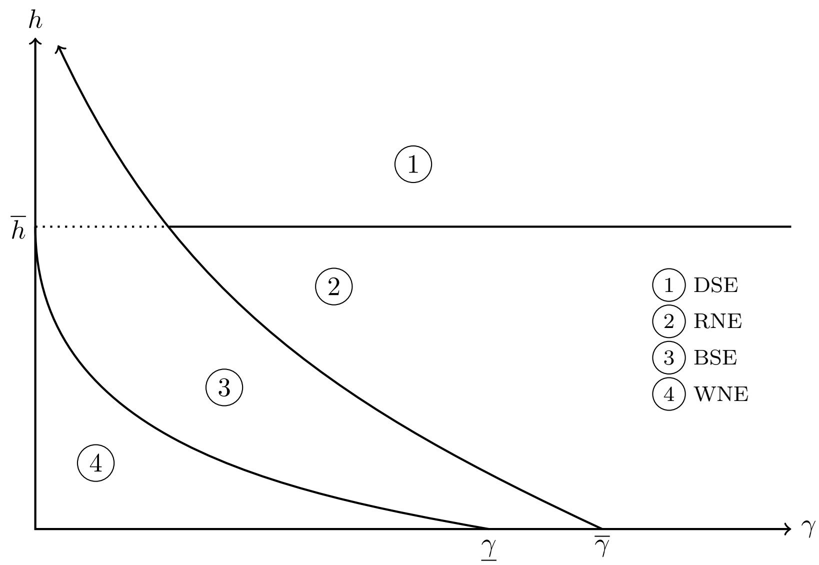


Figure 3.4: Equilibrium space over productivity (h) and propaganda technology (γ)

3.2.3 Overview of main findings

The model produces three main insights. First, an increase in the productivity effect of education from low to mid-h may be sufficient to induce investments in public education by a predatory ruler. However, this is made possible only by the feasibility of biased education policies, by which public education provision occurs as part of a biased schooling equilibrium [BSE] with propaganda. Nevertheless, by making investments in education appealing to predatory rulers, propaganda can make citizens better off. This outcome is more likely when political institutions are initially less inclusive (i.e. low q, σ), the ruler maintains high levels of trust (i.e. high π), and human capital is relatively valuable (i.e. high α, low w).

Then, as citizens become more productive ($\uparrow h$), an increasingly weaker propaganda technology is needed to sustain public education provision. In turn, one should see a shift away from biased to more neutral educational content among countries with public education. This effect is compounded as political institutions become more inclusive. However, investments in public education would nonetheless continue to improve a predatory ruler's expected payoffs, even as the frequency of political development driven by education increases.

Finally, in sufficiently high productivity settings ($h \geq \overline{h}$), public education provision is *always* incentive-compatible, even as the optimal propaganda technology *ex ante* approaches a perfectly weak technology (i.e. $\overline{\gamma} \to 0$). In this developmental schooling equilibrium [DSE], education policy is always characterized by neutral educational content, regardless of initial political institutions. Moreover, in the absence of propaganda, education always favors political development, inducing optimal political participation and in turn lower taxes.

3.2.4 Extension: A dynamic model of unproductive education

The findings of the previous section imply that a rise in the human capital from public education can decrease a predatory ruler's payoffs from providing it, by promoting transition from a biased schooling equilibrium [BSE] with propaganda to a developmental schooling equilibrium [DSE] with optimal political participation. Indeed, a predatory ruler may wish to prevent this. A temporary solution, of course, would be for him to simply use less propaganda for each subsequent cohort of citizens, as implied by Remark 3. However, a second approach would be for him to actually impede productivity growth. Take the following example:

Example 1. *Fix the ruler's propaganda technology at* $\gamma = \overline{\gamma}_{h=\overline{h}}$. *By Proposition 3, when* $h = \overline{h}$, *there is a DSE, while a lower* $h \in [\underline{h}, \overline{h})$, *conditional upon* $\gamma \geq \underline{\gamma}$, *induces a BSE. Then there exists some* $\eta = \hat{\eta} > 0$, *such that a predatory ruler's payoffs from a BSE with a productivity effect of education,* $h = \overline{h} - \eta$, *exceed his payoffs from a DSE with a greater productivity effect,* $h = \overline{h}$, *for any value of* $\eta \in (0, \hat{\eta})$.

Hence, under certain conditions, a predatory ruler will prefer propaganda over productivity. This points to an interesting dynamic problem, in which a predatory ruler has an incentive to sabotage increases in human capital. That is, he may want to choose an education policy that not only manipulates citizens' beliefs about the political environment, but also negatively impacts the productivity of education, so to sustain a BSE in the long-term. As counterintuitive as that may seem, the notion that propaganda-based education may also lower the return on education has received empirical support (Fuchs-Schundeln and Masella, 2016). In this section, I explore a dynamic extension of the model in which this arrangement is sometimes preferable from the standpoint of a predatory ruler. Furthermore, I will show that this enriches the equilibrium structure described in the proceeding sections.

There are now two time periods $t \in \{1, 2\}$, and human capital acquired in the first period increases the productivity of education in the second. A ruler with type $\theta \in \{0, 1\}$ rules over both periods. At the end of the first period, a citizen with a prior belief, $\pi_1 \in (0, 1)$, reproduces. Before dying, she transmits some information about the political environment to her offspring. To reflect this, if the ruler is a predatory type, then the second period citizen inherits a more informative prior, $\pi_2 < \pi_1$. Otherwise, $\pi_2 > \pi_1$. However, each citizen is myopic and makes decisions only with regard for her own payoffs.

If the ruler invests in education, $g_{\theta t} \in \{0, 1\}$, in any period, then an education system always exists thereafter, and he incurs a cost of 1 each remaining period. At the end of each period, he must balance the budget constraint, $R_{\theta t} + e_t g_{\theta t} \leq R^I_{\theta t} + \mathbb{E}(\tau_{\theta t}|\theta)$. He is endowed with new rents R^I at the beginning of each period, such that $R^I_{\theta 1} = R^I$ and $R^I_{\theta 2} = R_{\theta 1} + R^I$, with $R^I_{\theta t}$ restricted as before.

If he invests in education in the first period, then human capital acquisition promotes increased productivity in the second:

$$h_2 = h + (1 - \delta)h_1 > h_1,$$

where $\delta \in (0, 1)$. The ruler decides whether to invest in education in the first period given a lifetime payoff function,

$$V(\theta) = \theta U_1 + (1 - \theta)R_{\theta 1} + \beta[\theta(t - 1)U_2 + (1 - \theta)(R_{\theta 2} - R_{\theta 1})],$$

where $\beta \in (0, 1]$ is the time discount factor. This functional form assumes that a benevolent ruler internalizes citizens' myopia, while a predatory ruler does not double count the rents he acquired in the first period.

As before, if a citizen enrolls in a public school, $e_t \in \{0,1\}$, then she has the choice to join a political club, $c_t \in \{0,1\}$, and dedicate some of her time to developing politics-specific human capital. If $c_t = 1$, then she can later mount a political challenge against a predatory ruler, $r_t \in \{0,1\}$. If the citizen challenges a predatory ruler in period 1 and wins, then the ruler is replaced in period 2 by a new ruler who, with probability $1 - \pi_1$, is also predatory. However, a predatory ruler can try to influence this political club decision via education policy.

With the addition of dynamic human capital accumulation, I now allow education policy to vary over two dimensions. The curriculum continues to provide a signal of the ruler's type, $\hat{\theta}_t \in \{0,1\}$, so a predatory ruler may want to manipulate the curriculum to disseminate propaganda that targets the citizen's *beliefs* in period t, a choice given by $m^A_{\theta t} \in \{0,1\}$. However, if he opts to use propaganda, he may want to structure education such that propaganda also affects the *return* on education, suppressing human capital accumulation from period 1 to 2. I model this as a one-time choice, $m^B_\theta \in \{0,1\}$. This effect might be direct if the content of propaganda entails transmission of less productive values, or it might work through complementary teaching styles and reward structures that inhibit critical thinking or promote rent-seeking (Algan et al, 2013; Fuchs-Schundeln and Masella, 2016).

With probability $\gamma = Pr(\hat{\theta} = 1 | \theta = 0, m^A_{0t} = 1) \in (0,1]$, propaganda affects the citizen's beliefs about a predatory ruler in period t. If in addition $m^B_0 = 1$, then a portion $\rho \in (0,1)$ of education in period 1 is unproductive (i.e. $h_1 = (1-\rho)h$). Less human capital being acquired in period 1 results in a decreased productivity effect of education in period 2, relative to if $m^B_0 = 0$. I assume these two effects of propaganda to be orthogonal *ex ante*. However, they will nonetheless be highly interdependent in equilibria with propaganda, as will soon become clear.

The general structure of the game within each period remains unchanged. In turn, the results from the basic model regarding optimal taxes and rents for any schooling equilibrium [SE] will continue to hold for all t. I can now define the following:

Definition 3.2.6. An *unproductive biased schooling equilibrium* [UBSE] is a two period BSE in which (i) a predatory ruler chooses a biased education policy in both periods ($m_{0t}^{A*} = 1$ for all t); (ii) propaganda suppresses human capital accumulation from period 1 to 2, decreasing the return on education ($m_0^{B*} = 1$); (iii) a citizen joins a political club ($c_t^* = 1$) if and only if the curriculum is not ruler-favorable ($\hat{\theta}_t = 0$) for all t; and (iii) a predatory ruler chooses $\tau_{0t}^* = \tau_{0t}^{PA}$ if and only if $c_t^* = 1$ for all t.

Proposition 7. *When the propaganda technology is sufficiently strong such that $\gamma < \overline{\gamma}(h_1|m_0^B = 0)$ in period 1 but $\gamma \geq \overline{\gamma}(h_2|m_0^B = 0)$ in period 2, there exists an unproductive biased schooling equilibrium [UBSE]. In every UBSE:*

(i) There exists a positive level of unproductive education $\rho > \underline{\rho} \geq 0$ for which neither generation of citizen becomes politically active if exposed to a ruler-favorable curriculum ($\hat{\theta}_t = 0$).

(ii) If $\rho > \underline{\rho}$, then there exists an upper level of unproductive education $\rho = \overline{\rho} > 0$ below which a predatory ruler sets a biased education policy with unproductive education if he invests in education in period 1.

(iii) $\underline{\rho}$ is increasing in γ and h, while $\overline{\rho}$ is increasing in γ and decreasing in h.

The intuition behind this equilibrium is simple. Human capital accumulation (i.e. $h_2 > h_1$) decreases the effectiveness of propaganda in period 2 (i.e. $\overline{\gamma}(h_2) < \overline{\gamma}(h_1)$), making it more likely that education induces a DSE with optimal political participation and lower taxes. By letting propaganda not only impact beliefs but also lower the productivity of education in a small but non-zero way, this can be avoided. Although human capital still feeds into the productivity of education over time, it does not induce a transition from a BSE in period 1 to a DSE in period 2.

Nevertheless, the comparative statics generated by this exercise suggest that use of strategy will be short-lived. First, note that the two effects of propaganda – on beliefs and on the productivity of education – are complementary. A strong effect of propaganda on beliefs ($\uparrow \gamma$) enhances both the importance of and the incentives to use unproductive education ($\uparrow \underline{\rho}, \overline{\rho}$) for preventing transition to a DSE in period 2. Likewise, if propaganda cannot sufficiently lower the productivity of education (i.e. if $\rho \ngeq \underline{\rho}$), then a predatory ruler will prefer a weaker beliefs effect ($\downarrow \gamma$) in order to retain a BSE. Since this is more likely to occur if human capital continues to accumulate even a little (i.e. $\uparrow \underline{\rho}$ and $\downarrow \overline{\rho}$ if $\uparrow h$), a predatory ruler will eventually prefer little or no propaganda, as in the basic model.

Despite the simple environment assumed in this extension, it yields several interesting results that strengthen and expand the implications of the basic model. Moreover, it points to avenues for future theoretical work in which these assumptions are relaxed. For instance, it would be interesting to explore how propaganda's ability to impact citizens' beliefs may change over time, as well as how a predatory ruler might alter the propaganda technology to promote persistence in beliefs. I leave these as possibilities for future research.

3.3 Empirical evidence

Numerous studies have documented a robust relationship between education and democracy (Barro, 1999; Bobba and Coviello, 2007). Many others have helped illustrate, at least in part, a complex interplay between human or social capital and political development.[18] This section highlights some additional evidence relevant to this discussion, by testing two of the model's main relationships. One is a predatory ruler's incentives to use propaganda following investments in public education. Another is a citizen's tendency to be politically active after receiving an education, given her exposure to propaganda. Modeling each empirically, I establish strong and robust correlations that support the model's main predictions.

[18] See Gradstein and Justman (2000), Tabellini (2010), Algan et al (2013), Nannicini et al (2013).

3.3.1 Education provision and propaganda

To study the connection between public education and propaganda, I use data from the World Bank's (WB) World Development Indicators (WDI) index, spanning across 136 countries from 1997 to 2014. Since countries that use less propaganda exhibit political development in the model, I do not limit the sample to nondemocracies. In order to gauge a government's relative demand for public education, I use education expenditures as a percentage of total government expenditures as the explanatory variable. The response variable should thus indicate its investment in propaganda.

Unfortunately, there is no measure of propaganda within education systems across countries or over time. I therefore adopt a more general measure of propaganda – Freedom House's (FH) Freedom of the Press index, rescaled to be out of 10, with a larger value indicating more constraints on the press – as a proxy, implicitly assuming that press independence is perfectly correlated with academic independence. Nevertheless, to address concerns that this variable is too removed from the use of propaganda in schools, or similar only insofar as both are likely to be correlated with certain institutional qualities, I construct a cross-country "academic constraints" variable. This indicator takes a value of 1 only if a country appears in the Scholars at Risk (SAR) Incident Index as having punished scholars with imprisonment, prosecution, or travel restrictions.[19] In a cross section of countries, a 10% increase in a country's FH score corresponds to around a 7.5% increase in the probability that a country is assigned a 1 for academic constraints, even after controlling for general institutional qualities. Those estimates are available in Appendix C.

Additional cross-country variables are also derived from the WDI. To proxy for a country's productivity, I use GDP per capita (in USD 2010). To control for other potential sources of variation driving choices of press constraints, I include as covariates military expenditures as a percentage of GDP and government expenditures as a percentage of GDP.[20] To gauge

[19] I exclude incidents involving violence or loss of position, since the former usually follow general protests unrelated to academic content, while the latter typically reflect non-state retaliation.

[20] Cross-country observations are transformed logarithmically for greater normality. As a result, zero or

Table 3.1: Public education provision and media and press constraints

	Proxy for propaganda: press constraints and media censorship					
	(1a)	(1b)	(1c)	(1d)	(2)	(3)
lnEducation	1.850**	1.850**	1.958**	1.822**	2.011**	2.517***
	(0.731)	(0.717)	(0.762)	(0.782)	(0.817)	(0.921)
$lnEdu \times lnGDPpc$	−0.241***	−0.243***	−0.259***	−0.240**	−0.262**	−0.309***
	(0.092)	(0.091)	(0.096)	(0.098)	(0.102)	(0.116)
lnGDPpc	0.272	0.253	0.289	0.243	0.359	0.408
	(0.333)	(0.332)	(0.333)	(0.339)	(0.299)	(0.397)
Polity	−0.108***	−0.105***	−0.107***	−0.104***	−0.125***	−0.098***
	(0.020)	(0.020)	(0.022)	(0.023)	(0.024)	(0.026)
Crisis	–	0.176*	0.189*	0.194*	0.140	0.169
		(0.100)	(0.106)	(0.108)	(0.150)	(0.122)
lnMilitary	–	–	−0.025	0.008	0.029	0.040
			(0.102)	(0.118)	(0.096)	(0.134)
lnGovExpend	–	–	–	−0.045	−0.196	−0.183
				(0.215)	(0.175)	(0.237)
$lnEdu + lnEdu \times lnGDP_{min}$	0.588**	0.578**	0.605**	0.565*	0.641**	0.872***
	(0.271)	(0.264)	(0.280)	(0.292)	(0.305)	(0.330)
$lnEdu + lnEdu \times lnGDP_{10\%}$	0.363*	0.352*	0.365*	0.341	0.397*	0.586**
	(0.199)	(0.194)	(0.206)	(0.215)	(0.225)	(0.243)
$lnEdu + lnEdu \times lnGDP_{90\%}$	−0.714**	−0.734**	−0.790**	−0.732**	−0.773**	−0.790*
	(0.285)	(0.287)	(0.303)	(0.304)	(0.308)	(0.374)
Adj. R^2	0.22	0.23	0.23	0.22	0.25	0.21
Observations	1452	1452	1366	1343	2134	1092
Countries	136	136	132	131	132	85
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Balanced panel?	No	No	No	No	No	Yes
Mean imputation?	No	No	No	No	Yes	No

Standard errors are clustered by country, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. Column 2 applies a country mean imputation procedure to the education expenditure data and controls for a dummy, which takes a value of 1 when a value is missing. Rows labeled $lnEdu + lnEdu \times lnGDP$ report the combined effect (and standard errors) of *lnEducation* and $lnEdu \times lnGDPpc$ at different levels (sample minimum, 10% and 90%) within the sample distribution of *lnGDPpc*.

a country's political institutions, I include Polity IV's POLITY2 index, which ranges from −10 to 10 with 10 being most democratic. I also use POLITY2 to derive a "crisis" dummy, since propaganda may be more prevalent in more volatile times. This variable takes a value of one for any year Polity IV denotes a country as being in interregnum or transition – plus the years before and after an interregnum, transition, or foreign interruption – as well as a value of one for any year immediately preceding or following a change in POLITY2 score larger than 5.[21] Finally, I include country and year fixed effects and cluster standard errors by country.

With these data, I analyze two predictions regarding investment in propaganda: that (i) public education spending should be positively correlated with propaganda in less productive settings, and that (ii) this positive correlation should disappear with rises in productivity. Given these predictions, I include in each regression a term interacting education expenditures with real GDP per capita.

The estimates, as shown in Table 3.1, match these predictions. There is a positive and significant relationship between a government's education expenditures and its press constraints among observations in the bottom third of the distribution of real GDP per capita. This positive correlation disappears with productivity: the coefficient on the term interacting education expenditures and GDP per capita is negative and significant. This can also be seen by looking at the combined effect of education expenditures and its interaction with real GDP per capita. This is reported in Table 3.1 at different levels (sample minimum, 10%, and 90%) within the sample distribution of logged real GDP per capita. Finally, these results are also presented using standardized beta coefficients in Table C.6.

Since collecting and reporting on macroeconomic data is likely to be correlated with a country's institutions, I deal with concerns regarding missing education expenditures observations in columns (2) and (3). First, following Glaeser et al (2007), I impute missing

near-zero military expenditure values become outliers. In Table C.3, I repeat the main regressions with an alternative military expenditures variable correcting for this.

[21]Summary statistics can be found in Tables C.1 in Appendix C.

education expenditure values using country means, controlling for a dummy indicating when a value is missing. The estimates change little.[22] A second approach simply drops underreporting countries from the sample. These estimates in (3) remain substantively similar as well. Altogether, this exercise provides strong evidence in support of the model.

3.3.2 Education, propaganda, and political participation

To study how political participation levels following education vary with exposure to propaganda, I use individual-level data from the World Values Survey (WVS), spanning across 70 countries from 1994 to 2014, in addition to FH's press freedom and Polity IV's POLITY2 indexes. As before, I do not limit the sample based on polity.[23] Because the theory concerns the effects of propaganda observed while in school on political engagement thereafter, I match survey respondents with a FH score corresponding to the year of some *base age* (e.g. 18 years old), to proxy for exposure to propaganda during later adolescence. Since FH began publishing numeric scores in 1993, individuals in the sample range from 18 to 43 years of age.

To gauge adult political engagement, I use four WVS variables to create two composite response variables.[24] The first measures *passive* political participation and takes into account survey respondents' (i) interest in politics and (ii) beliefs about its importance. The most passively engaged respondents received a score of 6; the least, a score of 0. The second measures *active* participation and considers survey respondents' participation in (i) lawful demonstrations and (ii) petition signings. The most actively engaged respondents received a score of 4; the least, a score of 0.

Besides press constraints, the main explanatory variable is educational attainment, with responses ranging from 0 (incomplete primary) to 7 (university). In each specification with

[22]See Table C.4 for more on these relationships.

[23]I do however follow Glaeser et al (2007) in excluding countries with Polity IV autocracy scores that average > 4 over the period, since those may be more likely to mandate political participation.

[24]See Appendix C for explicit details about how variables were derived from WVS variables. See Table C.7 for some specifications run using each of the four original variables.

Table 3.2: Press constraints at age 18, education, and passive participation

	Measure of passive political participation					
	(1)	(2a)	(2b)	(3a)	(3b)	(3c)
EducationAttainment	0.088***	0.141***	0.097**	0.183***	0.172***	0.167***
	(0.010)	(0.029)	(0.044)	(0.045)	(0.040)	(0.040)
$EA \times PressConstraints_b$	–	−0.012**	−0.008	−0.057***	−0.053***	−0.052***
		(0.005)	(0.006)	(0.017)	(0.013)	(0.012)
$EA \times PressConstraints_b^2$	–	–	–	0.005***	0.005***	0.005***
				(0.002)	(0.001)	(0.001)
$EA \times \text{Polity}_t$	–	–	0.004	0.004	0.004**	0.005**
			(0.003)	(0.003)	(0.002)	(0.002)
$PressConstraints_b$	–	0.022	0.007	0.255***	0.251***	0.258***
		(0.025)	(0.028)	(0.097)	(0.071)	(0.064)
$PressConstraints_b^2$	–	–	–	−0.027**	−0.023***	−0.024***
				(0.011)	(0.007)	(0.006)
Polity_t	−0.003	0.011	−0.004	−0.003	0.025	0.019
	(0.018)	(0.028)	(0.033)	(0.034)	(0.024)	(0.023)
Age_t	0.054***	0.056***	0.057***	0.055***	0.053***	0.055***
	(0.018)	(0.021)	(0.021)	(0.021)	(0.018)	(0.018)
$\text{Age}_t{}^2$	−0.001**	−0.001*	−0.001**	−0.001*	−0.001*	−0.001**
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
$IncomeLevel_t$	0.033***	0.031***	0.031***	0.032***	0.025***	0.025***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.005)	(0.006)
Student_t	–	–	–	–	0.158***	0.156***
					(0.030)	(0.033)
Adj. R^2	0.13	0.13	0.13	0.13	0.11	0.11
Observations	43714	32432	32432	32432	49552	56129
Countries	65	65	65	65	70	70
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Non-students only?	Yes	Yes	Yes	Yes	No	No
Mean imputation?	No	No	No	No	No	Yes

Standard errors are clustered by country, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. Base year (b) is derived from the survey year (t), minus the respondent's age (Age_t), plus 18. All regressions control for sex, relationship status, number of children, and employment status. Columns (2c) and (2d) also apply a mean imputation procedure to educational attainment, income level, sex, relationship status, number of children, employment, and student status and control for dummies, which equal 1 when an observation is missing.

press constraints, I include a term interacting an individual's educational attainment with press constraints at her base age. Then, prompted by the model's prediction that the effects of propaganda will be non-monotonic, I allow press constraints to affect the correlation between education and political participation quadratically in columns (3a-3c). Finally, I include as controls a host of individual characteristics, as well as country and year fixed effects, and cluster standard errors by country.

Given this, I want to test three of the model's claims regarding the effects of propaganda: (i) the effect of educational attainment on political participation will be positive; (ii) exposure to more propaganda will decrease the effects of education on political participation; and (iii) the effect of propaganda on the relationship between education and political participation will decrease with the level of propaganda.

Using age 18 as the base age,[25] the estimates in specifications with passive political participation are consistent with those predictions. Although the overall effect of having press constraints on the correlation between education and political participation is at most slightly negative, quadratic specifications indicate a strong and significant effect when press constraints are relatively weak. As a level of constraints rises, this effect diminishes. If press constraints is a good proxy for propaganda, then this is reassuring evidence in favor of the model's predictions. When the response variable instead describes active participation as in Table 3.3, the results are similar, albeit at slightly lower levels of statistical significance. Results using standardized beta coefficients are presented in Table C.10.

Though the main specifications limit to sample to those who have completed schooling, the results should also hold for those still in school, for whom propaganda induces a greater allocation of time toward substitute activities. These estimates can be found in columns (3b) of Tables 3.2 and 3.3. They remain generally the same, with the correlation between being a student and being politically active being strong.

As before, many respondents correspond to missing observations. Luckily, there is more

[25] Is 18 the best base age if political socialization is still ongoing during college years? See Table C.8 for a version with 22 as the base age. The estimates are similar.

Table 3.3: Press constraints at age 18, education, and active participation

	Measure of active political participation					
	(1)	(2a)	(2b)	(3a)	(3b)	(3c)
$EducationAttainment$	0.111***	0.147***	0.115***	0.138***	0.129***	0.136***
	(0.006)	(0.016)	(0.039)	(0.043)	(0.037)	(0.033)
$EA \times PressConstraints_b$	–	−0.008***	−0.006	−0.019*	−0.019*	−0.022**
		(0.003)	(0.004)	(0.011)	(0.010)	(0.009)
$EA \times PressConstraints_b^2$	–	–	–	0.001	0.002*	0.002**
				(0.001)	(0.001)	(0.001)
$EA \times \text{Polity}_t$	–	–	0.003	0.003	0.003	0.002
			(0.003)	(0.003)	(0.002)	(0.002)
$PressConstraints_b$	–	0.033**	0.023	0.058	0.070	0.076
		(0.015)	(0.019)	(0.052)	(0.048)	(0.046)
$PressConstraints_b^2$	–	–	–	−0.004	−0.005	−0.005
				(0.005)	(0.005)	(0.004)
Polity_t	0.026	0.036	0.025	0.026	0.040**	0.043**
	(0.020)	(0.027)	(0.027)	(0.027)	(0.020)	(0.018)
Age_t	0.018	0.008	0.009	0.008	0.006	0.014
	(0.012)	(0.016)	(0.016)	(0.016)	(0.015)	(0.014)
$\text{Age}_t{}^2$	−0.0001	0.0001	0.0001	0.0001	0.0001	−0.00004
	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
$IncomeLevel_t$	0.002	0.002	0.002	0.002	0.003	0.004
	(0.006)	(0.007)	(0.007)	(0.007)	(0.005)	(0.005)
Student_t	–	–	–	–	0.193***	0.208***
					(0.034)	(0.040)
Adj. R^2	0.23	0.23	0.23	0.23	0.22	0.22
Observations	41221	30546	30546	30546	46492	52399
Countries	65	65	65	65	70	70
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Non-students only?	Yes	Yes	Yes	Yes	No	No
Mean imputation?	No	No	No	No	No	Yes

Standard errors are clustered by country, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. Base year (b) is derived from the survey year (t), minus the respondent's age (Age_t), plus 18. All regressions control for sex, relationship status, number of children, and employment status. Columns (2c) and (2d) also apply a mean imputation procedure to educational attainment, income level, sex, relationship status, number of children, employment, and student status and control for dummies, which equal 1 when an observation is missing.

flexibility in this setting, as countries have multiple observations per year. In turn, I am able to perform a country-year mean imputation procedure, generating greater variation among imputed values. There are no major differences between models with or without imputation. Columns (3c) summarize these results.

Another issue is that it is impossible to know whether a respondent was educated (and thus exposed to propaganda) in her country of current residence. Fortunately, WVS does ask some respondents about their immigration and citizenship statuses. I exclude from the analysis respondents identifying as immigrants or non-citizens, since those individuals are more likely to have characteristics that would bias the estimates, distracting from the mechanism of interest. This exclusion produces only trivial changes to the estimates. However, this is possibly due to the fact that the number of reported immigrants and non-citizens in the sample is relatively small.

Indeed, it is likely that other respondents are also immigrants. Moreover, it is also possible that some individuals might have been in the sample had they not migrated prior to the survey. Could the decision to migrate be correlated with education or press constraints? Fortunately, the existence of some respondents who were asked about immigration and citizenship allows me to make some rough inferences about the *types* of individuals who might have migrated. Being an immigrant or non-citizen, or both (i.e. "international"),[26] seems to be somewhat negatively correlated with educational attainment. However, each is similarly negatively correlated with political participation and does not appear to alter the relationship between education and political participation. As such, there is probably little reason for concern over unobservable inclusion or exclusion of migrants in the sample.

Lastly, it is worth noting that having greater press constraints at age 18 is positively correlated with passive (but not active) participation among individuals with less education (i.e. less then secondary). Whether this speaks to press censorship having a lesser effect on the uneducated or more government-favorable forms of political engagement among the

[26] See Tables C.9 in Appendix C.

uneducated deserves further study. A first step for future research into the connections between propaganda, education, and political participation would be the creation of a cross-country index of educational content. The current article cites a cross-sectional index of academic constraints as support for its use of press and media constraints to proxy for propaganda in educational settings. However, to my knowledge no systematic effort has been made to date to measure the relative prevalence of political propaganda or censorship in public schools across countries. Detailed analysis of state-sponsored curricula and teaching practices, including textual analysis of textbooks and education laws, is needed to better understand the importance of educational content for economic and political development. Finally, future research should also exploit educational reforms and regional variation in education policy in order to document causal effects of educational content on economic and political outcomes in specific settings.

3.4 Conclusion

Nondemocratic regimes face a tradeoff when investing in public education. On one hand, education promotes economic development. On the other hand, it may also facilitate political engagement and new social ties among citizens, elevating future political opposition movements. As a result, the economic incentives underlying public education provision may be insufficient to induce investment therein.

I present a model in which a ruler can relax this tradeoff by disseminating propaganda through the education system. I show that educational content will tend to be embedded with propaganda in less developed countries with public education. By making investments in education more appealing to a nondemocratic regime, however, propaganda can also make citizens better off. Further development should then coincide with decreases in propaganda. Perhaps the most important takeaway from this is that it may ultimately be a bad idea to promote "Western" (e.g. secular) education standards in illiberal countries. If successful, those efforts could drive down levels of investment in education, slowing development. Rather, if education does promote development, this transformation should eventually happen on its own.

Bibliography

[1] Abramitzky, R., L. P. Boustan, and K. Eriksson (2014). “A nation of immigrants: Assimilation and economic outcomes in the Age of Mass Migration,” *Journal of Political Economy*. 122:467-506.

[2] Acemoglu, D. (1995). “Reward structures and the allocation of talent,” *European Economic Review*. 39:17-33.

[3] Acemoglu, D. and M. Dell (2010). “Productivity differences between and within countries,” *American Economic Journal: Macroeconomics*. 2:169-88.

[4] Acemoglu, D., G. de Feo, and G. D. de Luca (2019). “Weak states: Causes and consequences of the Sicilian Mafia,” *Review of Economic Studies*. Forthcoming.

[5] Acemoglu, D., F. A. Gallego, and J. A. Robinson (2014). “Institutions, human capital, and development,” *Annual Review of Economics*. 6:875-912.

[6] Acemoglu, D., T. A. Hassan, and J. A. Robinson (2011). “Social structure and development: A legacy of the Holocaust in Russia,” *Quarterly Journal of Economics*. 126:895-946.

[7] Acemoglu, D., Johnson, S. and Robinson, J.A. (2001). “The colonial origins of comparative development: An empirical investigation,” *American Economic Review*. 91:1369-401.

[8] – (2002). “Reversal of fortune: Geography and institutions in the making of the modern world income distribution,” *Quarterly Journal of Economics*. 117:1231-94.

[9] Acemoglu, D., S. Johnson, J. A. Robinson, and P. Yared (2005). “From education to democracy?” *American Economic Review: Papers and Proceedings*. 95:44-9.

[10] Acemoglu, D., S. Naidu, P. Restrepo, and J. A. Robinson (2015). “Democracy, redistribution and inequality,” *Handbook of Income Distribution*. 2B:1885-1966.

[11] Aghion, P., X. Jaravel, T. Persson, and D. Rouzet (2019). “Education and military rivalry,” *Journal of the European Economic Association*. 17:376-412.

[12] Agnew, H. (2004). *The Czechs and the Lands of the Bohemian Crown*. Stanford: Hoover Institution Press.

[13] Akbulut-Yuksel, M. and M. Yuksel (2015). "The long-term direct and external effects of Jewish expulsion in Nazi Germany," *American Economic Journal: Economic Policy*. 7:58-85.

[14] Alesina, A., and N. Fuchs-Schundeln (2007). "Goodbye Lenin (or not?): The effect of Communism on people's preferences," *American Economic Review*. 97:1507-28.

[15] Alesina, A., P. Giuliano, and B. Reich (2019). "Nation-building and education," Working Paper.

[16] Algan, Y., P. Cahuc, and A. Shleifer (2013). "Teaching practices and social capital," *American Economic Journal: Applied Economics*. 5:189-210.

[17] Allen, T. and D. Donaldson (2018). "The geography of path dependence." Working Paper.

[18] Anbarci, N., M. Escaleras, and C. A. Register (2005). "Earthquake fatalities: The interaction of nature and political economy," *Journal of Public Economics*. 89:1907-33.

[19] von Arburg, A. (2001). *Die Besiedlung der Grenzgebiete der böhmischen Länder 1945-1950*. Vienna: University of Vienna.

[20] Bandiera, O., M. Mohnen, I. Rasul, and M. Viarengo (2017). "Nation-building through compulsory schooling during the age of mass migration." Working Paper.

[21] Barone, G., and S. Mocetti (2014). "Natural disasters, growth and institutions: A tale of two earthquakes," *Journal of Urban Economics*. 84:52-66.

[22] Barro, R. J. (1999). "Determinants of democracy," *Journal of Political Economy*. 107(S6):158-83.

[23] Bauer, T. K., S. T. Braun, and M. Kvasnicka (2013). "The economic integration of forced migrants: Evidence for postwar Germany," *Economic Journal*. 123:998-1024.

[24] Becker, S. O., K. Boeckh, C. Hainz, and L. Woessmann (2014). "The empire is dead, long live the empire! Long-run persistence of trust and corruption in the bureaucracy," *Economic Journal*. 126:40-74.

[25] Becker, S. O. and A. Ferrara (2019). "Consequences of forced migration: Recent findings," *Labour Economics*. Forthcoming.

[26] Becker, S. O., I. Grosfeld, P. Grosjean, N. Voigtländer, and E. Zhuravskaya (2019). "Forced migration and human capital: Evidence from post-WWII population transfers." Working Paper.

[27] Bell, B., F. Fasani, and S. Machin (2013). "Crime and immigration: Evidence from large immigrant waves." *Review of Economics and Statistics*. 95:1278-90.

[28] Belloc, M., F. Drago, and R. Galbiati (2016). "Earthquakes, religion, and transition to self-government in Italian cities," *Quarterly Journal of Economics*. 131:1875-926.

[29] Bernàšek, M. (1970). "Czechoslovak planning, 1945-48," *Soviet Studies*. 22:94-109.

[30] Bjork, C. (2005). *Indonesian Education: Teachers, Schools, and Central Bureaucracy*. New York: Routledge.

[31] Bleakley, H. and J. Lin (2012). "Portage and path dependence," *Quarterly Journal of Economics*. 127:587-644.

[32] Bobba, M., and D. Coviello (2007). "Weak instruments and weak identification in estimating the effects of education on democracy," *Economic Letters*. 96:301-7.

[33] Bosker, M., S. Brakman, H. Garretsen, and M. Schramm (2007). "Looking for multiple equilibria when geography matters: German city growth and the WWII shock," *Journal of Urban Economics*. 61:152-67.

[34] Bourguignon, F. and T. Verdier (2000). "Oligarchy, democracy, inequality, and growth," *Journal of Development Economics*. 62:285-313.

[35] Brakman, S., H. Garretsen, and M. Schramm (2004). "The strategic bombing of German cities during WWII and its impact on city growth," *Journal of Economic Geography*. 4:201-18.

[36] Braun, S. T., A. Kramer, and M. Kvasnicka (2017). "Local labor markets and the persistence of population shocks." Working Paper.

[37] Bryant, C. (2002). "Either German or Czech: Fixing nationality in Bohemia and Moravia, 1939-1946," *Slavic Review*. 61:683-706.

[38] – (2007). *Prague in Black: Nazi Rule and Czech Nationalism*. London: Harvard University Press.

[39] Bursztyn, L. (2016). "Poverty and the political economy of public education spending: evidence from Brazil," *Journal of the European Economic Association*. 14:1101-28.

[40] Campante, F. R. and D. Chor (2012). "Schooling, political participation, and the economy," *Review of Economic and Statistics*. 94:841-859.

[41] Campbell, D. E. (2008). "Voice in the classroom: How an open classroom climate fosters political engagement among adolescents," *Political Behavior*. 30:437-54.

[42] Cantoni, D., Y. Chen, D. Y. Yang, N. Yuchtman, and Y. J. Zhang (2017). "Curriculum and ideology," *Journal of Political Economy*. 125:338-92.

[43] Carvalho, J., and M. Koyama (2016). "Resisting education." Working paper.

[44] Carvalho, J., M. Koyama, and M. Sacks (2017). "Education, identity and community: Lessons from Jewish emancipation," *Public Choice*. 171:119-43.

[45] Castelló-Climent, A. (2008). "On the distribution of education and democracy," *Journal of Development Economics*. 87:179-90.

[46] Chaney, E. and R. Hornbeck (2016). "Economic dynamics in the Malthusian era: Evidence from the 1609 Spanish expulsion of the Moriscos," *Economic Journal*. 126:1404-40.

[47] Chong, A. and M. Gradstein (2015). "On education and democratic preferences," *Economics and Politics*. 27:362-88.

[48] Clots-Figueras, I. and P. Masella (2013). "Education, language, and identity," *Economic Journal*. 123:332-57.

[49] Cornwall, M. (1994). "The struggle on the Czech-German language border, 1880-1940," *English Historical Review*. 433:914-51.

[50] Crawford, V. P., and J. Sobel (1982). "Strategic information transmission," *Econometrica*. 50:1431-51.

[51] Daněk, P. (1995). "Towards cultural regionalization of the Czech lands: Sudeten half a century after the transfer," *Geography*. 25:41-60.

[52] Davis, D. R. and D. E. Weinstein (2002). "Bones, bombs, and break points: The geography of economic activity," *American Economic Review*. 92:1269-89.

[53] – (2008). "A search for multiple equilibria in urban industrial structure," *Journal of Regional Science*. 48:29-65.

[54] De la Croix, D., and M. Doepke (2009). "To segregate or to integrate: Education politics and democracy," *Review of Economic Studies*. 76:597-628.

[55] Dee, T. S. (2004). "Are there civic returns to education?" *Journal of Public Economics*. 88:1697-720.

[56] Dell, M. (2010). "The persistent effects of Peru's mining *mita*," *Econometrica*. 78:1863-1903.

[57] Dell, M. and B. A. Olken (2019). "The development effects of the extractive colonial economy: The Dutch Cultivation System in Java," *Review of Economic Studies*. Forthcoming.

[58] Droller, F. (2017). "Migration, population composition, and long-run economic development: Evidence from settlements in the Pampas," *Economic Journal*. 127:1-32.

[59] Duranton, G. and D. Puga (2004). "Micro-foundations of urban agglomeration economies," *Handbook of Regional and Urban Economics*. 4:2063-117.

[60] Edmond, C. (2013). "Information manipulation, coordination, and regime change," *Review of Economic Studies*. 80:1422-58.

[61] Ellison, G., E. L. Glaeser, and W. R. Kerr (2010). "What causes industry agglomeration? Evidence from coagglomeration patterns," *American Economic Review*. 100:1195-213.

[62] Fidrmuc, J. (2004). "Labor mobility during transition: Evidence from the Czech Republic," *Journal of Labor Market Research*. 37:437-50.

[63] Fidrmuc, J. and P. Huber (2007). "The willingness to migrate in the CEECs: Evidence from the Czech Republic," *Empirica*. 34:351-69.

[64] Foged, M. and G. Peri (2016). "Immigrants' effect on native workers: New analysis on longitudinal data," *American Economic Journal: Applied Economics*. 8:1-34.

[65] Fouka, V. (2019). "Backlash: The unintended effects of language prohibition in US schools after World War I," *Review of Economic Studies*. Forthcoming.

[66] Friedman, W., M. Kremer, E. Miguel, and R. Thornton (2016). "Education as liberation?" *Economica*. 83:1-30.

[67] Fuchs-Schundeln, N. and P. Masella (2016). "Long-lasting effects of socialist education," *Review of Economics and Statistics*. 98:428-41.

[68] Galor, O. and O. Moav (2006). "Das human-kapital: A theory of the demise of the class structure," *Review of Economic Studies*. 73:85-117.

[69] Gardner, M., J. Roth, and J. Brooks-Gunn (2008). "Adolescents' participation in organized activities and developmental success 2 and 8 years after high school: Do sponsorship, duration, and intensity matter?" *Developmental Psychology*. 44:814-30.

[70] Gehlbach, S., K. Sonin, and M. Svolik (2016). "Formal models of nondemocratic politics," *Annual Review of Political Science*. 19:565-84.

[71] Gelman, A. and G. Imbeds (2018). "Why high-order polynomials should not be used in regression discontinuity designs," *Journal of Business and Economic Statistics*. Forthcoming.

[72] Gerlach, D. (2017). *The Economy of Ethnic Cleansing: The Transformation of the German-Czech Borderlands after WWII*. New York: Cambridge University Press.

[73] Glaeser, E. L., G. Ponzetto, and A. Shleifer (2007). "Why does democracy need education?" *Journal of Economic Growth*. 12:77-99.

[74] Glaeser, E., R. La Porta, F. Lopez-de-Silanes, and A. Shleifer (2004). "Do institutions cause growth?" *Journal of Economic Growth*. 9:271-303.

[75] Glassheim, E. (2001). "The mechanics of ethnic cleansing: The expulsion of Germans from Czechoslovakia, 1945-47," in P. Ther and A. Siljak, eds., *Redrawing Nations: Ethnic Cleansing in East-Central Europe*. Lanham, MD: Rowman & Littlefield.

[76] – (2016). *Cleansing the Czechoslovak Borderlands: Migration, Environment, and Health in the Former Sudetenland*. Pittsburg: University of Pittsburg Press.

[77] Goldstein, E. (1999). "Neville Chamberlain, the British official mind and the Munich crisis," *Diplomacy & Statecraft*. 10:276-92.

[78] Gradstein, M. and M. Justman (2002). "Education, social cohesion, and economic growth," *American Economic Review*. 92:1192-204.

[79] – (2000). "Human capital, social capital, and public schooling," *European Economic Review*. 44:879-90.

[80] – (2005). "The melting pot and school choice," *Journal of Public Economics*. 89:871-96.

[81] Greenstone, M., R. Hornbeck, and E. Moretti (2010). "Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings," *Journal of Political Economy*. 118:536-98.

[82] Greger D., M. Levínská, and I. Smetáčková (2012). "Priority education policies in the Czech Republic: Redesigning equity policies in the post-communist transformation," in M. Demeuse, D. Frandji, D. Greger, and J.Y. Rochex, eds., *Educational Policies and Inequalities in Europe*. London: Palgrave Macmillan.

[83] Grosfeld, I., A. Rodnyansky, and E. Zhuravskaya (2013). "Persistent antimarket culture: A legacy of the Pale of Settlement after the Holocaust," *American Economic Journal: Economic Policy*. 5:189-226.

[84] Helliwell, J. F. and R. D. Putnam (2007). "Education and social capital," *Eastern Economic Journal*. 33:1-19.

[85] Hornung, E. (2014). "Immigration and the diffusion of technology: The Huguenot diaspora in Prussia," *American Economic Review*. 104:84-122.

[86] Illner, M. and A. Andrle (1994). "The regional aspect of post-communist transformation in the Czech Republic," *Czech Sociological Review*. 2:107-27.

[87] Jedwab, R. and A. Moradi (2016). "The permanent effects of transportation revolutions in poor countries: Evidence from Africa," *Review of Economics and Statistics*. 98:268-84.

[88] Jedwab, R., N. Johnson, and M. Koyama (2016). "Bones, bacteria and break points: The heterogeneous spatial effects of the Black Death and long-run growth." Working Paper.

[89] – (2019). "Pandemics, places, and populations: The effects of the Black Death on urban development." Working Paper.

[90] Johnson, N. D. and M. Koyama (2017). "Jewish communities and city growth in preindustrial Europe," *Journal of Development Economics*. 127:339-54.

[91] Jelínek, J. (2011). *Analýza Topografických Map 1:50000 z Období Protektorátu Čechy a Morava*. Brno: Masaryk University.

[92] Judson, P. M. (2006). *Guardians of the Nation: Activists on the Language Frontiers of Imperial Austria*. Cambridge: Harvard University Press.

[93] Kahne, J., D. Crow, and N. Lee (2013). "Different pedagogy, different politics: High school learning opportunities and youth political engagement," *Political Psychology*. 34:419-41.

[94] Kamenica, E. and M. Gentzkow (2011). "Bayesian persuasion," *American Economic Review*. 101:2590-615.

[95] Karaja, E. and J. Rubin (2017). "The cultural transmission of trust norms: Evidence from a lab in the field on a natural experiment." Working Paper.

[96] King, J. (2002). *Budweisers into Czechs and Germans: A Local History of Bohemian Politics, 1848-1948*. Princeton: Princeton University Press.

[97] Kline, P. and E. Moretti (2014). "Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority," *Quarterly Journal of Economics*. 129:275-331.

[98] Korbel, J. (1959). *The Communist Subversion of Czechoslovakia, 1938-1948*. Princeton: New Jersey.

[99] Krugman, P. (1991). "Increasing returns and economic geography," *Journal of Political Economy*. 99:483-99.

[100] Lipset, S. (1959). "Some social requisites of democracy: Economic development and political legitimacy." *American Political Science Review*. 53:69-105.

[101] Lott, Jr., J. R. (1999). "Public schooling, indoctrination, and totalitarianism," *Journal of Political Economy*. 107:127-57.

[102] Maloney, W. F. and F. V. Caicedo (2015). "The persistence of (subnational) fortune," *Economic Journal*. 126:2363-401.

[103] Marshall, A. (1920). *Principles of Economics*. London: MacMillan.

[104] McFarland, D. A. and R. J. Thomas (2006). "Bowling young: How youth voluntary associations influence adult political participation," *American Sociological Review*. 71:401-25.

[105] Meyersson, E. (2014). "Islamic rule and the empowerment of the poor and pious," *Econometrica*. 82:229-69.

[106] Michalopoulos, S. and E. Papaioannou (2013). "Precolonial ethnic institutions and contemporary African development," *Econometrica*. 81:113-52.

[107] Michalopoulos, S., N. Benos, S. Karagiannis and E. Papaioannou (2019). "Refugees in the Mediterranean: The impact of 1920's forced population exchange between Greece and Turkey on the Greek economy." Working Paper.

[108] Miguel, E. and G. Roland (2011). "The long-run impact of bombing Vietnam," *Journal of Development Economics*. 96: 1-15.

[109] Milligan, K., E. Moretti, and P. Oreopoulos (2004). "Does education improve citizenship? Evidence from the US and the UK," *Journal of Public Economics*. 88:1667-95.

[110] Moretti, E. (2004). "Human capital externalities in cities," in J. V. Henderson and J. Thisse, eds., *Handbook of Regional and Urban Economics*, Volume 4. New York: North Holland.

[111] Moscona, J., N. Nunn, and J. A. Robinson (2018). "Social structure and conflict: Evidence form sub-Saharan Africa." Working Paper.

[112] Murard, E. and S. O. Sakalli (2018). "Mass refugee inflow and long-run prosperity: Lessons from the Greek population resettlement." Working Paper.

[113] Myerson, R. B. (2008). "The autocrat's credibility problem and foundations of the constitutional state," *American Political Science Review*. 102:125-39.

[114] Nannicini, T., A. Stella, G. Tabellini, and U. Troiano (2013). "Social capital and political accountability," *American Economic Journal: Economic Policy*. 5:222-50.

[115] Nie, N. and D. Hillygus (2001). "Education and democratic citizenship: Explorations into the effects of what happens in pursuit of the baccalaureate," in D. Ravitch and J. Viteritti, (eds.), *Education and Civil Society*. New Haven: Yale University Press.

[116] Nie, N., J. Junn, and K. Stehlik-Barry (1996). *Education and Democratic Citizenship in America*. Chicago: University of Chicago Press.

[117] Nunn, N. (2007). "Historical legacies: A model linking Africa's past to its current underdevelopment," *Journal of Development Economics*. 83:157-75.

[118] – (2008). "The long term effects of Africa's slave trades," *Quarterly Journal of Economics*. 123:139-76.

[119] – (2009). "The importance of history for economic development," *Annual Review of Economics*. 1:65-92.

[120] – (2014). "Historical development," *Handbook of Economic Growth*. 2A:347-402.

[121] *Odsun: Die Vertreibung der Sudetendeutschen* (1995). Munich: Sudetendeutsches Archiv.

[122] Pascali, L. (2016). "Banks and development: Jewish communities in the Italian Renaissance and current economic performance," *Review of Economics and Statistics*. 98(1):140-58.

[123] Plutzer, E. (2002). "Becoming a habitual voter: Inertia, resources, and growth in young adulthood," *American Political Science Review*. 96:41-56.

[124] Quintelier, E. (2008). "Who is politically active: The athlete, the scout member, or the environmental activist?" *Acta Sociologica*. 51:355-70.

[125] Radvanovský, Z. (2001). "The social and economic consequences of resettling Czechs into Northwestern Bohemia, 1945-47," in P. Ther and A. Siljak, eds., *Redrawing Nations: Ethnic Cleansing in East-Central Europe*. Lanham, MD: Rowman & Littlefield.

[126] Rappaport, J. and J. D. Sachs (2003). "The United States as a coastal nation," *Journal of Economic Growth*. 8:5-46.

[127] Redding, S. J. (2010). "The empirics of new economic geography," *Journal of Regional Science*. 50:297-311.

[128] Rocha, R., C. Ferraz, and R. R. Soares (2017). "Human capital persistence and development," *American Economic Journal: Applied Economics*. 9:105-36.

[129] Rosenthal, S. S. and W. C. Strange (2004). "Evidence on the nature and sources of agglomeration economies," *Handbook of Regional and Urban Economics*. 4:2119-171.

[130] Ruiz, I. and C. Vargas-Silva (2013). "The economics of forced migration," *Journal of Development Studies*. 6:772-84.

[131] Schumann, A. (2014). "Persistence of population shocks: Evidence from the occupation of West Germany after WWII," *American Economic Journal: Applied Economics*. 6:189-205.

[132] Semrad, A. (2015). "Immigration and educational spillovers: Evidence from Sudeten German expellees in post-war Bavaria." Working Paper.

[133] Sequeira, S., N. Nunn, and N. Qian (2019). "Immigrants and the making of America," *Review of Economic Studies*. Forthcoming.

[134] Skaperdas, S. (1992). "Cooperation, conflict, and power in the absence of property rights," *American Economic Review*. 82:720-39.

[135] – (1996). "Contest success functions," *Economic Theory*. 7:283-290.

[136] Spurný, M. (2013). "Reliability and the border: The discourse of the Czech borderlands, 1945-49," *Austrian Journal of Political Science*. 42:83-94.

[137] Swee, E. L. (2015). "Together or separate? Post-conflict partition, ethnic homogenization, and the provision of public schooling," *Journal of Public Economics*. 128:1-15.

[138] Tabellini, G. (2010). "Culture and institutions: Economic development in the regions of Europe," *Journal of the European Economic Association*. 8:677-716.

[139] Tampke, J. (2003). *Czech-German Relations and the Politics of Central Europe*. New York: Palgrave Macmillan.

[140] Taylor, T. (1980). *Munich: The Price of Peace*. New York: Random House.

[141] Testa, P. A. (2019). "The economic legacy of expulsion: Lessons from postwar Czechoslovakia." Working Paper.

[142] – (2019). "Shocks and the spatial distribution of economic activity: The role of institutions." Working Paper.

[143] Voigtlander, N. and H. Voth (2015). "Nazi indoctrination and anti-Semitic beliefs in Germany," *Proceedings of the National Academy of Sciences*. 112:7931-6.

[144] Waldinger, F. (2010). "Quality matters: The expulsion of professors and the consequences for PhD student outcomes in Nazi Germany," *Journal of Political Economy*. 118:787-831.

[145] Winters, J. A. (2011). *Oligarchy*. New York: Cambridge University Press.

[146] Wiskemann, E. (1938). *Czechs and Germans: A Study of the Struggle in the Historic Provinces of Bohemia and Moravia*. New York: Oxford University Press.

[147] Woodson, C. G. (1915). *The Education of the Negro Prior to 1861*. New York: G.P. Putnam's Sons.

[148] Yanagizawa-Drott, D. (2014). "Propaganda and conflict: Evidence from the Rwandan Genocide," *Quarterly Journal of Economics*. 129:1947-94.

[149] Zahra, T. (2008). *Kidnapped Souls: National Indifference and the Battle for Children in the Bohemian Lands, 1900-48*. Ithaca: Cornell University Press.

[150] de Zayas, A. (1989). *A Terrible Revenge: The Ethnic Cleansing of the East European Germans, 1944-50*. New York: St. Martin's Press.

Appendix A

Appendix for Chapter 1

Table A.1: Summary statistics

Variable	Observations	Mean	Std. Dev.	Min.	Max.
		Pre-war data (1923-33)			
% German	325	35.666	41.132	.024	98.845
% Roma	325	.002	.012	0	.129
% Jewish (ethnic group)	325	.145	.316	0	2.825
% Jewish (religion)	325	.522	.645	0	6.981
% Literate	325	98.516	.709	94.94	99.65
Convictions per capita	318	7.274	2.057	2.484	16.318
% Taxpayers	147	5.698	1.786	2.33	11.65
Income per capita (100 Kčs)	145	9.469	4.105	3.669	30.775
ln Population density	325	4.724	.623	3.336	9.002
ln Labor force density	325	3.951	.698	2.413	8.476
Labor force participation	325	46.468	5.318	33.339	61.908
Unemployment	146	13.125	10.186	1.417	58.796
% Agricultural sector	325	28.077	13.991	.769	60.529
% Secondary sector	325	40.602	13.672	16.967	76.309
% Industry	325	33.557	14.183	10.753	72.145
% Mining and other extraction	325	3.59	5.199	.26	36.218
% Metallurgy and metalwork	325	4.411	3.499	1.429	24.611
% Manufacturing	325	2.283	2.21	.311	16.322
% Glasswork	325	1.133	3.687	0	33.643
% Textiles	325	7.13	10.822	.029	54.741
% Other industry	325	15.01	6.149	6.809	62.857
% Construction	325	7.045	2.364	2.806	17.536
% Transport sector	325	3.473	2.012	1.131	13.615
% Business sector	325	5.999	2.497	2.592	20.841
% Finance and insurance	325	.401	.279	0	3.084
% Trade	325	5.597	2.301	2.512	19.469
% Other service	325	6.694	3.779	3.071	29.368
		Geographic data			
Elevation (m)	6206	410.505	144.345	121.833	1144.601
Ruggedness (°)	6206	6.422	3.001	1.053	20.725
Precipitation (mm)	6206	53.047	6.98	40.494	100.068
Temperature (°C)	6206	7.581	.82	3.262	9.534
Rivers/km^2 (km)	6206	1.183	.52	0	5.1
% Arable land, 1945	159	45.39	14.569	7.938	77.664

Table A.1: Summary statistics (II)

Variable	Observations	Mean	Std. Dev.	Min.	Max.
		Post-expulsion/resettlement data (mid-1947)			
ln Population density	325	4.396	.742	2.1	8.913
% Agricultural sector	325	29.417	15.433	1.096	68.175
% Secondary sector	325	47.046	14.501	14.138	81.535
% Transport sector	325	5.136	2.776	1.373	17.252
% Business sector	325	6.372	1.877	2.262	15.599
General enrollment per 100, 5-14	160	56.201	5.287	47.15	73.243
Civic enrollment per 100, 10-14	160	52.363	8.596	27.805	75.51
Agricultural enroll. per 100, 15-19	160	6.528	4.976	0	22.097
Vocational enroll. per 100, 15-19	160	15.212	10.74	0	57.857
College enrollment per 100, 15-24	160	2.14	1.506	.093	10.552
		Contemporary data (2011)			
Unemployment	6206	11.511	5.686	0	52.222
ln Population density	6206	3.987	.946	-3.965	7.894
ln Labor force density	6206	3.251	.967	-5.064	7.214
% Agricultural sector	6206	7.111	6.195	0	100
% Industry	6206	25.769	8.119	0	58.974
% Construction	6206	7.135	3.148	0	30.769
% Transport sector	6206	5.301	2.75	0	27.273
% Finance and insurance	6206	1.381	1.262	0	10
% Hospitality	6206	2.435	2.341	0	41.213
% Auto trade and repair	6206	7.861	3.322	0	29.167
% Public	6206	4.54	2.664	0	64.029
% Communications	6206	1.33	1.399	0	14.085
% Education	6206	4.012	2.199	0	22.222
% Healthcare	6206	4.525	2.864	0	54.412
% Other service	6206	4.599	2.806	0	33.333
% Primary education or less	6206	21.797	5.698	0	68.908
% Secondary education	6206	65.919	5.317	25.21	86.111
% Tertiary education	6206	8.364	4.214	0	33.741
		Panel data (1921-2011)			
% Agricultural sector	657	18.399	13.032	.426	56.506
% Industry	657	36.892	10.585	11.95	70.679
% Service sector	657	28.909	11.846	7.479	63.622
ln Population density	730	4.727	.647	3.4	7.431
ln Labor force density	657	3.979	.684	2.347	6.786
Education index	584	-.024	.929	-4.183	3.882
% Secondary education	438	46.962	18.93	6.498	66.538
% Tertiary education	438	7.079	4.441	1.407	28.584
Net migrants per capita	511	-.032	.537	-2.383	2.634
In migrants per capita	511	1.945	1.326	.568	9.378
Out migrants per capita	511	1.977	1.301	.682	10.16

For variable descriptions, see the end of these supplementary materials. This table omits Prague and Polish Zaolzie since they are excluded from all analyses. Sample is otherwise not limited, including by bandwidth or by the extent of overlap with the Munich Agreement line, except: 1933 income per capita data are missing for a handful of political districts in the Prague area (Praha-venkov, Ricany, and Jilove). Unemployment data for 1933 political districts are missing for Praha-venkov, while labor force data are combined for Olomouc and Olomouc-venkov. 1923-7 convictions data merge several districts into larger criminal jurisdictions in the Brno, Zlin, and Prague urban areas. In 1945, political districts Lanskroun and Usti nad Orlici were merged, so I manually merge them for the 1945 arable land variable.

Table A.2: Geographic balance tests

	Borderlands	Interior	Difference S.E.	Borderlands	Interior	Difference S.E.
Elevation	407.243	401.409	(12.065)	434.071	398.881	(5.147)***
Ruggedness	6.554	6.455	(.253)	7.373	6.093	(.106)***
Precipitation	53.471	53.920	(.612)	54.610	53.104	(.259)***
Temperature	7.517	7.590	(.076)	7.244	7.650	(.028)***
Rivers/km^2	1.045	1.115	(.0 46)	1.163	1.141	(.019)
Observations	224	322	546	1102	2947	4049
Bandwidth	2 km	2 km	2 km	25 km	25 km	25 km
Arable land	46.974	50.279	(6.374)	39.600	50.547	(3.462)***
Observations	11	14	25	30	38	68
Bandwidth	10 km	10 km	10 km	25 km	25 km	25 km
Year	1945	1945	1945	1945	1945	1945

Mean difference standard errors reported in parentheses, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All t-tests exclude Prague and Polish Zaolzie.

Table A.3: WWII deaths by group

Group	Cause of wartime death	Casualties	Source	Notes
Sudeten Germans	Military deaths	∼180,000; <206,000	Die Deutschen Vertreibungsverluste: Bevölkerungsbilanzen für die Deutschen Vertreibungsgebiete, 1939-50 (1958); Overmans (2004)	Includes Sudeten German servicemen who died during the liberation of Czechoslovakia in May of 1945, which also marked the start of the expulsions. Though impossible to know theexact number killed during the liberation, it was a violent event that left hundreds of thousands dead. Overmans estimates 206,000 Germans dead from all territories annexed by Germany in WWII.
	Civilian casualties	?		Uncertain how many Sudeten Germans died in the bombings that hit Czechoslovakia during the war. However, few bombs struck the country, and most were in the interior (see Figure A.2). About 30,000 Sudeten German civilian deaths, including about 7000 murders at Czech hands, occurred during the expulsion itself (Gerlach, 2017).
Jews	Executed by Nazis or died from forced labor	270,000	Erlikhman (2004)	Previously smooth through MAL.
Roma	Executed by Nazis or died from forced labor	8000	Erlikhman (2004)	Previously smooth through MAL.
Other Czechs	Military deaths	35,000	Erlikhman (2004)	
	Civilian casualities	10,000	Erlikhman (2004)	
	Executed by Nazis or died from forced labor	32,000	Erlikhman (2004)	

Overmans (2004) refers to *Deutsche Militärische Verluste im Zweiten Weltkrieg*, Munich: Oldenbourg. Erlikhman (2004) refers to *Poteri Narodonaseleniia v XX Veke: Spravochnik*, Moscow: Russkaia Panorama.

Table A.4: RD robustness, balance tests

	% German	Literacy	ln Pop. density	Unemploy.	$Income_{pc}$	% Taxpayer	Agri. sector
	(1a)	(1b)	(1c)	(1d)	(1e)	(1f)	(1g)
	Local conditional mean comparison, 10km bandwidth						
In borderlands	74.186	.084	-.162	.676	-1.249	.422	-1.149
	(5.605)***	(.173)	(.153)	(2.148)	(2.630)	(.832)	(2.731)
R^2	.924	.288	.355	.769	.361	.513	.404
Observations	70	70	70	20	21	21	70
Clusters	53	53	53	–	–	–	53
Border segments	4	4	4	4	4	4	4
	Local linear in longitude and latitude						
In borderlands	73.575	.264	-.108	1.781	-.392	.391	-1.264
	(3.853)***	(.152)*	(.116)	(2.449)	(1.490)	(.484)	(2.577)
R^2	.933	.515	.482	.645	.387	.541	.535
Observations	165	165	165	104	104	105	165
Clusters	98	98	98	–	–	–	98
Border segments	24	24	24	16	16	16	16
Bandwidth	25 km	25 km	25 km	50 km	50 km	50 km	25 km
	Cubic in distance from Munich Agreement line, no bandwidth						
In borderlands	68.325	-.309	-.433	-3.091	-6.904	-1.109	.266
	(7.478)***	(.269)	(.330)	(6.404)	(4.694)	(1.508)	(6.385)
R^2	.951	.495	.389	.692	.447	.591	.485
Observations	272	272	272	110	109	111	272
Clusters	138	138	138	–	–	–	138
Border segments	24	24	24	16	16	16	24
	Mining and extraction	Metals	Manu-facturing	Glass	Textiles	Transport sector	Business sector
	(2a)	(2b)	(2c)	(2d)	(2e)	(2f)	(2g)
	Local conditional mean comparison, 10km bandwidth						
In borderlands	-.789	.920	-.438	1.248	-1.591	-.046	.360
	(1.181)	(1.355)	(.309)	(1.333)	(2.581)	(.496)	(.571)
R^2	.224	.32	.329	.284	.464	.287	.189
Observations	70	70	70	70	70	70	70
Clusters	53	53	53	53	53	53	53
Border segments	4	4	4	4	4	4	4
	Local linear in longitude and latitude, 25km bandwidth						
In borderlands	-.982	-.414	-.697	.230	.758	-.114	.483
	(.970)	(.911)	(.324)**	(.479)	(1.552)	(.420)	(.464)
R^2	.395	.339	.311	.467	.64	.321	.341
Observations	165	165	165	165	165	165	165
Clusters	98	98	98	98	98	98	98
Border segments	24	24	24	24	24	24	24
	Cubic in distance from Munich Agreement line, no bandwidth						
In borderlands	-1.751	2.432	.801	.898	-6.562	.040	-1.262
	(2.505)	(2.450)	(.999)	(2.365)	(4.682)	(.891)	(1.225)
R^2	.251	.251	.263	.287	.554	.289	.339
Observations	272	272	272	272	272	272	272
Clusters	138	138	138	138	138	138	138
Border segments	24	24	24	24	24	24	24

Robust standard errors clustered by political district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie and control for elevation, ruggedness, precipitation, temperature, and river density.

Table A.5: Pre-expulsion RD pre-trends, 1921-30

	Literacy	ln Population density	ln Labor force density	Agricultural sector
	(1a)	(1b)	(1c)	(1d)
$\Delta_{`30-`21}$In borderlands	.576	.019	.00005	-1.640
	(.163)***	(.021)	(.037)	(1.591)
R^2	.95	.519	.495	.895
	Industry	Construction	Transport sector	Business sector
	(2a)	(2b)	(2c)	(2d)
$\Delta_{`30-`21}$In borderlands	.597	-.049	-.226	-.283
	(1.418)	(.307)	(.215)	(.257)
R^2	.435	.876	.683	.862
Observations	330	330	330	330
Clusters	98	98	98	98
Border segments×1930	24	24	24	24
Geo. controls	Yes	Yes	Yes	Yes
District and year fixed effects	Yes	Yes	Yes	Yes
Bandwidth	25 km	25 km	25 km	25 km
Year	1921-30	1921-30	1921-30	1921-30

Robust standard errors are clustered by political district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie, include exogenous controls for elevation, ruggedness, precipitation, temperature, and river density interacted with year, and utilize a local linear running variable of distance from the Munich Agreement line interacted with the treatment and year. Since there were some splits and mergers of judicial districts between 1921 and 1930, I perform areal interpolation in ArcGIS to reshape a few 1921 districts into 1930 ones (see Table A.14 for details on procedure).

Table A.6: Balance tests (extended sample)

	% German	Literacy	ln Pop. density	Unemploy.	Income$_{pc}$	% Taxpayer	Agri. sector
	(1a)	(1b)	(1c)	(1d)	(1e)	(1f)	(1g)
In borderlands	66.681 (5.102)***	-.097 (.186)	-.301 (.189)	-2.048 (2.377)	-.053 (1.477)	.612 (.532)	4.666 (3.390)
R^2	.934	.541	.456	.68	.343	.507	.495
Observations	191	191	191	119	120	121	191
Clusters	104	104	104	–	–	–	104
Border segments	24	24	24	16	16	16	24
Bandwidth	25 km	25 km	25 km	50 km	50 km	50 km	25 km
Year	1930	1930	1930	1933	1933	1933	1930
	Mining and extraction	Metals	Manu-facturing	Glass	Textiles	Transport sector	Business sector
	(2a)	(2b)	(2c)	(2d)	(2e)	(2f)	(2g)
In borderlands	-.927 (1.180)	.679 (1.233)	-.414 (.583)	.922 (1.692)	-4.471 (2.355)*	-.775 (.615)	-.512 (.808)
R^2	.377	.319	.291	.35	.636	.294	.332
Observations	191	191	191	191	191	191	191
Clusters	104	104	104	104	104	104	104
Border segments	24	24	24	24	24	24	24
Bandwidth	25 km	25 km	25 km	25 km	25 km	25 km	25 km
Year	1930	1930	1930	1930	1930	1930	1930
Geo. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors clustered by political district, with *** and * denoting significance at the 1% and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie, include controls for elevation, ruggedness, precipitation, temperature, and river density, and utilize a local linear running variable of distance from the Munich Agreement line (MAL) interacted with the treatment. Relative to the main sample, this also includes districts lying mostly but not entirely in the borderlands that nonetheless had >80% Germans in 1930 (i.e. treated in spite of overlap) as well as those lying mostly but not entirely in the interior that nonetheless had <20% Germans.

Table A.7: Extent of ethnic diversity by region, 1930

	Ethnic fractionalization			
	(1a)	(1b)	(1c)	(1d)
In borderlands	.240 (.023)***	.265 (.040)***	.199 (.026)***	.204 (.031)***
R^2	.605	.629	.403	.564
Observations	70	165	123	218
Clusters	53	98	68	107
Border segments	4	24	4	24
Include overlapping districts?	No	No	Yes	Yes
Including distance polynomial?	No	Yes	No	Yes
Geo. controls	Yes	Yes	Yes	Yes
Bandwidth	10 km	25 km	10 km	25 km
Year	1930	1930	1930	1930

Robust standard errors are clustered by political district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie, include exogenous controls for elevation, ruggedness, precipitation, temperature, and river density. (1b) and (1d) utilize a local linear running variable of distance from the Munich Agreement line interacted with the treatment. (1c) and (1d) include districts that overlap the Munich Agreement line, on the basis that they are relevant as they are likely to be ethnically mixed. The ethnic fractionalization measure used here only takes into account the share of the population that was German (g) or Czechoslovak (c) on the 1930 census. Other ethnic groups in the Czech lands were of trivial size. Hence, this measure is given by $1 - g^2 - c^2$.

Table A.8: What kinds of places tended to be mixed?

	Literacy	ln Pop. density	Unemployment	Income_{pc}	Agri. sector	Mining and extraction
	(1a)	(1b)	(1c)	(1d)	(1e)	(1f)
Ethnic frac.	-.175	.739	1.491	6.415	-15.918	1.099
	(.299)	(.324)**	(4.174)	(2.779)**	(4.976)***	(2.456)
R^2	.462	.464	.596	.368	.481	.326
Observations	218	218	97	98	218	218
Clusters	107	107	–	–	107	107
Border segments	24	24	16	16	24	24
Year	1930	1930	1933	1933	1930	1930
	Metals	Manu-facturing	Glass	Textiles	Transport sector	Business sector
	(2a)	(2b)	(2c)	(2d)	(2e)	(2f)
Ethnic frac.	-.421	-.313	2.252	.452	.519	3.055
	(1.272)	(.617)	(1.886)	(2.971)	(.726)	(1.107)***
R^2	.302	.252	.354	.611	.295	.312
Observations	218	218	218	218	218	218
Clusters	107	107	107	107	107	107
Border segments	24	24	24	24	24	24
Year	1930	1930	1930	1930	1930	1930
Geo. controls	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	25 km	25 km	25 km	25 km	25 km	25 km

Robust standard errors are clustered by political district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie and control for elevation, ruggedness, precipitation, temperature, and river density. Regressions include all districts with centroids within 25 km of the Munich Agreement line, on the basis that those are the places most likely to be mixed in the Czech lands. The ethnic fractionalization measure used here only takes into account the share of the population that was German (g) or Czechoslovak (c) on the 1930 census. Other ethnic groups in the Czech lands were of trivial size. Hence, this measure is given by $1 - g^2 - c^2$.

Table A.9: Balance tests (concrete sample)

	% German	Literacy	ln Pop. density	Unemploy.	Income$_{pc}$	% Taxpayer	Agri. sector
	(1a)	(1b)	(1c)	(1d)	(1e)	(1f)	(1g)
In borderlands	86.052	-.253	-.179	-5.253	2.778	1.624	2.939
	(1.398)***	(.314)	(.183)	(5.421)	(1.846)	(.762)**	(3.359)
R^2	.996	.376	.474	.683	.557	.648	.498
Observations	105	105	105	60	60	60	105
Clusters	65	65	65	–	–	–	65
Border segments	4	4	4	4	4	4	4
Bandwidth	25 km	25 km	25 km	50 km	50 km	50 km	25 km
Year	1930	1930	1930	1933	1933	1933	1930
	Mining and extraction	Metals	Manu-facturing	Glass	Textiles	Transport sector	Business sector
	(2a)	(2b)	(2c)	(2d)	(2e)	(2f)	(2g)
In borderlands	-1.278	.026	-.270	.472	-6.266	-.461	-.079
	(.778)	(1.326)	(.515)	(1.785)	(4.077)	(.588)	(1.178)
R^2	.071	.173	.142	.188	.477	.173	.323
Observations	105	105	105	105	105	105	105
Clusters	65	65	65	65	65	65	65
Border segments	4	4	4	4	4	4	4
Bandwidth	25 km	25 km	25 km	25 km	25 km	25 km	25 km
Year	1930	1930	1930	1930	1930	1930	1930
Geo. controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors clustered by political district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie, include exogenous controls for elevation, ruggedness, precipitation, temperature, and river density, and utilize a local linear running variable of distance from the Munich Agreement line (MAL) interacted with the treatment. Due to the loss of various stretches of the MAL, I aggregate border segments into four, associated with North and South Bohemia and Moravia. The concrete sample was generated as follows:

1. First, I drop from the sample all districts that lie entirely in the borderlands or the interior which are not "homogeneous." I call these "mixed" districts. For the borderlands, I define "homogeneous" as being >80% German; for the interior, <20% German (or >80% Czechoslovak). However, stopping here would be problematic. Recall that the goal of this exercise is to check that places with many Germans (i.e. exposed to expulsion) and places with few were indeed otherwise *ex ante* similar around the MAL, while at the same time minimizing the likelihood that borderland Czechs and pre-treatment sorting around the MAL may have been biasing local district-level differences toward zero. Yet given what we know from history and Tables A.7-8 – that the borderlands was more mixed than the interior, and that borderland Czechs selected into urban areas – dropping only these mixed districts will bias the remaining borderlands sample toward being poorer and more rural on average relative to that of the interior.

2. Hence, we must also drop the other areas around the MAL in the neighborhood of these mixed districts – namely, the interior districts which correspond to them on the other side of the MAL that are not mixed yet are likely to be fundamentally similar, given the estimates in Tables 1.2 and 1.3. To do this, I first discretize the MAL in ArcGIS into just over 100,000 unique points.

3. I then perform a proximity analysis, wherein if a point on the MAL is nearer to the centroid of a homogeneous judicial district in the borderlands than that of a mixed district, I consider it to be part of a "concrete stretch" of the MAL (note: since judicial districts are less likely to be mixed in the interior and those which were "language islands" were not close to the MAL, I need not perform this for both regions). I then generate two files: one of concrete stretches of points and another of non-concrete.

4. But being on a concrete stretch need not mean the district which is closest will necessarily be concrete, even if that district is homogeneous; the district which is closest to that point may itself be closer to a different point. To determine whether a given *district* is concrete or not, I perform another proximity analysis among districts (note: for political districts, I use the same set of concrete and non-concrete points as generated by the less aggregated judicial district data). If a district is closer to a concrete point of the MAL, then I say that that district lies on a concrete stretch.

5. Finally, I drop all remaining districts that do not lie on a concrete stretch of the MAL. See Figures A.10-11 for the final concrete sample alongside a map showing the spatial distribution of Germans at the village level in the Czech lands prior to 1938. **We are now comparing only homogeneous parts of the borderlands with nearby homogeneous parts of the interior.**

Table A.10: RD robustness, long-run effects

	Unemployment	ln Population density	Agricultural sector	Finance and insurance	Auto repair and trade
	(1a)	(1b)	(1c)	(1d)	(1e)
	Local conditional mean comparison, 5km bandwidth				
In borderlands	3.042	-.224	-.565	-.271	-.649
	(.462)***	(.086)***	(.364)	(.082)***	(.270)**
R^2	.481	.457	.385	.174	.22
Observations	1201	1201	1201	1201	1201
Clusters	46	46	46	46	46
	Local linear in longitude and latitude, 25km bandwidth				
In borderlands	3.623	-.251	-.752	-.386	-.933
	(.520)***	(.084)***	(.457)	(.073)***	(.252)***
R^2	.398	.4	.304	.134	.201
Observations	4049	4049	4049	4049	4049
Clusters	71	71	71	71	71
	Cubic in distance from Munich Agreement line, no bandwidth				
In borderlands	1.827	-.264	-.788	-.250	-.320
	(.589)***	(.110)**	(.432)*	(.097)**	(.296)
R^2	.41	.378	.312	.205	.225
Observations	6112	6112	6112	6112	6112
Clusters	76	76	76	76	76
	Communications	Education	Healthcare	% Primary education or less	% Tertiary education
	(2a)	(2b)	(2c)	(2d)	(2e)
	Local conditional mean comparison, 5km bandwidth				
In borderlands	-.272	-.648	-.685	4.403	-1.743
	(.078)***	(.197)***	(.258)**	(.572)***	(.343)***
R^2	.201	.156	.236	.387	.281
Observations	1201	1201	1201	1201	1201
Clusters	46	46	46	46	46
	Local linear in longitude and latitude, 25km bandwidth				
In borderlands	-.381	-.791	-.780	4.965	-2.270
	(.082)***	(.145)***	(.233)***	(.543)***	(.375)***
R^2	.202	.085	.139	.298	.269
Observations	4049	4049	4049	4049	4049
Clusters	71	71	71	71	71
	Cubic in distance from Munich Agreement line, no bandwidth				
In borderlands	-.217	-.735	-.655	3.836	-1.334
	(.101)**	(.209)***	(.237)***	(.657)***	(.412)***
R^2	.337	.07	.095	.276	.326
Observations	6112	6112	6112	6112	6112
Clusters	76	76	76	76	76
Border segments	50	50	50	50	50
Geo. controls	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes
Year	2011	2011	2011	2011	2011

Robust standard errors clustered by district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie and include exogenous controls for elevation, ruggedness, precipitation, temperature, and river density.

Table A.11: Geographic balance tests (no mountainous stretches)

	Borderlands	Interior	Difference S.E.	Borderlands	Interior	Difference S.E.
Elevation	347.153	355.892	(9.160)	353.893	384.911	(5.772)***
Ruggedness	5.905	6.012	(.209)	6.541	6.211	(.132)**
Precipitation	50.654	50.598	(.385)	51.006	51.150	(.257)
Temperature	7.865	7.849	(.058)	7.673	7.727	(.034)
Rivers/km^2	.927	.977	(.039)	1.042	1.056	(.024)
Observations	284	424	708	728	1778	2506
Bandwidth	5 km	5 km	5 km	25 km	25 km	25 km
Arable land	55.782	51.672	(7.843)	42.572	49.172	(4.948)
Observations	8	7	15	19	22	41
Bandwidth	10 km	10 km	10 km	25 km	25 km	25 km
Year	1945	1945	1945	1945	1945	1945

Mean difference standard errors reported in parentheses, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All t-tests exclude Prague and Polish Zaolzie. Mountainous areas excluded include stretches of the Munich Agreement line that visibly closely follow the Sudete and Sumava ranges, as well as low-lying parts of the Ore range (see Figure A.9).

Table A.12: Long-run effects (no mountainous stretches)

	Unemploy.	ln Pop. density	ln Labor force density	% Primary edu. or less	% Second. education	% Tert. education
	(1a)	(1b)	(1c)	(1d)	(1e)	(1f)
In borderlands	2.973	-.338	-.352	4.608	-3.632	-1.772
	(.732)***	(.127)***	(.129)***	(.890)***	(.694)***	(.526)***
R^2	.446	.396	.4	.353	.223	.331
	Agri. sector	Auto repair and trade	Communi- cations	Finance and insurance	Education	Healthcare
	(2a)	(2b)	(2c)	(2d)	(2e)	(2f)
In borderlands	-.921	-.589	-.312	-.441	-.782	-1.017
	(.585)	(.349)*	(.104)***	(.099)***	(.210)***	(.319)***
R^2	.308	.217	.271	.144	.1	.145
Observations	2506	2506	2506	2506	2506	2506
Clusters	57	57	57	57	57	57
Geo. controls	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Border fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	25 km	25 km	25 km	25 km	25 km	25 km
Year	2011	2011	2011	2011	2011	2011

Robust standard errors are clustered by district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie, include exogenous controls for elevation, ruggedness, precipitation, temperature, and river density, and utilize a local linear running variable of distance from the Munich Agreement line interacted with the treatment. A municipality is omitted if it lies closer to one of the mountainous stretches highlighted in Figure A.9 than to any other part of the Munich Agreement line. 50 border segment dummies are included, though 19 are dropped by removing mountainous stretches.

Table A.13: Relative net population decline, 1930 to mid-1947

	Labor force % change	Agricultural % change	Secondary sector % change	Transport sector % change	Business sector % change
	(1a)	(1b)	(1c)	(1d)	(1e)
In borderlands	-12.408	-9.290	-13.220	-27.023	-32.913
	(3.175)***	(3.313)***	(5.314)**	(7.180)***	(4.455)***
R^2	.732	.549	.727	.572	.844
Observations	165	165	165	165	165
Clusters	98	98	98	98	98
Border segments	24	24	24	24	24
Geo. controls	Yes	Yes	Yes	Yes	Yes
Bandwidth	25 km	25 km	25 km	25 km	25 km
Year	1930-47	1930-47	1930-47	1930-47	1930-47

Robust standard errors are clustered by political district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie, include exogenous controls for elevation, ruggedness, precipitation, temperature, and river density, and utilize a local linear running variable of distance from the Munich Agreement line interacted with the treatment.

Table A.14: Relative density change, 1921-2011

	ln Population density	ln Labor force density	ln Pop. density	ln L.F. density
	(1a)	(1b)	(2a)	(2b)
	Assignment by no overlap MAL, 50km		Assignment by majority overlap	
In borderlands×1921	-.004 (.034)	-.030 (.045)	-.011 (.015)	-.014 (.020)
In borderlands×1930	0	0	0	0
In borderlands×1947	-.443 (.054)***	-.367 (.065)***	-.310 (.036)***	-.281 (.036)***
In borderlands×1950	-.415 (.060)***	-.388 (.074)***	-.298 (.036)***	-.301 (.035)***
In borderlands×1961	-.391 (.076)***	–	-.299 (.054)***	–
In borderlands×1970	-.312 (.070)***	-.299 (.069)***	-.272 (.056)***	-.260 (.055)***
In borderlands×1980	-.252 (.070)***	-.262 (.073)***	-.238 (.058)***	-.243 (.059)***
In borderlands×1991	-.214 (.075)***	-.222 (.079)**	-.214 (.060)***	-.216 (.061)***
In borderlands×2001	-.214 (.073)***	-.231 (.076)***	-.203 (.059)***	-.208 (.061)***
In borderlands×2011	-.315 (.084)***	-.353 (.094)***	-.245 (.056)***	-.267 (.062)***
R^2	.607	.809	.423	.701
Observations	410	369	730	657
Clusters	41	41	73	73
District fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes

Robust standard errors are clustered by district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie and include controls for longitude, latitude, and each interacted with census. To construct common district boundaries used for this panel analysis and others, I use the "intercept" tool in ArcGIS software to interpolate population and subpopulations. I use 1991 boundaries as the standard, since districts were arguably at their highest level of aggregation that year, minimizing error. The "intercept" tool creates subsets of districts based on where a given census year's district boundaries overlapped with those from 1991. For example, if a 1921 judicial district lied completely within a 1991 district, that judicial district would only have one subset: itself. If it straddled the line of two 1991 districts, it would have two subsets. Adopting the assumption that a given census' district's subpopulations were uniformly distributed within its boundaries, I estimate the number of individuals in various subpopulations (e.g. number of farmers) within each district subset. I then add up these estimates within the boundaries of each 1991 district.

Table A.15: Relative sectoral change, 1921-2011

	Agricultural sector	Industry	Service sector	Agricultural sector	Industry	Service sector
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
	Assignment by no overlap MAL, 50km			Assignment by majority overlap		
In borderlands×'21	-2.550 (1.607)	.374 (1.386)	1.709 (.572)***	-.954 (.739)	.115 (.691)	.835 (.323)**
In borderlands×'47	4.521 (2.038)**	–	-2.933 (1.227)**	3.277 (.847)***	–	-1.466 (.662)**
In borderlands×'50	5.916 (1.914)***	-3.518 (2.047)*	-2.831 (1.667)*	3.725 (.889)***	-2.676 (1.026)**	-1.272 (.844)
In borderlands×'61	–	-7.883 (2.437)***	–	–	-4.639 (1.238)***	–
In borderlands×'70	5.141 (1.355)***	-7.347 (2.068)***	-1.566 (1.295)	3.025 (.982)***	-4.297 (1.201)***	-.284 (.904)
In borderlands×'80	6.341 (1.258)***	-6.849 (2.422)***	-1.831 (1.415)	3.186 (1.125)***	-3.539 (1.339)***	-.862 (.862)
In borderlands×'91	6.470 (1.407)***	-7.484 (2.617)***	-1.556 (1.485)	3.148 (1.200)**	-3.975 (1.450)***	-.523 (.842)
In borderlands×'01	7.429 (2.202)***	-4.862 (4.644)	-5.100 (1.857)***	2.839 (1.600)*	-1.840 (1.998)	-2.282 (.952)**
In borderlands×'11	7.750 (2.495)***	-6.292 (4.657)	-8.830 (2.080)***	2.845 (1.748)	-2.621 (2.057)	-5.050 (1.070)***
Constant	36.156 (.998)***	32.361 (.903)***	14.310 (.486)***	35.754 (.829)***	32.413 (.715)***	14.686 (.340)***
1930	-12.308 (.779)***	2.593 (.721)***	4.329 (.277)***	-11.329 (.589)***	2.160 (.559)***	4.155 (.202)***
1947	-14.627 (1.118)***	–	12.003 (.521)***	-12.899 (.820)***	–	11.478 (.355)***
1950	-16.406 (1.131)***	13.175 (.923)***	14.054 (.759)***	-14.429 (.894)***	11.878 (.692)***	13.563 (.479)***
1961	–	13.317 (1.064)***	–	–	11.080 (.819)***	–
1970	-20.384 (1.303)***	13.535 (1.022)***	.756 (.761)	-17.627 (1.125)***	11.450 (.950)***	.281 (.627)
1980	-24.049 (1.326)***	13.259 (1.247)***	19.767 (.739)***	-21.167 (1.233)***	11.111 (1.092)***	18.937 (.555)***
1991	-24.829 (1.398)***	9.773 (1.355)***	23.021 (.840)***	-21.920 (1.278)***	7.650 (1.176)***	22.021 (.606)***
2001	-33.732 (1.721)***	1.616 (2.315)	33.655 (1.101)***	-31.000 (1.572)***	.108 (1.593)	32.314 (.716)***
2011	-36.176 (1.813)***	-4.684 (2.262)**	30.026 (1.287)***	-33.580 (1.653)***	-6.402 (1.549)***	28.674 (.806)***
R^2	.891	.738	.957	.872	.719	.958
Observations	369	369	369	657	657	657
Clusters	41	41	41	73	73	73
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors are clustered by district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie and include controls for longitude, latitude, and each interacted with census. To construct common district boundaries used for this panel analysis and others, I perform areal interpolation as described in Table A.14.

Table A.16: Short-run supply of education, mid-1947

	General schools per 100 pupils	General teachers per 100 pupils	Civic schools per 100 pupils	Civic teachers per 100 pupils
	(1a)	(1b)	(2a)	(2b)
In borderlands	.056 (.221)	-.162 (.147)	.069 (.062)	-.234 (.235)
R^2	.46	.529	.296	.243
Observations	115	115	115	115
	Agricultural schools per 100 pupils	Agricultural teachers per 100 pupils	Vocational schools per 100 pupils	Vocational teachers per 100 pupils
	(3a)	(3b)	(4a)	(4b)
In borderlands	-1.228 (.359)***	-1.933 (1.854)	.257 (.173)	2.537 (1.343)*
R^2	.397	.287	.195	.241
Observations	104	99	97	97
Border segments	16	16	16	16
Geo. controls	Yes	Yes	Yes	Yes
Bandwidth	50 km	50 km	50 km	50 km
Year	1947	1947	1947	1947

Robust standard errors reported in brackets, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie, include controls for elevation, ruggedness, precipitation, temperature, and river density, and utilize a local linear running variable of distance from the Munich Agreement line interacted with the treatment. Note that some districts have no vocational or agricultural schools. Also note that some agricultural teacher data is missing for a few larger cities with few (e.g. 1) agricultural folk schools. If I assign a value of 1 for teachers in these cities, the coefficient remains negative and insignificant.

Table A.17: Regional differences in education, 1921-2011

	Education index	% Second. education	% Tert. education	Edu. index	% Second. education	% Tert. education
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
	Assignment by no overlap MAL, 50km			Assignment by majority overlap		
In borderlands×'21	-1.090	–	–	-.755	–	–
	(.104)***			(.103)***		
In borderlands×'30	0	–	–	0	–	–
In borderlands×'61	-1.759	-2.087	-1.312	-1.204	-1.611	-1.180
	(.255)***	(.486)***	(.508)**	(.228)***	(.294)***	(.314)***
In borderlands×'70	-1.562	-2.735	-.272	-1.156	-2.618	-.374
	(.314)***	(1.118)**	(.258)	(.257)***	(.676)***	(.176)**
In borderlands×'80	-1.968	-3.930	-.487	-1.324	-2.884	-.529
	(.280)***	(.871)***	(.355)	(.272)***	(.611)***	(.231)**
In borderlands×'91	-2.527	-4.745	-.968	-1.686	-3.242	-.926
	(.253)***	(.497)***	(.495)*	(.277)***	(.438)***	(.301)***
In borderlands×'01	-3.002	-3.859	-2.603	-1.891	-2.403	-1.876
	(.266)***	(.522)***	(.736)***	(.283)***	(.362)***	(.415)***
In borderlands×'11	-3.093	-2.233	-4.603	-1.926	-1.318	-3.131
	(.286)***	(.8205)**	(1.263)***	(.271)***	(.461)***	(.587)***
R^2	.627	.989	.802	.433	.988	.826
Observations	328	246	246	584	438	438
Clusters	41	41	41	73	73	73
District fixed effects	Yes	No	No	Yes	No	No
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors are clustered by district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie and include controls for longitude, latitude, and each interacted with census. The education index uses prewar literacy and postwar post-primary education data transformed into standard deviations from census year district means. To construct common district boundaries used for this panel analysis and others, I perform areal interpolation as described in Table A.14.

Table A.18: RD robustness, short-run effects

	ln Population density	Agricultural sector	Secondary sector	Transport sector	Business sector
	(1a)	(1b)	(1c)	(1d)	(1e)
	Local conditional mean comparison, 10km bandwidth				
$\Delta_{`47-`30}$In borderlands	-.270	5.088	-2.308	-.276	-1.889
	(.033)***	(1.222)***	(1.366)*	(.306)	(.340)***
R^2	.898	.421	.612	.58	.497
Observations	140	140	140	140	140
Clusters	53	53	53	53	53
Border segments×1947	4	4	4	4	4
	Local linear in longitude and latitude, 25km bandwidth				
$\Delta_{`47-`30}$In borderlands	-.296	6.390	-3.389	-.003	-1.959
	(.028)***	(1.129)***	(1.064)***	(.294)	(.244)***
R^2	.914	.584	.748	.685	.625
Observations	330	330	330	330	330
Clusters	98	98	98	98	98
Border segments×1947	24	24	24	24	24
	Cubic in distance from Munich Agreement line, no bandwidth				
$\Delta_{`47-`30}$In borderlands	-.188	2.407	2.707	-.145	-1.214
	(.078)**	(2.648)	(2.559)	(.573)	(.768)
R^2	.908	.518	.721	.651	.651
Observations	544	544	544	544	544
Clusters	138	138	138	138	138
Border segments×1947	24	24	24	24	24
District fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
	Enrollment, $\text{general}_{5\text{-}14}$	Enrollment, $\text{civic}_{10\text{-}14}$	Enrollment, $\text{agricultural}_{15\text{-}19}$	Enrollment, $\text{vocational}_{15\text{-}19}$	Enrollment, $\text{college}_{15\text{-}24}$
	(2a)	(2b)	(2c)	(2d)	(2e)
	Local conditional mean comparison, 10km bandwidth				
In borderlands	7.805	-11.840	4.386	-11.910	-2.540
	(1.236)***	(3.239)***	(1.672)**	(5.706)*	(.349)***
R^2	.779	.611	.63	.33	.746
Observations	25	25	25	25	25
Border segments	4	4	4	4	4
	Local linear in longitude and latitude, 50km bandwidth				
In borderlands	6.959	-7.669	5.562	-6.293	-2.574
	(.728)***	(2.395)***	(1.374)***	(4.398)	(.374)***
R^2	.86	.591	.517	.214	.724
Observations	115	115	115	115	115
Border segments	16	16	16	16	16
	Cubic in distance from Munich Agreement line, no bandwidth				
In borderlands	4.899	-7.067	7.725	-26.282	-3.405
	(2.180)**	(5.545)	(4.009)*	(12.246)**	(.901)***
R^2	.87	.597	.5	.28	.73
Observations	122	122	122	122	122
Border segments	16	16	16	16	16
Geo. controls	Yes	Yes	Yes	Yes	Yes
Year	1947	1947	1947	1947	1947

Robust standard errors clustered by political district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie and control for elevation, ruggedness, precipitation, temperature, and river density.

Table A.19: Short-run effects (extended sample)

	ln Population density	Agricultural sector	Secondary sector	Transport sector	Business sector
	(1a)	(1b)	(1c)	(1d)	(1e)
$\Delta_{'47-'30}$In borderlands	-.192	3.689	-.650	-.382	-1.138
	(.045)***	(1.816)**	(1.801)	(.345)	(.440)**
R^2	.914	.563	.745	.697	.644
Observations	382	382	382	382	382
Clusters	104	104	104	104	104
Border segments×1947	24	24	24	24	24
Bandwidth	25 km	25 km	25 km	25 km	25 km
	Enrollment, $\text{general}_{5\text{-}14}$	Enrollment, $\text{civic}_{10\text{-}14}$	Enrollment, $\text{agricultural}_{15\text{-}19}$	Enrollment, $\text{vocational}_{15\text{-}19}$	Enrollment, $\text{college}_{15\text{-}24}$
	(2a)	(2b)	(2c)	(2d)	(2e)
In borderlands	5.492	-8.011	5.372	-10.818	-2.558
	(.874)***	(2.422)***	(1.726)***	(4.366)**	(.366)***
R^2	.868	.562	.496	.214	.661
Observations	131	131	131	131	131
Border segments	16	16	16	16	16
Bandwidth	50 km	50 km	50 km	50 km	50 km
Geo. controls	Yes	Yes	Yes	Yes	Yes
Year	1947	1947	1947	1947	1947

Robust standard errors clustered by political district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie, include controls for elevation, ruggedness, precipitation, temperature, and river density, and utilize a local linear running variable of distance from the Munich Agreement line (MAL) interacted with the treatment. Relative to the main sample, this also includes districts lying mostly but not entirely in the borderlands that nonetheless had >80% Germans in 1930 (i.e. treated in spite of overlap) as well as those lying mostly but not entirely in the interior that nonetheless had <20% Germans.

Table A.20: Pre-expulsion agglomeration economies

	Income$_{pc}$		Agricultural sector		Industry	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)
ln Pop density	5.036 (.586)***	5.247 (.630)***	-13.729 (2.311)***	-13.607 (2.293)***	8.004 (2.947)***	7.905 (3.004)***
In borderlands	–	1.973 (1.335)	–	-1.417 (2.368)	–	-.745 (3.022)
R^2	.757	.767	.767	.781	.65	.653
Observations	104	104	165	165	165	165
Clusters	–	–	98	98	98	98
Border segments	16	16	24	24	24	24
Bandwidth	50 km	50 km	25 km	25 km	25 km	25 km
Year	1933	1933	1930	1930	1930	1930
	Construction		Transport sector		Business sector	
	(4a)	(4b)	(5a)	(5b)	(6a)	(6b)
ln Pop density	-.987 (.279)***	-.996 (.276)***	1.135 (.374)***	1.125 (.375)***	3.034 (.316)***	3.018 (.302)***
In borderlands	–	-.183 (.709)	–	-.099 (.629)	–	.317 (.564)
R^2	.378	.383	.403	.404	.669	.692
Observations	165	165	165	165	165	165
Clusters	98	98	98	98	98	98
Border segments	24	24	24	24	24	24
Bandwidth	25 km	25 km	25 km	25 km	25 km	25 km
Year	1930	1930	1930	1930	1930	1930
Geo. controls	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors are clustered by political district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie and control for elevation, ruggedness, precipitation, temperature, and river density. Regressions (b) utilize a local linear running variable of distance from the Munich Agreement line interacted with the treatment.

Table A.21: Net migration

	Net migrat.	In migrat.	Outmigrat.	Net mig.	In mig.	Outmig.
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
	Assignment by no overlap, 50km			Assignment by majority overlap		
In borderlands×1950	-.434	3.067	3.501	-.470	1.760	2.230
	(.349)	(.471)***	(.512)***	(.221)**	(.348)***	(.312)***
In borderlands×1961	-.507	.928	1.435	-.492	.477	.969
	(.284)*	(.288)***	(.226)***	(.161)***	(.180)***	(.138)***
In borderlands×1970	-.488	.667	1.155	-.329	.385	.713
	(.145)***	(.305)**	(.264)***	(.097)***	(.156)**	(.137)***
In borderlands×1980	-.616	-.142	.475	-.316	-.056	.260
	(.121)***	(.146)	(.149)***	(.090)***	(.100)	(.087)***
In borderlands×1991	.006	.005	-.001	.023	.003	-.019
	(.067)	(.141)	(.131)	(.041)	(.072)	(.060)
In borderlands×2001	-.549	-.604	-.055	-.269	-.280	-.011
	(.232)**	(.289)**	(.111)	(.096)***	(.122)**	(.062)
In borderlands×2011	-.907	-.958	-.051	-.417	-.485	-.068
	(.263)***	(.336)***	(.109)	(.124)***	(.149)***	(.071)
R^2	.232	.846	.904	.182	.781	.859
Observations	287	287	287	511	511	511
Clusters	41	41	41	73	73	73
District fixed effects	No	No	No	No	No	No
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors are clustered by district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie and include controls for longitude, latitude, and each interacted with census. To construct common district boundaries used for this panel analysis and others, I perform areal interpolation as described in Table A.14.

Table A.22: Heterogeneous effects, natural geography

	Unemployment	ln Population density	% Primary education or less
	(1a)	(1b)	(1c)
In borderlands×River density	-.956	.230	-1.474
	(.630)	(.077)***	(.415)***
R^2	.405	.401	.3
	(2a)	(2b)	(2c)
In borderlands×Ruggedness	.004	.034	-.129
	(.081)	(.019)*	(.100)
R^2	.404	.4	.298
Observations	4049	4049	4049
Clusters	71	71	71
Border segments	50	50	50
Geo. controls	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes
Bandwidth	25 km	25 km	25 km
Year	2011	2011	2011

Robust standard errors are clustered by district, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. All regressions exclude Prague and Polish Zaolzie, include controls for elevation, ruggedness, precipitation, temperature, and river density, and utilize a local linear running variable of distance from the Munich Agreement line interacted with the treatment. For these regressions, river density and ruggedness are mean-normalized.

Table A.23: Post-transition trends, 2001-11

	Unemployment	ln Pop. density	Agricultural sector	Industry	Construction
	(1a)	(1b)	(1c)	(1d)	(1e)
$\Delta_{'11-'01}$In borderlands	-.205	.005	-.307	-.501	-.904
	(.368)	(.012)	(.665)	(.602)	(.276)***
R^2	.392	.318	.632	.522	.369
	Auto repair and trade	Transport+ comm.	Public	Education+ healthcare	% Primary edu. or less
	(2a)	(2b)	(2c)	(2d)	(2e)
$\Delta_{'11-'01}$In borderlands	-.021	.147	-.125	-.266	-.483
	(.242)	(.292)	(.190)	(.219)	(.321)
R^2	.094	.133	.076	.162	.757
Observations	8088	8088	8088	8088	8088
Clusters	71	71	71	71	71
Border segments×2011	50	50	50	50	50
Geo. controls	Yes	Yes	Yes	Yes	Yes
District fixed effects×2011	Yes	Yes	Yes	Yes	Yes
Municipality fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Bandwidth	25 km	25 km	25 km	25 km	25 km

Robust standard errors are clustered by district, with *** denoting significance at the 1% level. All regressions exclude Prague and Polish Zaolzie, include exogenous controls for elevation, ruggedness, precipitation, temperature, and river density interacted with year, and utilize a local linear running variable of distance from the Munich Agreement line interacted with the treatment and year. Since a few municipalities split or merged between 2001 and 2011, I manually aggregate these and their data.

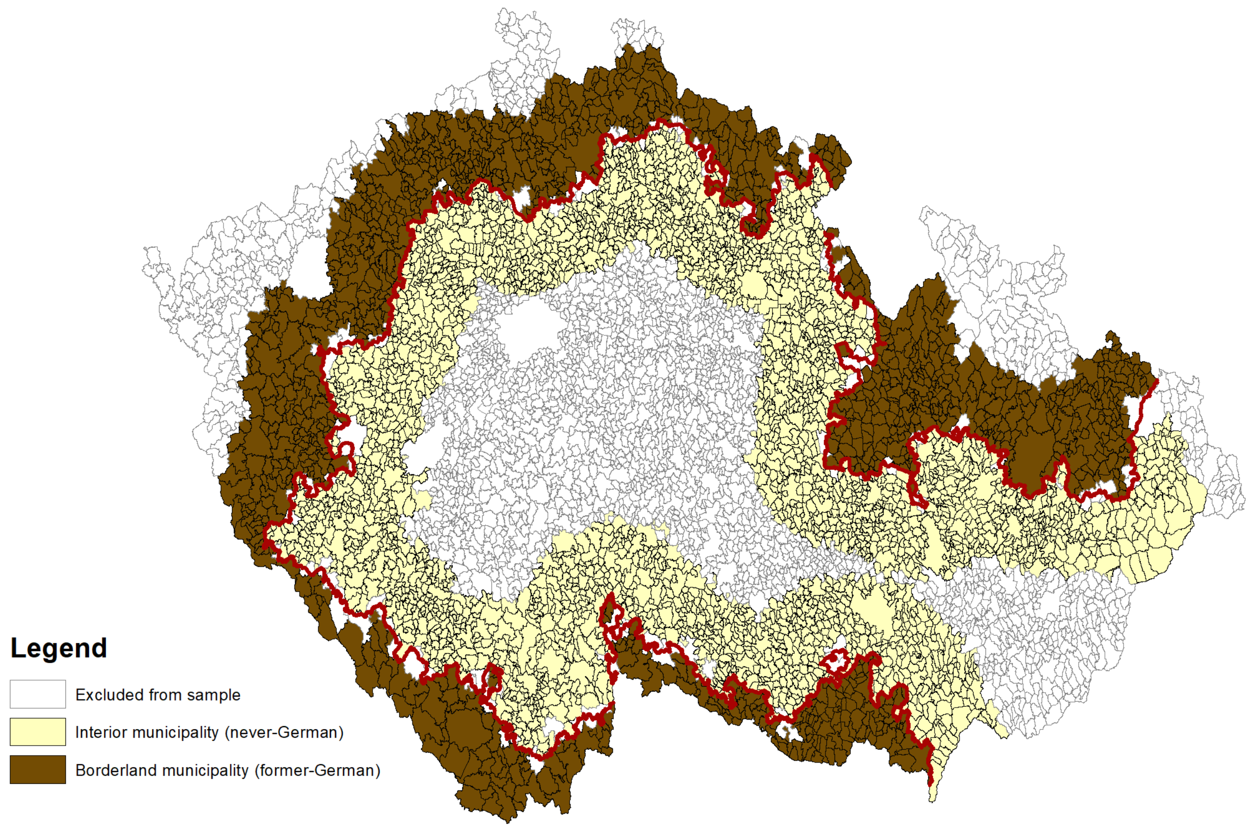
Table A.24: Heterogeneous effects, Eastern Bloc

	Unemploy.	ln Pop. density	ln L.F. density	% Primary edu. or less	% Second. education	% Tertiary education
	(1a)	(1b)	(1c)	(1d)	(1e)	(1f)
In borderlands ×Eastern Bloc	.780	.009	-.013	1.669	-.959	-.669
	(.983)	(.165)	(.168)	(1.234)	(.983)	(.740)
R^2	.406	.402	.403	.301	.201	.272
	Agricultural sector	Auto repair and trade	Commun- ications	Finance and insurance	Education	Healthcare
	(2a)	(2b)	(2c)	(2d)	(2e)	(2f)
In borderlands ×Eastern Bloc	.123	.429	-.007	.047	.221	.539
	(.824)	(.521)	(.169)	(.164)	(.332)	(.379)
R^2	.305	.204	.206	.137	.086	.143
Observations	4049	4049	4049	4049	4049	4049
Clusters	71	71	71	71	71	71
Border segments	50	50	50	50	50	50
Geo. controls	Yes	Yes	Yes	Yes	Yes	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	25 km	25 km	25 km	25 km	25 km	25 km
Year	2011	2011	2011	2011	2011	2011

Robust standard errors clustered by district. A municipality is dummied 1 if it lied closer to East Germany/Poland pre-1989 than West Germany/Austria. All regressions exclude Prague and Polish Zaolzie, include exogenous controls for elevation, ruggedness, precipitation, temperature, and river density, and use a linear running variable of distance from MAL interacted with the treatment, both interacted with Eastern Bloc.

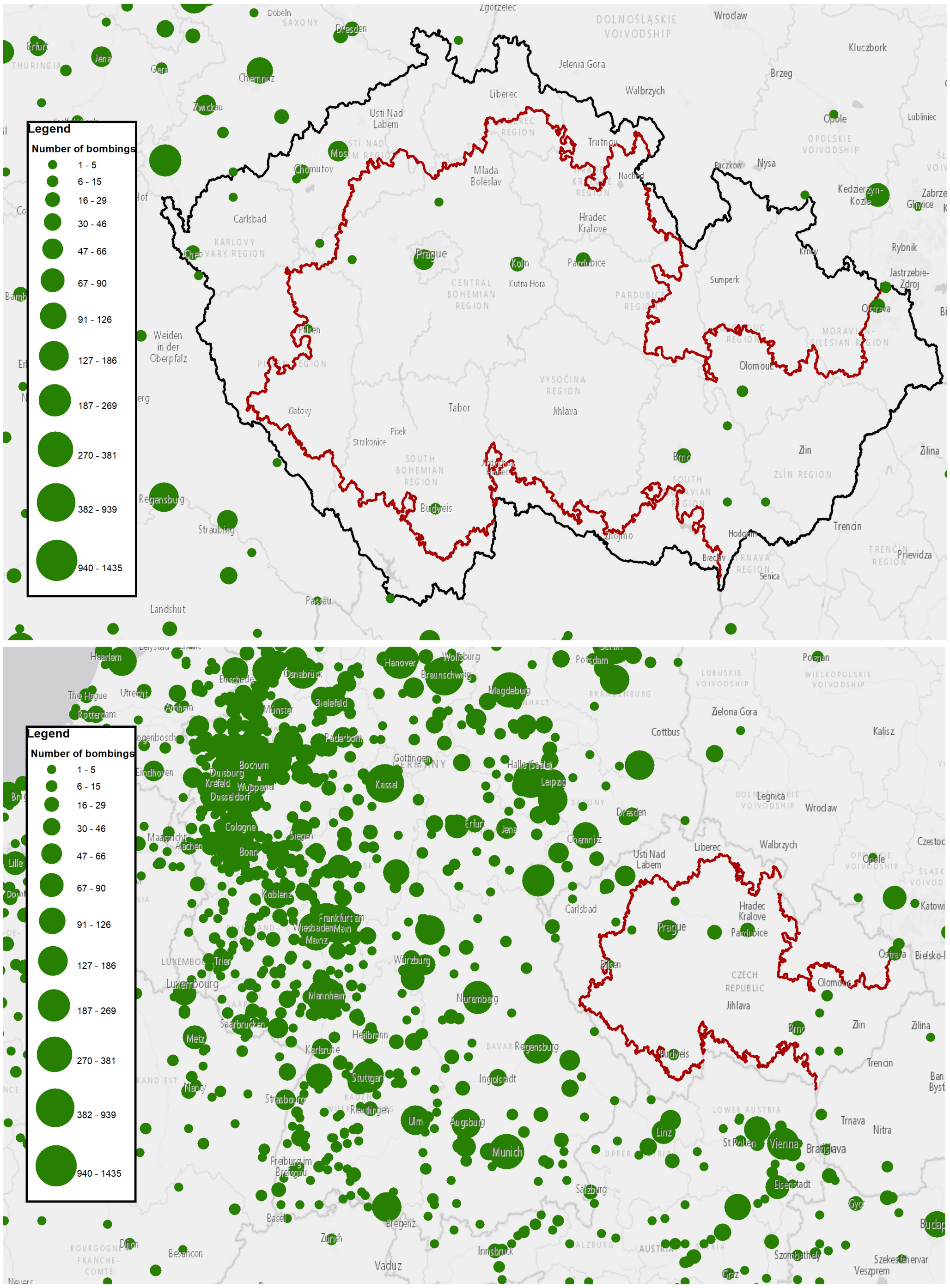
Figures

Figure A.1: Municipalities in main sample



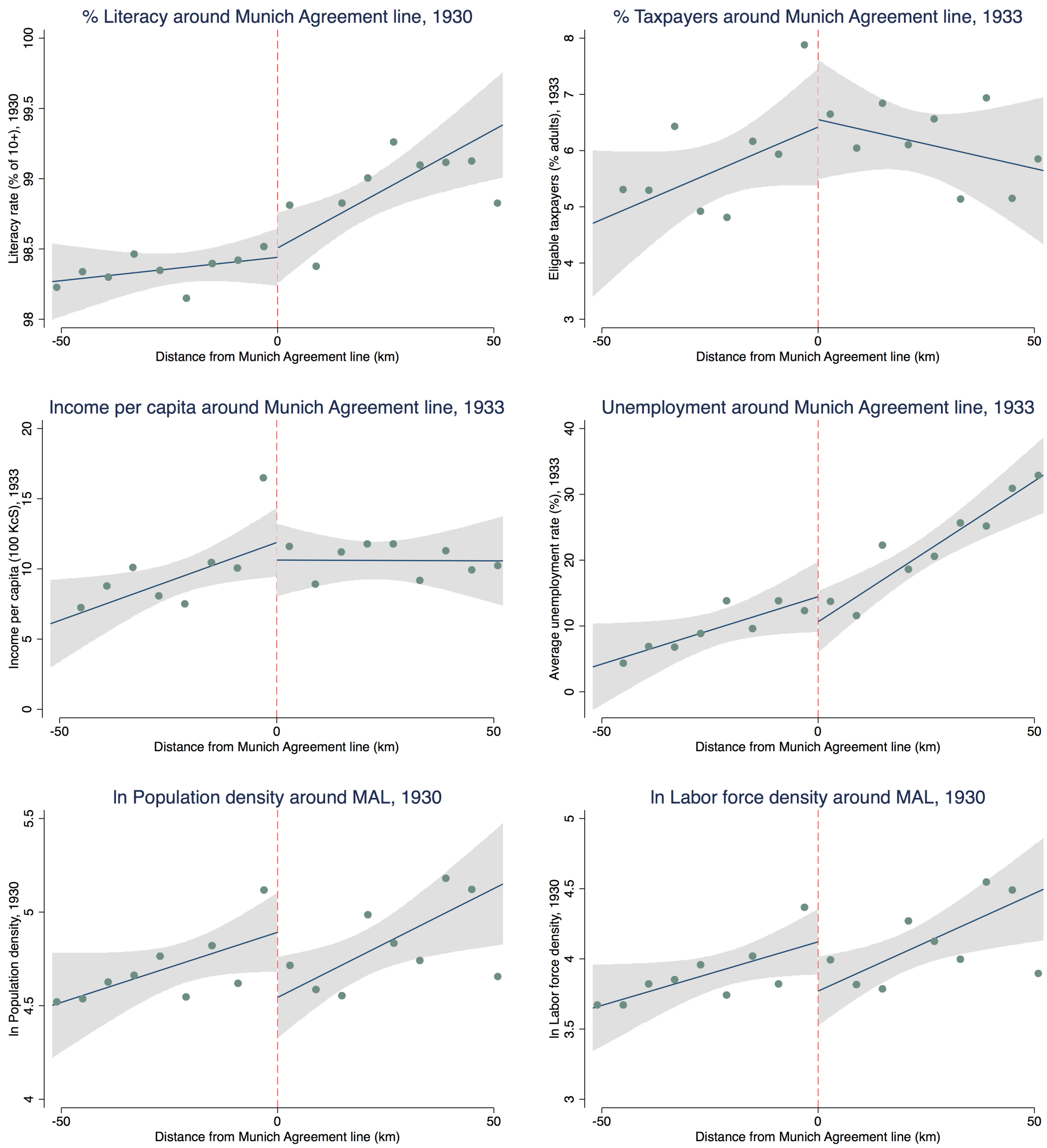
Note: 94 municipalities for which only some parts were annexed are dropped. Municipalities in Polish Zaolzie (i.e. the strip of white municipalities to the right of the Munich Agreement line on the far right of the map) are also excluded from all analyses.

Figure A.2: Allied bombings during World War II



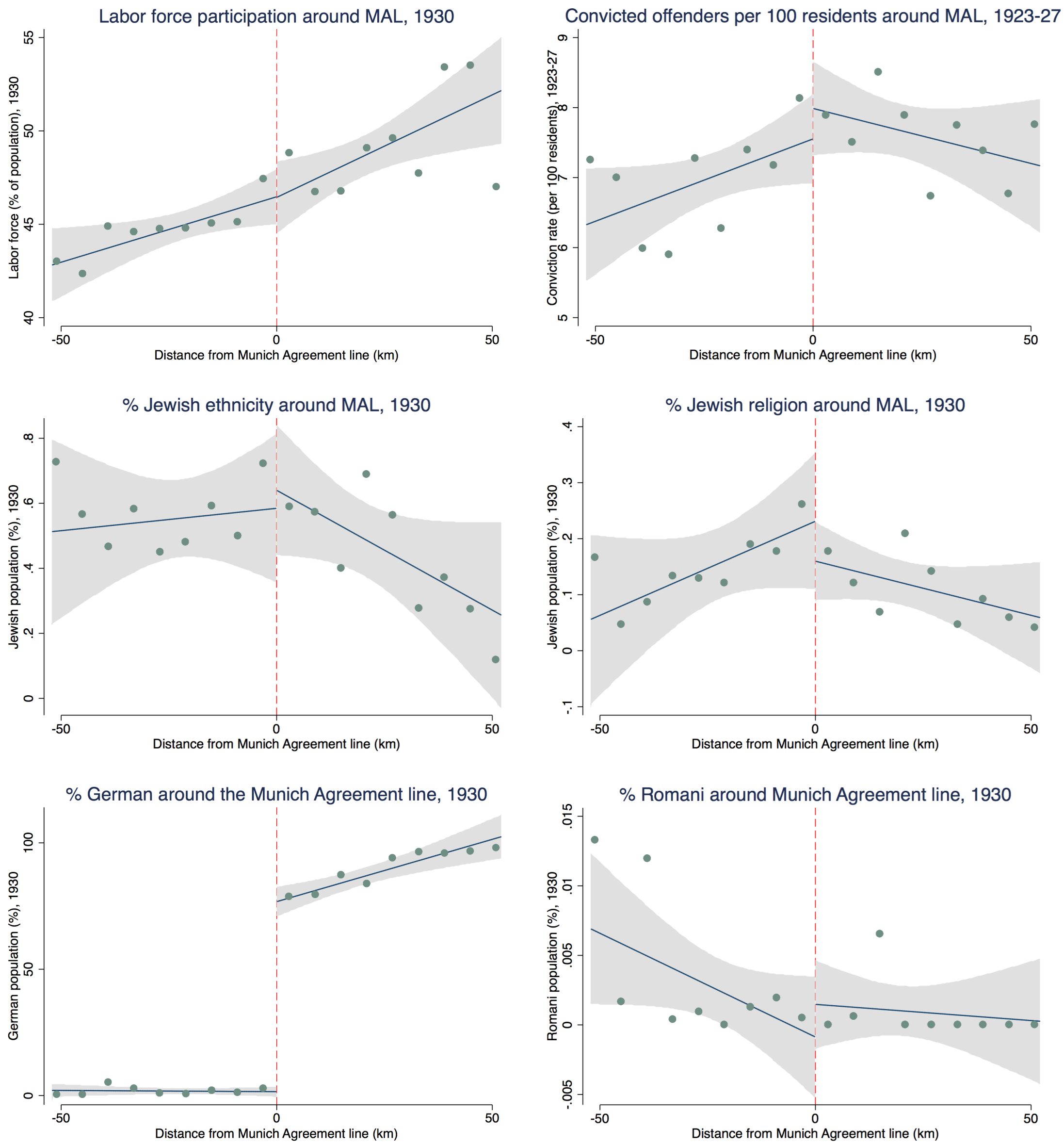
This map shows confirmed Allied bombing sites during World War II, first relative to the Munich Agreement line and then relative to relevant nearby territories (source: Theatre History of Operations Reports (THOR), 2019). Nearly all took place in late 1944 or 1945. Note: this drops observations for which coordinates were not specified.

Figure A.3: Balance tests, RD plots



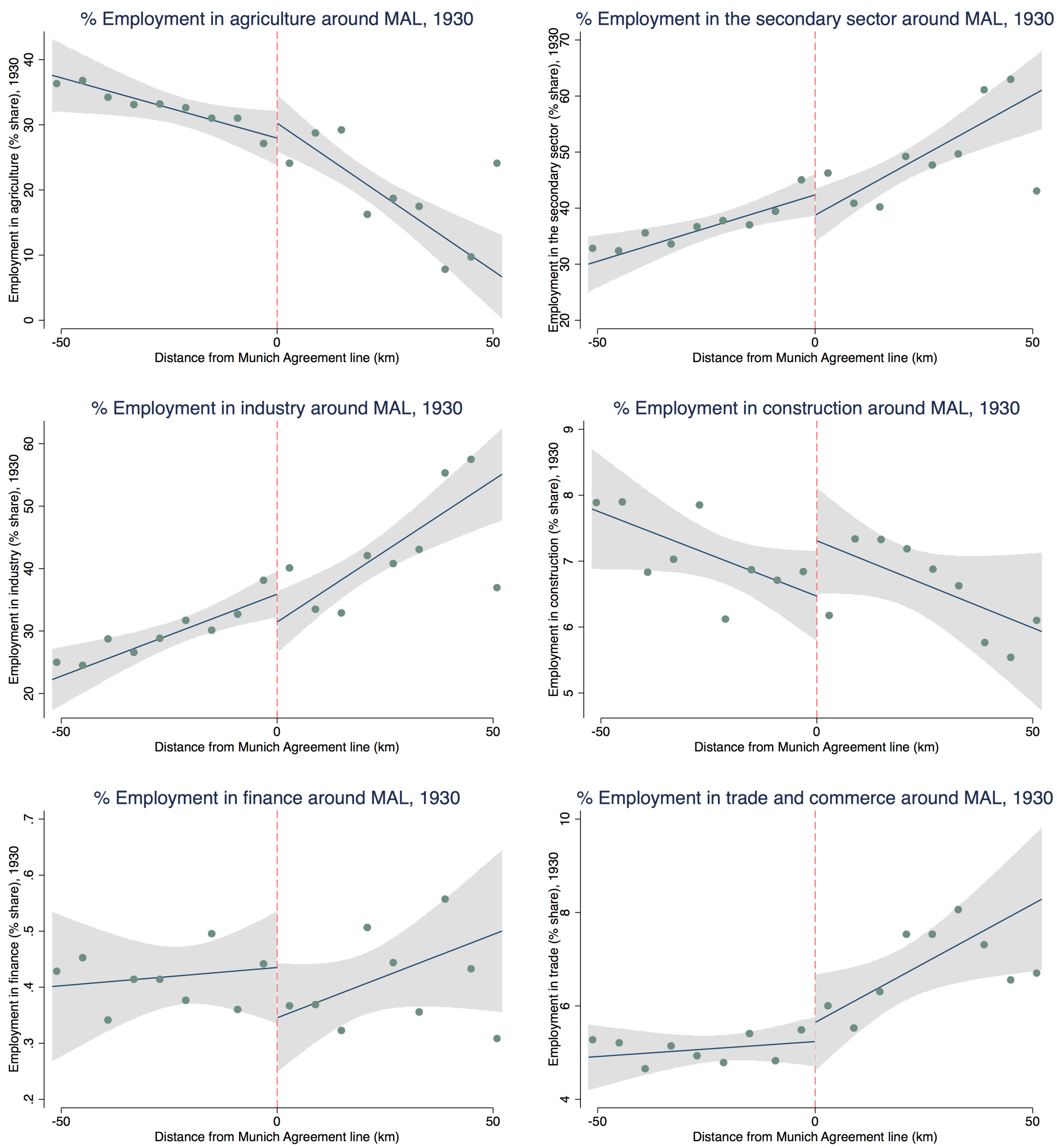
Trend lines are linear over the domain on each side of the Munich Agreement line. Data binned by 6 km intervals to produce 18 mean data points. All plots exclude Prague, Polish Zaolzie, and administrative units that overlap the Munich Agreement line. Shaded areas represent 95% confidence intervals.

Figure A.3: Balance tests, RD plots (II)



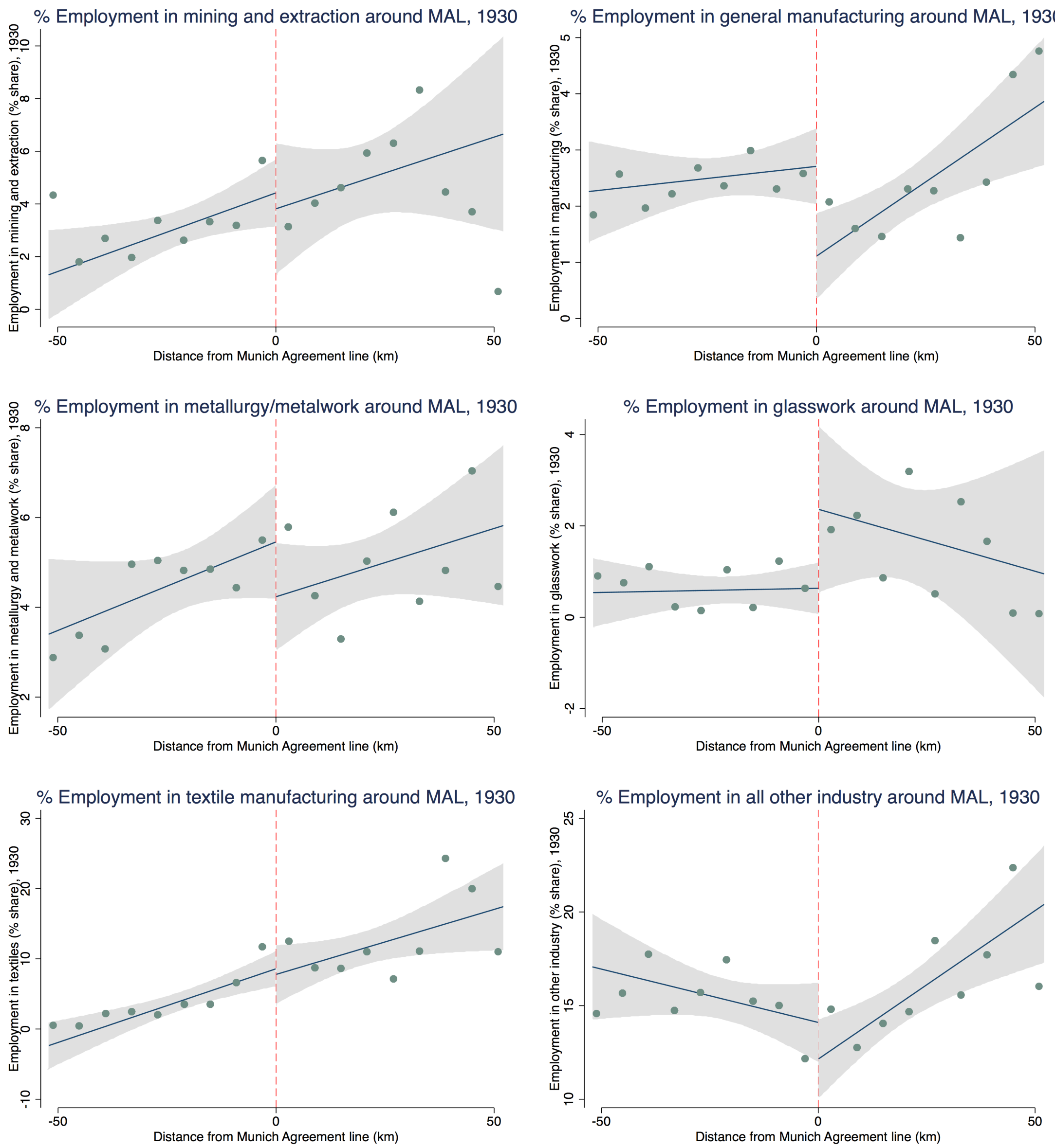
Trend lines are linear over the domain on each side of the Munich Agreement line. Data binned by 6 km intervals to produce 18 mean data points. All plots exclude Prague, Polish Zaolzie, and administrative units that overlap the Munich Agreement line. Shaded areas represent 95% confidence intervals.

Figure A.3: Balance tests, RD plots (III)



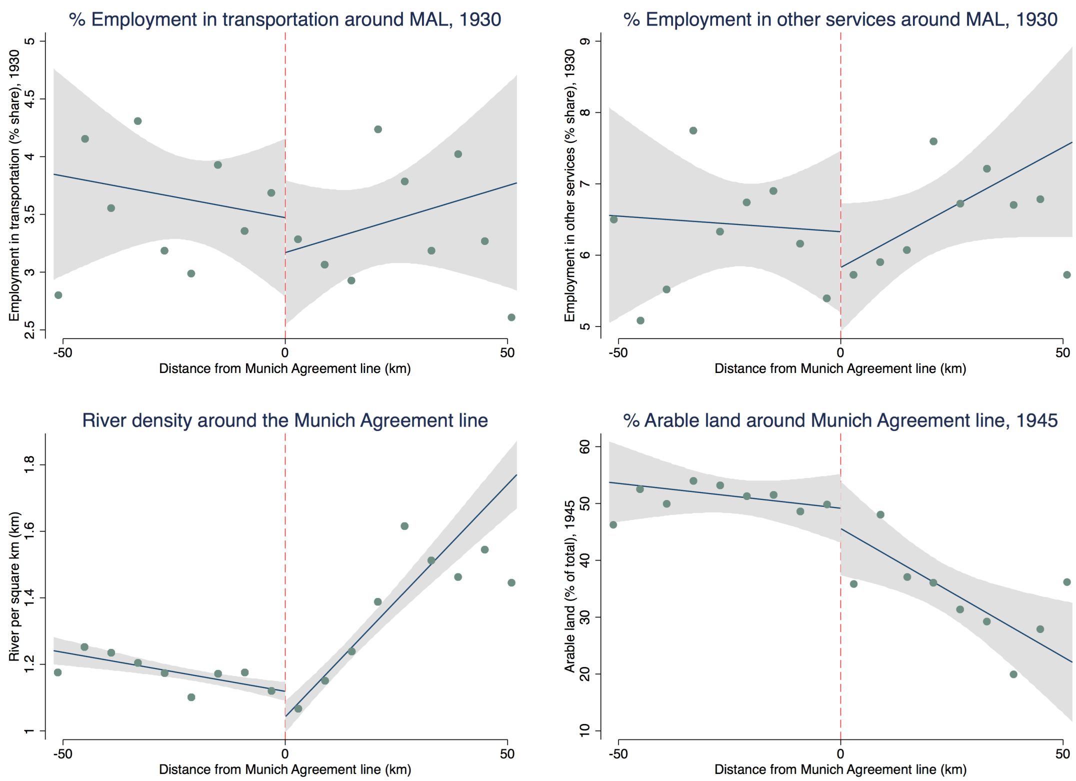
Trend lines are linear over the domain on each side of the Munich Agreement line. Data binned by 6 km intervals to produce 18 mean data points. All plots exclude Prague, Polish Zaolzie, and administrative units that overlap the Munich Agreement line. Shaded areas represent 95% confidence intervals.

Figure A.3: Balance tests, RD plots (IV)



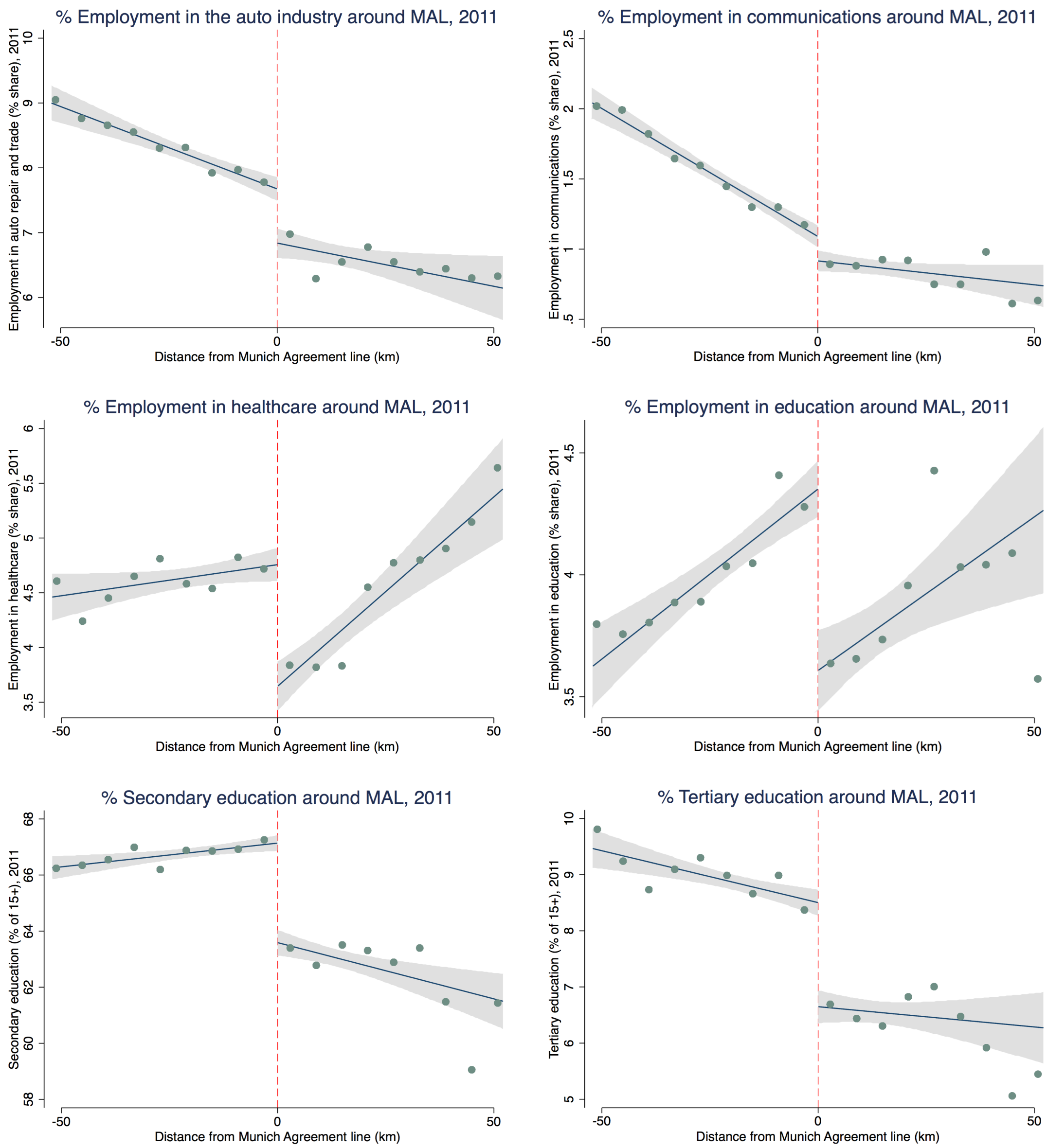
Trend lines are linear over the domain on each side of the Munich Agreement line. Data binned by 6 km intervals to produce 18 mean data points. All plots exclude Prague, Polish Zaolzie, and administrative units that overlap the Munich Agreement line. Shaded areas represent 95% confidence intervals.

Figure A.3: Balance tests, RD plots (V)



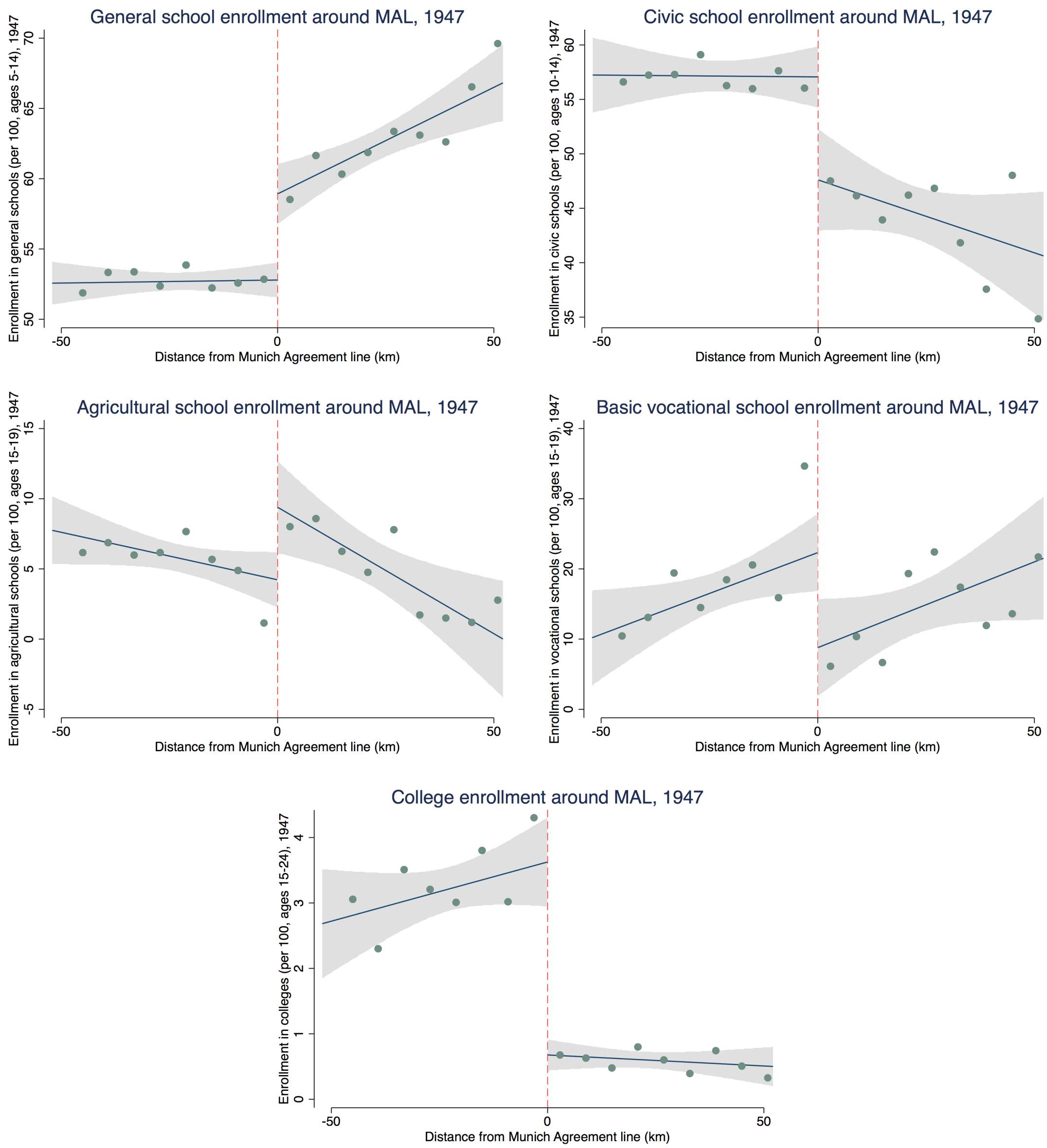
Trend lines are linear over the domain on each side of the Munich Agreement line. Data binned by 6 km intervals to produce 18 mean data points. All plots exclude Prague, Polish Zaolzie, and administrative units that overlap the Munich Agreement line. Shaded areas represent 95% confidence intervals.

Figure A.4: Other long-run effects, RD plots



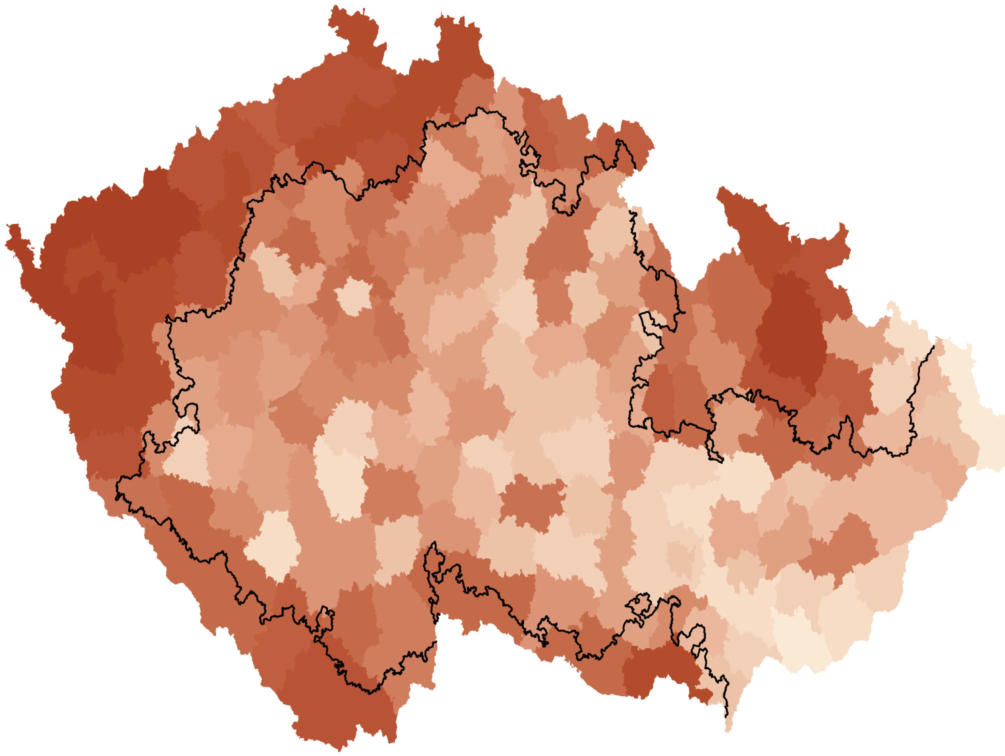
Trend lines are linear over the domain on each side of the Munich Agreement line. Data binned by 6 km intervals to produce 18 mean data points. All plots exclude Prague, Polish Zaolzie, and administrative units that overlap the Munich Agreement line. Shaded areas represent 95% confidence intervals.

Figure A.5: Other short-run effects, RD plots



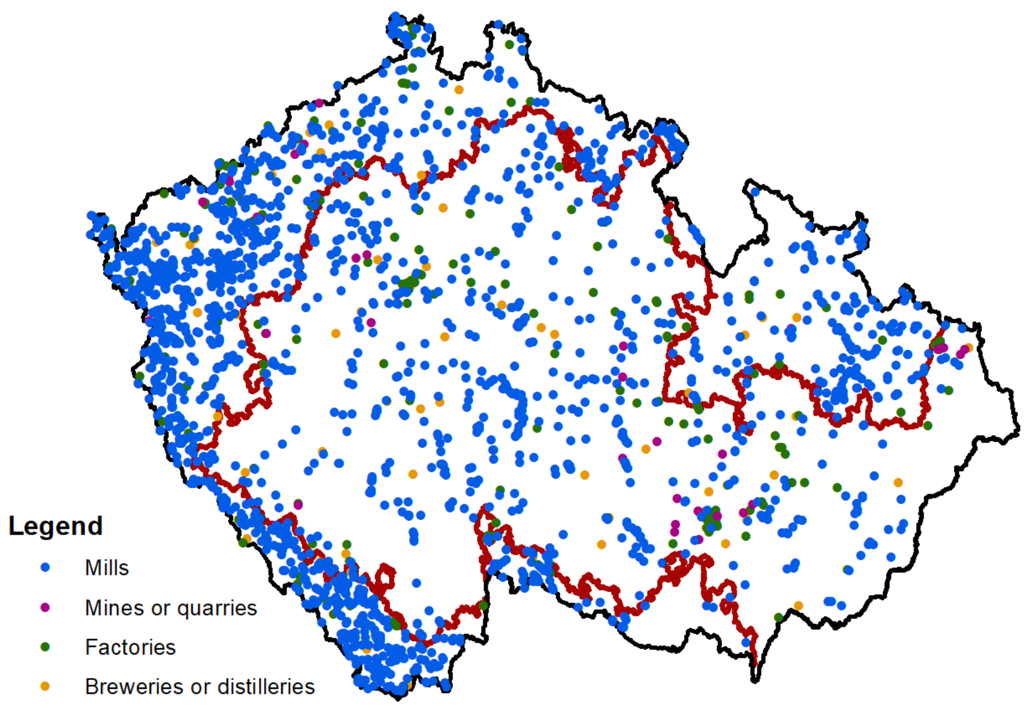
Trend lines are linear over the domain on each side of the Munich Agreement line. Data binned by 6 km intervals to produce 18 mean data points. All plots exclude Prague, Polish Zaolzie, and administrative units that overlap the Munich Agreement line. Shaded areas represent 95% confidence intervals.

Figure A.6: Migration out of the borderlands, 1950



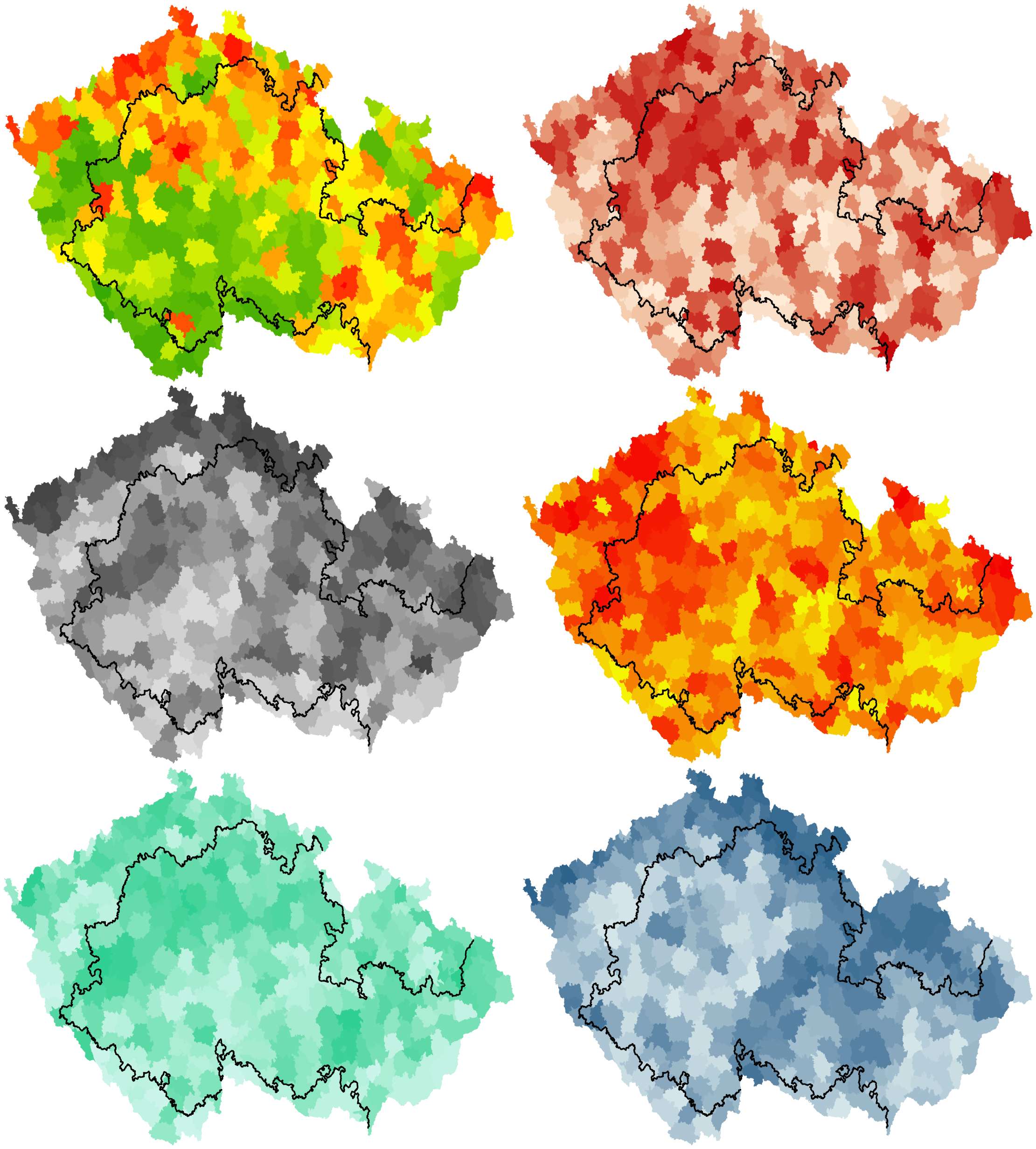
Outmigration heat map for 1950-districts (earliest outmigration data in the post-expulsion period) relative to the Munich Agreement line. Darker implies higher per capita outmigration.

Figure A.7: An "industrial graveyard"



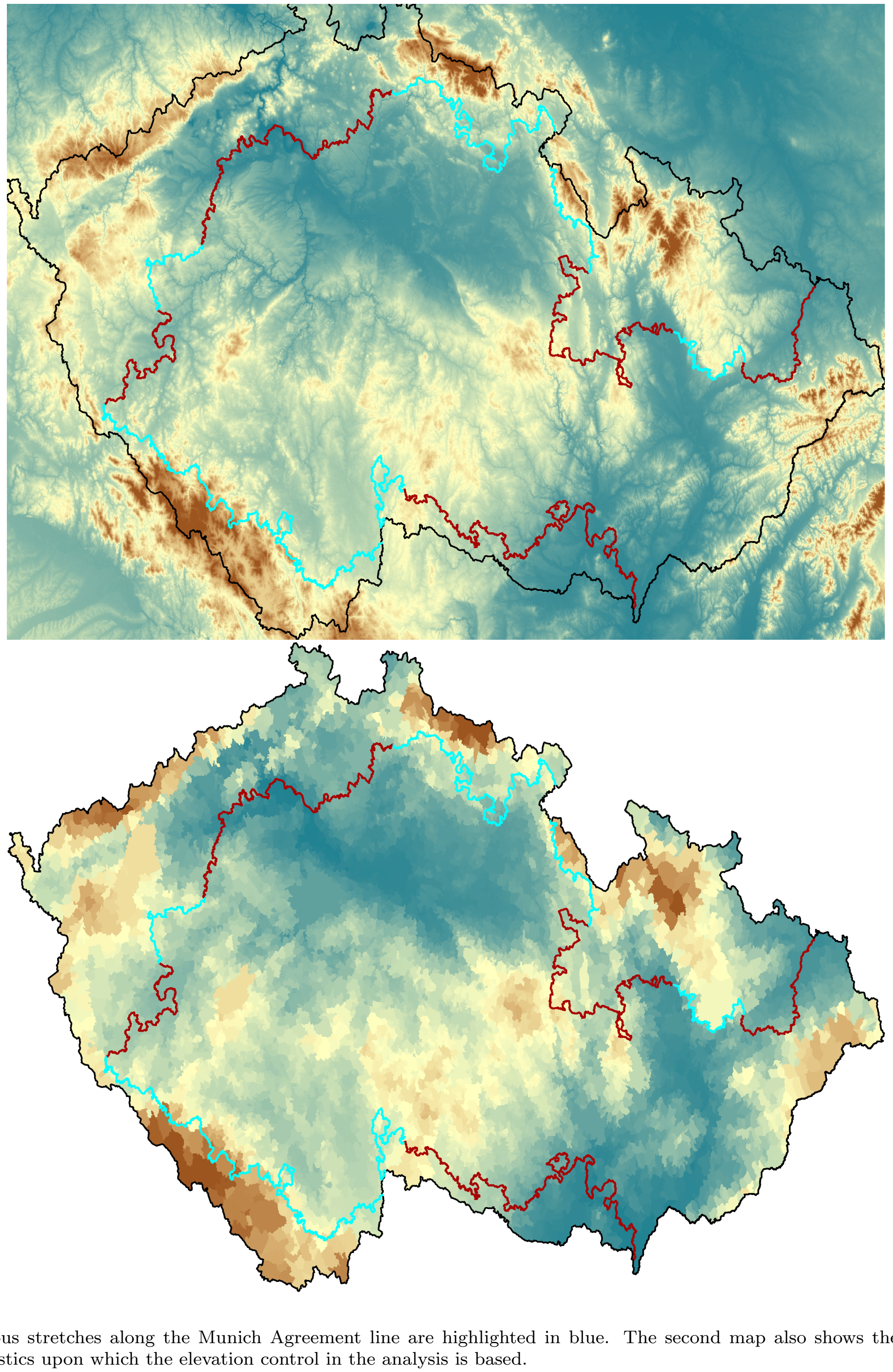
Sites destroyed or abandoned in the 20th century (source: zanikleobce.cz).

Figure A.8: Other localization patterns, 1930



Heatmaps of 1930 population density, share of employment in tranportation, mining and other extraction, textiles, general manufacturing, and overall industry (clockwise from top left). Darker shades indicate larger shares. Note that transportation and general manufacturing tend to be located wherever population is denser; mining, stone, and soil extraction are more common in Northwest Bohemia as well as Eastern Moravia, which are both mineral rich and not necessarily densely populated; and textile manufacturing is more common in Northern Moravia, in a mixture of densely and not-so-densely populated areas. None appear to be discontinuous through the eventual Munich Agreement line.

Figure A.9: Elevation (raster and zonal statistics)



Mountainous stretches along the Munich Agreement line are highlighted in blue. The second map also shows the elevation zonal statistics upon which the elevation control in the analysis is based.

Figure A.10: German language frontiers, post-1918

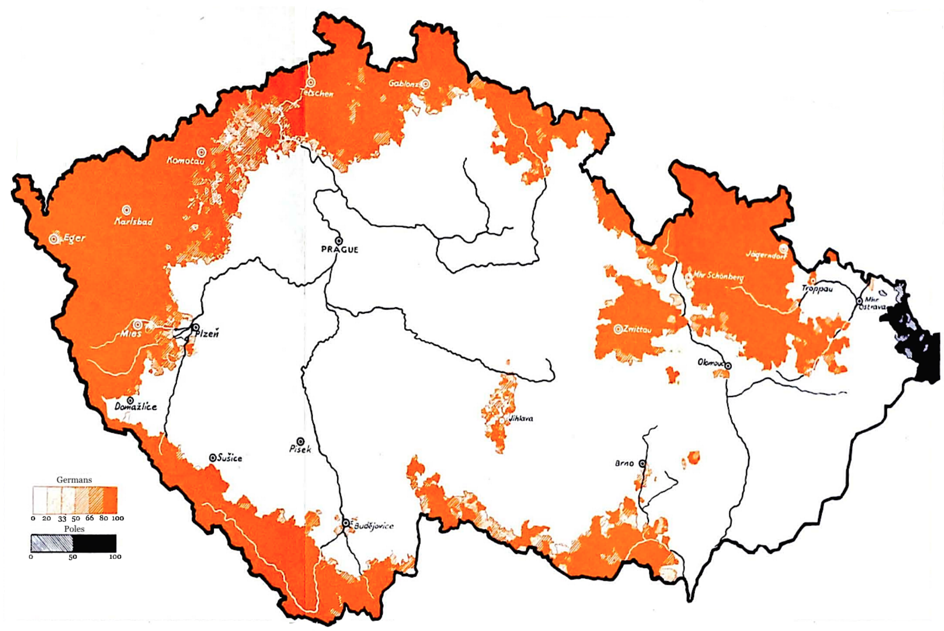
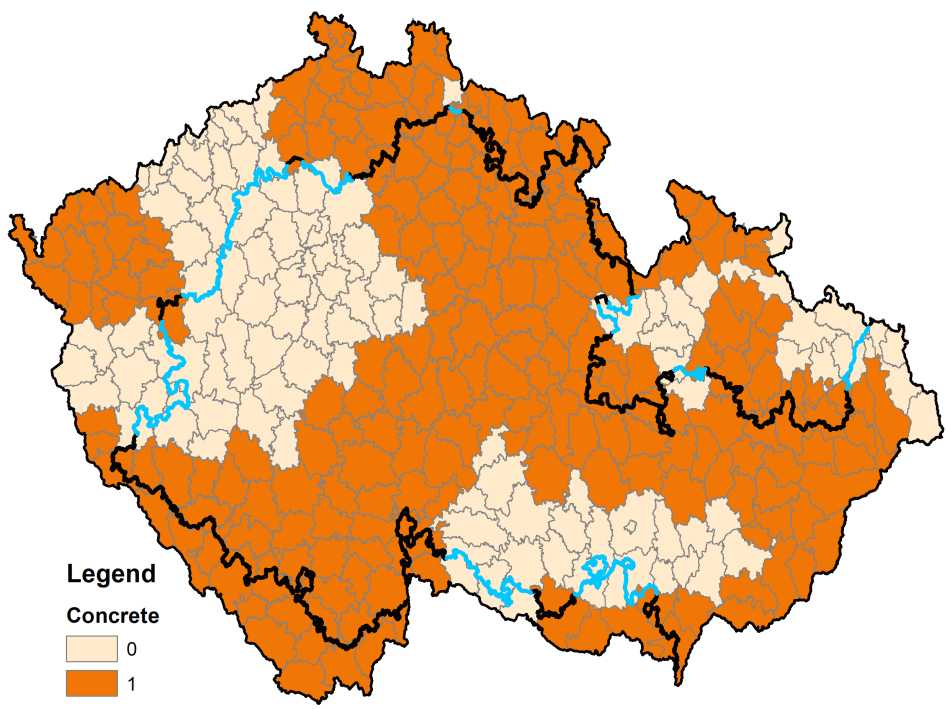


Figure A.11: "Concrete" sections of the language border, 1930



Non-concrete (i.e. mixed) stretches of the Munich Agreement line (MAL) are highlighted in blue. Based on their proximity to these stretches, beige districts are henceforth dropped for the analysis in Table A.9, which seeks to compare the parts of the sample for which the borderlands was more homogeneous (i.e. > 80% German) near the MAL, using the algorithm described below Table A.9. Compare to Figure A.10, which shows village-level ethnic composition prior to the expulsion, from Wiskemann (1938).

Variable descriptions

- **In borderlands**: if the majority of a district or municipality's area lies in the parts of the Czech lands (i.e. the modern day Czech Republic) that were annexed by Germany in 1938, then it is said to be in the borderlands (i.e. the Sudetenland), as opposed to the interior (i.e. the Protectorate of Bohemia and Moravia). The main specifications include any district or municipality in which > 95% of its area lie in either the borderlands or the interior. Additional specifications relax this if that district was nonetheless ethnically homogeneous in 1930. Calculated in ArcGIS.

- **Distance from Munich Agreement line**: a district or municipality's centroid's minimum distance to the Munich Agreement line. Calculated in ArcGIS.

- **% German**, 1930/50: the percentage of the total district population identified as being German on the census.

- **% Roma**, 1930: the percentage of the total judicial district population identified as being Romani on the 1930 census.

- **% Jewish (ethnic group)**, 1930: the percentage of the total judicial district population identified as being Jewish in nationality on the 1930 census.

- **% Jewish (religion)**, 1930: the percentage of the total judicial district population identified as being Jewish in religion on the 1930 census.

- **Ethnic fractionalization**, 1930: takes into account the share of the population that was German (g) or Czechoslovak (c) in judicial districts on the 1930 census. Other ethnic groups in the Czech lands were of trivial size. Hence, this measure is given by $1 - g^2 - c^2$.

- **Convictions per 100**, 1923-27: the number of convicted offenders in Czech criminal districts between 1923 and 1927 as a proportion of the total population in 1930. These data merge several judicial districts into larger jurisdictions in the Brno, Zlin, and Prague urban areas.

- **% Taxpayers**, 1933: the number of eligible taxpayers per 100 individuals in 1933 political districts, as reported in *Statistika daně důchodové placené přímo, daně z vyššího služného, daně rentové placené přímo, všeobecné a zvláštní daně výdělkové podle předpisu za rok 1933*, a Czechoslovak taxation report published by the State Bureau of Statistics in 1938. Not reported on its own for Praha-venkov (i.e. a suburban political district near Prague).

- **Income per capita (100 Kčs)**, 1933: average income per capita in 1933 political districts in 100 Czechoslovak koruna, as reported in the same taxation report. Not reported on its own for Praha-venkov, Ricany, and Jilove (i.e. all Prague suburban political districts).

- **ln Population density**, 1921-2011: the log of population counts per square kilometer in a district or municipality as reported in each census. Note that for 2011, three municipalities (Brezina, Brdy, and Modrava) designated for military purposes have low population counts and therefore have negative values, though removing these does not affect estimates.

- **Labor force**, 1921-2011: the total number of employed and unemployed. For 2011, the census reports the number of employed only. I use the number of unemployed from the same month as reported by the Czech Ministry of Labor and Social Affairs to derive the full labor force count.

- **ln Labor force density**, 1921-2011: the log of the above value divided by the square kilometer size of a district or municipality as reported in each census. Note the same three negative values here as in ln population density.

- **Labor force participation rate**, 1930/2011: the total labor force count divided by the total population of a district or municipality, as reported in the census.

- **Unemployment**, 1933/2011: the number of registered unemployed as a share of the total labor force. For 1933, the number of unemployed in a political district is taken from the social insurance report, *Nezaměstnanost a podpůrná péče v Československu*, written by Minister of Social Welfare Jaromír Nečas and published by the Social Institute of the Czechoslovak Republic in 1938. The 1930 labor force count from the census is used as the denominator. For 2011, the number of unemployed in a municipality for the month of March is scraped from the Czech Ministry of Labor and Social Affairs website. The 2011 labor force count described above is used as the denominator.

- **% Agricultural sector**, 1921-2011: the total number of workers in a district or municipality employed in agricultural work, fishing, hunting, or forestry as a share of the total labor force, as reported in the census.

- **% Secondary sector**, 1921-2011: the total number of workers in a district or municipality employed in the secondary sector (i.e. industry and construction) as a share of the total labor force, as reported in the census.

- **% Industry**, 1921-2011: the total number of workers in a district or municipality employed in the six industrial sectors below as a share of the total labor force, as reported in the census. Note that in the 1961 census, this was reported as a percentage instead of as the number of workers.

- **% Mining and other extraction**, 1930: the total number of workers in a district or municipality employed in mineral, stone, and soil extraction as a share of the total labor force, as reported in the census.

- **% Metallurgy and metalwork**, 1930: the total number of workers in a district or municipality employed in metallurgy and metalworking as a share of the total labor force, as reported in the census.

- **% Manufacturing**, 1930: the total number of workers in a district or municipality employed in manufacturing of machinery, equipment, and transportation devices as a share of the total labor force, as reported in the census.

- **% Glasswork**, 1930: the total number of workers in a district or municipality employed in the production of glass and glass products as a share of the total labor force, as reported in the census.

- **% Textiles**, 1930: the total number of workers in a district or municipality employed in textile manufacturing as a share of the total labor force, as reported in the census.

- **% Other industry**, 1930: the total number of workers in a district or municipality employed in other industrial sectors (i.e. chemical, gas, water, and electric industries; leather, clothing, and footwear manufacturing, lumber, paper, and printing industries; and food and beverage production) as a share of the total labor force, as reported in the census.

- **% Construction**, 1921-2011: the total number of workers in a district or municipality employed in construction as a share of the total labor force, as reported in the census.

- **% Service sector**, 1921-2011: the total number of workers in a district or municipality employed in the service sector (i.e. transport, business, and other service sectors below) as a share of the total labor force, as reported in the census.

- **% Transport sector**, 1921-2011: the total number of workers in a district or municipality employed in the transport sector (i.e. post, storage and shipping, rail, and bus) as a share of the total labor force, as reported in the census.

- **% Business sector**, 1921-2011: the total number of workers in a district or municipality employed in the business sector (i.e. finance and insurance as well as work in trade and commerce) as a share of the total labor force, as reported in the census.

- **% Finance and insurance**, 1930/2011: the total number of workers in a district or municipality employed in finance, accounting, and insurance as a share of the total labor force, as reported in the census.

- **% Trade**, 1930: the total number of workers in a district or municipality employed in trade and commerce (i.e. hospitality and food, auto trade and repair, and other commerce) as a share of the total labor force, as reported in the census.

- **% Hospitality and food services**, 2011: the total number of workers in a district or municipality employed in hotel and hospitality as well as food and catering services as a share of the total labor force, as reported in the census.

- **% Auto trade and repair**, 2011: the total number of workers in a district or municipality employed in auto retail trade and repair as a share of the total labor force, as reported in the census.

- **% Public**, 2011: the total number of workers in a district or municipality employed in public administration and defense as a share of the total labor force, as reported in the census.

- **% Communications**, 2011: the total number of workers in a district or municipality employed in communications and other information industries as a share of the total labor force, as reported in the census.

- **% Education**, 2011: the total number of workers in a district or municipality employed in education as a share of the total labor force, as reported in the census.

- **% Healthcare**, 2011: the total number of workers in a district or municipality employed in social and healthcare as a share of the total labor force, as reported in the census.

- **% Other service**, 1930/2011: the total number of workers in a district or municipality employed in all other service industries as a share of the total labor force, as reported in the census. For 1930, this includes public administrative and defense, education, healthcare, and domestic services. For 2011, this includes real estate, administrative and support fields, and scientific and technical activities.

- **% Literate**, 1921/30: the percentage of the population over the age of 10 that can read and write, as reported in the census.

- **% Primary education or less**, 1961-2011: the percentage of the population over the age of 15 that has at most primary education or less, as reported in the census.

- **% Secondary education**, 1961-2011: the percentage of the population over the age of 15 that has a secondary education (i.e. vocational, lower professional, or gymnasium, with or without exams) but no more, as reported in the census.

- **% Tertiary education**, 1961-2011: the percentage of the population over the age of 15 that has a tertiary education (i.e. higher professional education, some college, a bachelor degree, or more), as reported in the census.

- **Education index**, 1921-2011: uses prewar literacy and postwar post-primary education data transformed into standard deviations from census year district means.

- **General enrollment per 100, 5-14**, 1947: the number of individuals in a political district enrolled in general schools (i.e. schools which offer both primary schooling as well as terminal lower secondary education) as a share of the total population between the age of 5 and 14, as reported in the report, *Zprávy státního úřadu statistického republiky Československé*, published by the State Bureau of Statistics in 1948.

- **General schools per 100 pupils**, 1947: the number of general schools in a political district per 100 pupils that live there who are enrolled in a general school, as reported in the same statistical report.

- **General teachers per 100 pupils**, 1947: the number of general school teachers in a political district per 100 pupils that live there who are enrolled in a general school, as reported in the same statistical report.
- **Civic enrollment per 100, 10-14**, 1947: the number of individuals in a political district enrolled in civic schools (i.e. a form of lower secondary education that leads into higher forms) as a share of the total population between the age of 10 and 14, as reported in the same statistical report.
- **Civic schools per 100 pupils**, 1947: the number of civic schools in a political district per 100 pupils that live there who are enrolled in a civic school, as reported in the same statistical report.
- **Civic teachers per 100 pupils**, 1947: the number of civic school teachers in a political district per 100 pupils that live there who are enrolled in a civic school, as reported in the same statistical report.
- **Agricultural enrollment per 100, 15-19**, 1947: the number of individuals in a political district enrolled in agricultural folk schools (i.e. a common form of higher secondary education that focuses on agricultural and related skills) as a share of the total population between the age of 15 and 19, as reported in the same statistical report.
- **Agricultural folk schools per 100 pupils**, 1947: the number of agricultural folk schools in a political district per 100 pupils that live there who are enrolled in a agricultural folk school, as reported in the same statistical report.
- **Agricultural teachers per 100 pupils**, 1947: the number of agricultural folk school teachers in a political district per 100 pupils that live there who are enrolled in a agricultural folk school, as reported in the same statistical report.
- **Basic vocational enrollment per 100, 15-19**, 1947: the number of individuals in a political district enrolled in basic vocational schools (i.e. a common form of higher secondary education that focuses on more technical applied skills) as a share of the total population between the age of 15 and 19, as reported in the same statistical report.
- **Basic vocational schools per 100 pupils**, 1947: the number of basic vocational schools in a political district per 100 pupils that live there who are enrolled in a basic vocational school, as reported in the same statistical report.
- **Basic vocational teachers per 100 pupils**, 1947: the number of basic vocational school teachers in a political district per 100 pupils that live there who are enrolled in a basic vocational school, as reported in the same statistical report.
- **College enrollment per 100, 15-24**, 1947: the number of individuals in a political district enrolled in colleges as a share of the total population between the age of 15 and 19, as reported in the same statistical report.

- **In migrants per capita**, 1950-2011: the number of individuals who are immigrants in a given year into a district, divided by the total population size of that district in that year, as reported in official annual population journals, *Pohyb obyvatelstva v republice Československé*, available online from the Czech Statistical Office.

- **Out migrants per capita**, 1950-2011: the number of individuals who are emigrants in a given year from a district, divided by the total population size of that district in that year, as reported in the same statistical journal.

- **Net migrants per capita**, 1950-2011: the net of the previous two measures.

- **Border segments**: a variable whose value corresponds to the "segment" of the Munich Agreement line to which a district or municipality is closest. This is derived as follows: for 2001/11 municipalities, each of the three unique continuous stretches of the Munich Agreement line – in Bohemia, Northern Moravia, and Southern Moravia – is divided into 25, 13, and 12 segments, respectively, in ArcGIS, each about 50 km in length. For 1930/47 judicial districts, each is divided into 12, 6, and 6 segments, respectively, in ArcGIS, each about 100 km in length. For 1930/47 political districts, each is divided into 8, 4, and 4 segments, respectively, in ArcGIS, each about 150 km in length. For some small sample sensitivity tests, I also generate a version with just 4 border segments, corresponding to the three regions with Bohemia divided into Northern and Southern sections.

- **Prague**: a dummy that equals 1 if a district or municipality corresponds to the city of Prague.

- **Polish Zaolzie**: a dummy that equals 1 if a district or municipality lies in the strip of land in the easternmost part of the Czech lands, which was annexed by Poland in 1938 and has historically been predominantly Polish-speaking.

- **Eastern Bloc**: a dummy that equals 1 if a municipality lies closer to Poland or the former East Germany than to Austria or the former West Germany, as calculated in ArcGIS.

- **Longitude and latitude**: measures of longitude and latitude for district and municipality centroids, calculated in ArcGIS using a WGS 1984 projection, each normalized around the sample mean.

- **Elevation** (m): 1 arc sec elevation data are derived from Japan Aerospace Exploration Agency's Advanced Land Observing Satellite (2017) maps, with district- and municipality-specific mean values estimated in ArcGIS using zonal statistics.

- **Ruggedness** (°): 1 arc sec ruggedness data are derived from Japan Aerospace Exploration Agency's Advanced Land Observing Satellite (2017) maps, with district- and municipality-specific mean values estimated in ArcGIS using zonal statistics.

- **Temperature** (°C): 30 arc sec temperature data (1970-2000) are derived from Worldclim (2016) maps, with district- and municipality-specific mean values estimated in ArcGIS using zonal statistics.

- **Precipitation** (mm): 30 arc sec precipitation data (1970-2000) are derived from Worldclim (2016) maps, with district- and municipality-specific mean values estimated in ArcGIS using zonal statistics.

- **River density** (rivers (km)/km^2): detailed GIS shapefiles of river networks provided by Geofabrik (2017) are converted to a equidistant cylindrical projection in ArcGIS. The "intercept" tool is used to determine in what districts and municipalities a given river segment lies. I then sum the total length for all river segments within each municipality. Using the district or municipality area calculated in ArcGIS from files with a cylindrical equal area projection, I then calculate river density values.

- **% Arable land**, 1945: the number of square kilometers of arable land in 1945 political districts divided by the total number of square kilometers, as reported in the report, *Zprávy státního úřadu statistického republiky Československé*, published by the State Bureau of Statistics in 1947.

Appendix B

Appendix for Chapter 2

Proofs

Proof of Proposition 1

Proof. Suppose

$$V_r(h) = V_r(u) \Leftrightarrow a_r \frac{K}{\lambda M_r} h^{1+\gamma} m_r^{\gamma} = \frac{K}{\beta M_r}, \tag{B.1}$$

where m_r gives the number of productive workers in region r (relative to unproductive workers in r as well as workers in region $-r$). Hence, any arbitrarily small positive perturbation to m_r will induce

$$a_r \frac{K}{\lambda M_r} h^{1+\gamma} m_r^{\gamma} > \frac{K}{\beta M_r}, \tag{B.2}$$

such that the remainder of workers shift to production iteratively, until the total share of workers in r are all engaged in production and $m_r = M_r \equiv m_r^*$. I define the m_r that solves (B.1) to be:

$$\widehat{m_r} \equiv \left(\frac{\lambda}{\beta a_r h^{1+\gamma}} \right)^{\frac{1}{\gamma}},$$

above which all workers in region r prefer to specialize in production over unproductive activities.

Because $M_r \geq m_r$ by definition, it follows that $M_r \geq \widehat{m_r}$ must hold for this equilibrium to be feasible. If instead $\widehat{m_r} > M_r$, then $\widehat{m_r} > m_r$ always. □

Proof of Remark 1

Proof. In a HPSE, $m_r^* = M_r$, and

$$a_r \frac{K}{\lambda M_r} h^{1+\gamma} M_r^{\gamma} \geq \frac{K}{\beta M_r}.$$

Suppose $\gamma = 0$. Then this condition becomes $\beta a_r h \geq \lambda$, which is invariant to M_r. Hence, when $\gamma = 0$, if a HPSE exists in region r for some population share M_r, then it also exists for all M_r'. □

Proof of Proposition 2

Proof. Let $\gamma \in (0,1)$. Suppose $M_r > \widehat{m_r}$ and $m_r^* = M_r$ in each region r, such that there is a locally stable HPSE in each region. Then

$$\begin{aligned}\dot{M}_r &= M_r(V_r(h) - \overline{V}(h)) = M_r(1-M_r)(V_r(h) - V_{-r}(h)) \\ &= M_r(1-M_r)(a_r \frac{K}{\lambda} h^{1+\gamma} M_r^{\gamma-1} - a_{-r}\frac{K}{\lambda} h^{1+\gamma} M_{-r}^{\gamma-1}), \quad \text{(B.3)}\end{aligned}$$

where $M_{-r} = 1 - M_r$. By inspection, $V_r(h)$ is strictly decreasing in M_r for all M_r, with $\lim_{M_r\to 0} V_r(h) = \infty$ and $\lim_{M_r \to 1} V_r(h) = a_r \frac{K}{\lambda} h^{1+\gamma}$, while $V_{-r}(h)$ is strictly increasing in M_r for all M_r, with $\lim_{M_r\to 0} V_{-r}(h) = a_{-r}\frac{K}{\lambda}h^{1+\gamma}$ and $\lim_{M_r\to 1} V_{-r}(h) = \infty$. Hence, $\dot{M}_r = 0$ if $V_r(h) = V_{-r}(h)$, where $V_r(h) = V_{-r}(h)$ if and only if

$$M_r = \frac{a_r^{\frac{1}{1-\gamma}}}{a_r^{\frac{1}{1-\gamma}} + a_{-r}^{\frac{1}{1-\gamma}}} \equiv M_r^* \in (0,1)$$

for regions r and $-r \neq r$, and $\frac{\partial \dot{M}_r}{\partial M_r}|_{M_r = M_r^*} < 0$. Hence, if $m_r^* = M_r$ and $M_r^* > \widehat{m_r}$ in each region r, then there is a locally stable HPLE, with a unique interior steady state population $M_r^* \in (0,1)$ when $\gamma \in (0,1)$.

(ii) Suppose there is a sufficiently large negative population shock in one region, say 2 (without loss of generality), such that (B.2) no longer holds and $m_2^* = 0$. However, suppose $M_1 > \widehat{m_1}$ and $m_1^* = M_1$ still. Then

$$\dot{M}_1 = M_1(1-M_1)(V_1(h) - V_2(u)) = M_1(1-M_1)\left(a_1 \frac{K}{\lambda} h^{1+\gamma} M_1^{\gamma-1} - \frac{K}{\beta(1-M_1)}\right). \quad \text{(B.4)}$$

By inspection, $V_1(h)$ is strictly decreasing in M_1 for all M_1, with $\lim_{M_1\to 0} V_1(h) = \infty$ and $\lim_{M_1\to 1} V_1(h) = a_1 \frac{K}{\lambda} h^{1+\gamma}$, while $V_2(u)$ is strictly increasing in M_2 for all M_2, with $\lim_{M_1 \to 0} V_2(u) = K/\beta$ and $\lim_{M_1\to 1} V_2(u) = \infty$. Hence, $\dot{M}_1 = 0$ if $V_1(h) = V_2(u)$; $V_1(h) = V_2(u)$ for a unique $M_1 \equiv M_1^{**} \in (0,1)$; and $\frac{\partial \dot{M}_r}{\partial M_1}|_{M_r = M_1^{**}} < 0$. Hence, if $m_1^* = M_1$, $m_2^* = 0$, and $M_1^{**} > \widehat{m_1}$, then there is a locally stable ALE, with a unique interior steady state $M_1^{**} \in (0,1)$ when $\gamma \in (0,1)$.

Next, it is straightforward to show that the steady state population share in a productive (unproductive) region will be greater (lower) in a ALE than in a HPLE: $M_1^{**} > M_1^*$ and $M_2^{**} < M_2^*$.

Suppose to the contrary that $M_1^{**} \leq M_1^*$. When $M_1 = M_1^*$,

$$a_1 \frac{K}{\lambda} h^{1+\gamma} M_1^{*\gamma-1} = a_2 \frac{K}{\lambda} h^{1+\gamma} (1 - M_1^*)^{\gamma-1}. \quad \text{(B.5)}$$

If $M_1^{**} \leq M_1^*$, then when $M_1 = M_1^*$ holding all else fixed,

$$a_1 \frac{K}{\lambda} h^{1+\gamma} M_1^{*\gamma-1} \leq \frac{K}{\beta(1-M_1^*)} \Leftrightarrow \dot{M}_1 < 0. \tag{B.6}$$

Together (B.5) and (B.6) imply that when $M_1 = M_1^*$, if $M_1^{**} \leq M_1^*$ then

$$a_2 \frac{K}{\lambda} h^{1+\gamma} (1-M_1^*)^{\gamma-1} \leq \frac{K}{\beta(1-M_1^*)} \Leftrightarrow M_2^* \leq \widehat{m_2}.$$

However, by Definition 2.2.3, $M_r^* > \widehat{m_r}$ for all r in any locally stable HPLE. Hence, $M_1^{**} > M_1^*$ by contradiction, and $M_2^{**} = 1 - M_1^{**} < 1 - M_1^* = M_2^*$ by symmetry.

To check comparative statics for M_1^{**}, implicitly differentiate $V_1(h) - V_2(u) = 0$ with respect to each variable of interest. Doing so yields:

$$a_1 : \frac{1}{\lambda} h^{1+\gamma} M_1^{**\gamma-1} + (\gamma-1)\frac{\partial M_1^{**}}{\partial a_1} a_1 \frac{1}{\lambda} h^{1+\gamma} M_1^{**\gamma-2} = \frac{\partial M_1^{**}}{\partial a_1} \frac{1}{\beta(1-M_1^{**})^2},$$

$$h : (1+\gamma) a_1 \frac{1}{\lambda} h^{\gamma} M_1^{**\gamma-1} + (\gamma-1)\frac{\partial M_1^{**}}{\partial h} a_1 \frac{1}{\lambda} h^{1+\gamma} M_1^{**\gamma-2} = \frac{\partial M_1^{**}}{\partial h} \frac{1}{\beta(1-M_1^{**})^2},$$

$$\lambda : -a_1 \frac{1}{\lambda^2} h^{1+\gamma} M_1^{**\gamma-1} + (\gamma-1)\frac{\partial M_1^{**}}{\partial \lambda} a_1 \frac{1}{\lambda} h^{1+\gamma} M_1^{**\gamma-2} = \frac{\partial M_1^{**}}{\partial \lambda} \frac{1}{\beta(1-M_1^{**})^2},$$

$$\beta : (\gamma-1)\frac{\partial M_1^{**}}{\partial \beta} a_1 \frac{1}{\lambda} h^{1+\gamma} M_1^{**\gamma-2} = -\frac{1}{\beta^2(1-M_1^{**})} + \frac{\partial M_1^{**}}{\partial \beta} \frac{1}{\beta(1-M_1^{**})^2},$$

which imply $\frac{\partial M_1^{**}}{\partial a_1}, \frac{\partial M_1^{**}}{\partial h}, \frac{\partial M_1^{**}}{\partial \beta} > 0$ and $\frac{\partial M_1^{**}}{\partial \lambda} < 0$ when $\gamma \in (0,1)$. □

Proof of Remark 2

Proof. By Remark 1, population shocks cannot shift a region from one short-run equilibrium to another when $\gamma = 0$. Hence, if the steady state population distribution M_r^* is determined by (B.3) or (B.4) prior to a population shock, then all else fixed, it will also be determined by (B.3) or (B.4) after a shock, respectively, when $\gamma = 0$. □

Proof of Proposition 3

Proof. (i) Suppose $\gamma > 1$, $M_r > \widehat{m_r}$, and $m_r^* = M_r$ for all regions r. By inspection, $V_r(h) = a_r \frac{K}{\lambda} h^{1+\gamma} M_r^{\gamma-1}$ is now strictly increasing in M_r for all M_r, with $\lim_{M_r \to 0} V_r(h) = 0$ and $\lim_{M_r \to 1} V_r(h) = a_r \frac{K}{\lambda} h^{1+\gamma}$, while $V_{-r}(h) = a_{-r} \frac{K}{\lambda} h^{1+\gamma} (1-M_r)^{\gamma-1}$ is strictly decreasing in M_r for all M_r, with $\lim_{M_r \to 0} V_{-r}(h) = a_{-r} \frac{K}{\lambda} h^{1+\gamma}$ and $\lim_{M_r \to 1} V_{-r}(h) = 0$. Assuming positive parameters, $\dot{M}_r = 0$ at only one interior M_r^*, when $V_r(h) = V_{-r}(h)$. However, $\frac{\partial \dot{M}_r}{\partial M_r}|_{M_r = M_r^*} > 0$. Hence, there are no locally stable HPLE.

Note that $\lim_{M_r \to 0} \dot{M}_r = \lim_{M_r \to 1} \dot{M}_r = 0$; $\frac{\partial \dot{M}_r}{\partial M_r}|_{M_r \to 0} = V_r(h)|_{M_r \to 0} - V_{-r}(h)|_{M_r \to 0} < 0$; and $\frac{\partial \dot{M}_r}{\partial M_r}|_{M_r \to 1} = -V_r(h)|_{M_r \to 1} + V_{-r}(h)|_{M_r \to 1} < 0$, which both imply local stability. However, these "black hole" outcomes each violate the necessary condition for a HPSE for one of the regions $-r$ that $M_{-r} \geq \widehat{m_{-r}}$, such that $m_{-r}^* = 0$ for that region and $V_{-r} \neq V_{-r}(h)$. Hence, these cannot be long-run equilibria either.

(ii) Suppose instead that region 1 (without loss of generality) specializes in production while region 2 specializes in unproductive activities: $M_1 > \widehat{m_1}$, $m_1^* = M_1$, and $m_2^* = 0$. Then

$$\dot{M}_1 = M_1(1 - M_1)(V_1(h) - V_2(u)) = M_1(1 - M_1)\left(a_1\frac{K}{\lambda}h^{1+\gamma}M_1^{\gamma-1} - \frac{K}{\beta(1-M_1)}\right).$$

By inspection, $V_1(h)$ is strictly increasing in M_1 for all M_1, with $\lim_{M_1\to 0} V_1(h) = 0$ and $\lim_{M_1\to 1} V_1(h) = a_1\frac{K}{\lambda}h^{1+\gamma}$, while $V_2(u)$ is strictly increasing in M_1 for all M_1, with $\lim_{M_1\to 0} V_2(u) = K/\beta$ and $\lim_{M_1\to 1} V_2(u) = \infty$. Hence, assuming all positive parameters,

$$\lim_{M_1\to 0} V_2(u) > \lim_{M_1\to 0} V_1(h) \text{ and } \lim_{M_1\to 1} V_2(u) > \lim_{M_1\to 1} V_1(h).$$

This implies that $\frac{\partial \dot{M}_1}{\partial M_1}|_{M_1\to 0} = V_1(h)|_{M_1\to 0} - V_2(u)|_{M_1\to 0} < 0$ and $\frac{\partial \dot{M}_1}{\partial M_1}|_{M_1\to 1} = -V_1(h)|_{M_1\to 1} + V_2(u)|_{M_1\to 1} > 0$. Moreover, $\lim_{M_1\to 0}\dot{M}_1 = 0$ and $\lim_{M_1\to 1}\dot{M}_1 = -K/\beta$. Hence, black hole outcomes remain unstable.

However, it can be shown that there are at most two interior ALE, with steady state populations in the productive region, 1, of $\underline{M}_1^*$ and $\overline{M}_1^*$. First, $\dot{M}_1 = 0$ if and only if

$$a_1\frac{K}{\lambda}h^{1+\gamma}M_1^{\gamma-1} - \frac{K}{\beta(1-M_1)} = 0 \Leftrightarrow M_1^{\gamma-1}(1-M_1) - \frac{\lambda}{\beta a_1 h^{1+\gamma}} = 0, \quad \text{(B.7)}$$

which is never satisfied as $M_1 \to 0$ or $M_1 \to 1$.

Suppose $M_1^{\gamma-1}(1-M_1) > \frac{\lambda}{\beta a_1 h^{1+\gamma}}$ for some M_1. Then there are two M_1 that solve (B.7). To see this, note that the right equation is strictly increasing from the origin if and only if

$$(\gamma-1)M_1^{\gamma-2}(1-M_1) > M_1^{\gamma-1} \Leftrightarrow M_1 < \frac{\gamma-1}{\gamma}.$$

Otherwise, it is strictly decreasing until $M_1 = 1$. Hence, if $M_1^{\gamma-1}(1-M_1) > \frac{\lambda}{\beta a_1 h^{1+\gamma}}$ when $M_1 = \frac{\gamma-1}{\gamma}$, there is a lower root $\underline{M_1}^* \in (0, \frac{\gamma-1}{\gamma})$ that solves (B.7) as well as an upper root $\overline{M_1}^* \in (\frac{\gamma-1}{\gamma}, 1)$.

(iii) Recall that $M_1^{\gamma-1}(1-M_1) - \frac{\lambda}{\beta a_1 h^{1+\gamma}}$ is increasing through its lower root $\underline{M_1}^*$, and decreasing through its upper root, $\overline{M_1}^*$. It is straightforward to show that if $M_1^{\gamma-1}(1 - M_1) - \frac{\lambda}{\beta a_1 h^{1+\gamma}}$ is increasing (decreasing) through one of its roots, $a_1\frac{K}{\lambda}h^{1+\gamma}M_1^{\gamma-1} - \frac{K}{\beta(1-M_1)}$ and in turn $\dot{M}_1$ must be increasing (decreasing) through that same root.

To show this, I evaluate the more general claim that if $f(x) - q(x)$ is increasing (decreasing) in x at the x^* that solves $f(x^*) = q(x^*)$, then $g(x)[f(x) - q(x)]$ must also be increasing (decreasing) in x at x^* for all continuously differentiable $f(x)$, $q(x)$, and $g(x) > 0$. The derivative of $f(x) - q(x)$ with respect to x at $x = x^*$ is

$$[f'(x) - q'(x)]|_{x=x^*}. \quad \text{(B.8)}$$

The derivative of $g(x)[f(x) - q(x)]$ with respect to x at $x = x^*$ is

$$(g'(x)[f(x) - q(x)] + g(x)[f'(x) - q'(x)])|_{x=x^*} = g(x)[f'(x) - q'(x)]|_{x=x^*},$$

which has the same sign as (B.8) for all $g(x) > 0$. Let $x = M_1$, $f(\cdot) = a_1 \frac{K}{\lambda} h^{1+\gamma} M_1^{\gamma-1}$, $q(\cdot) = \frac{K}{\beta(1-M_1)}$ and $g(\cdot) = \frac{\lambda(1-M_1)}{a_1 h^{1+\gamma} K}$. Hence, $\frac{\partial \dot{M}_r}{\partial M_1}|_{M_1 = \underline{M_1}^*} > 0$ and $\frac{\partial \dot{M}_r}{\partial M_1}|_{M_1 = \overline{M_1}^*} < 0$. □

Appendix C

Appendix for Chapter 3

Proofs

Proof of Lemma 1

Proof. Suppose an educated ($e = 1$), politically active ($c = 1$) citizen observes $\tau_\theta = \tau_0$, revealing $\theta = 0$. Her expected payoff from challenging the ruler ($r = 1$) is $q\sigma(w + \alpha h - \alpha)$. Otherwise, it is $w + \alpha h - \alpha - \tau_0$. The latter exceeds the former if and only if $\tau_0 \leq (1 - q\sigma)(w + \alpha h - \alpha)$.

A predatory ruler's expected payoff when $e = 1$, $c = 1$, and $r = 1$ is $R^I + (1 - q)\sigma(w + \alpha h - \alpha) - 1$. A predatory ruler's expected payoff when $e = 1$ and $c = 1$ if he sets the highest possible τ_0 at which $r^* = 0$ is $R^I + (1 - q\sigma)(w + \alpha h - \alpha) - 1$. The latter exceeds the former for all $\sigma < 1$. Hence, a predatory ruler always prefers $\tau_0^* = (1-q\sigma)(w+\alpha h-\alpha) \equiv \tau_0^{PA}$ when the citizen is educated ($e = 1$) and politically active ($c = 1$), where $\tau_0^{PA} < w + \alpha h \equiv \tau_0^{PI}$, such that the citizen never mounts a political challenge ($r^* = 0$) thereafter. □

Proof of Proposition 1

Proof. (i) *Political club participation ($c^* = 1$)*: Fix education policy at neutral ($m_0 = m_1 = 0$). Suppose the citizen observes public education provision ($g_\theta = 1$) and enrolls ($e = 1$). Then $\hat{\pi}(\theta|\hat{\theta}) = 1$. Suppose $\theta = 1$. Her payoff from $c = 1$ is $w + \alpha h - \alpha - \tau_1^{BE}$. Her payoff from $c = 0$ is $w + \alpha h - \tau_1^{BE}$. Hence, $c^* = 0$ whenever $\theta = 1$. Now suppose $\theta = 0$. Her payoff from $c = 1$ is $q\sigma(w + \alpha h - \alpha)$. Her payoff from $c = 0$ is 0. Hence, $c^* = 1$ whenever $\theta = 0$. Hence, when $g_0 = g_1 = e = 1$ and $m_0 = m_1 = 0$, $c^* = 1$ if and only if $\theta = 0$.

Enrollment ($e^* = 1$), Case 1: suppose the citizen knows that $g_0 = 0$ and $g_1 = 1$, such that $g_\theta = 1$ implies ruler type. Then $c^* = 0$ whenever $e = 1$, such that her payoff from $e = 1$ after observing $g_\theta = 1$ is $w + \alpha h - \max\{1 - R^I, 0\}$. Her reservation payoff is w, where $e^* = 1$ if and only if the former exceeds the latter.

Case 2: suppose that the citizen knows that $g_0 = g_1 = 1$. Then $c^* = 1$ if and only if $\theta = 0$, such that her expected payoff from $e = 1$ is

$$Pr(\hat{\theta} = 0)\hat{\pi}(0|0)q\sigma(w + \alpha h - \alpha) + Pr(\hat{\theta} = 1)\hat{\pi}(1|1)(w + \alpha h - \max\{1 - R^I, 0\}),$$

where $e^* = 1$ if and only if this is greater than πw.

Public education provision ($g_0^* = 1$): By Lemma 1, a predatory ruler's expected payoff from $g_0 = 1$ when $e^* = 1$ and $c^* = 1$ if and only if $\theta = 0$ is $R^I + (1 - q\sigma)(w + \alpha h - \alpha) - 1$. His reservation payoff is $R^I + w$. Suppose $g_0 = 1$.

$$R^I + (1 - q\sigma)(w + \alpha h - \alpha) - 1 \geq R^I + w \Leftrightarrow h \geq \frac{1 + \alpha(1 - q\sigma) + q\sigma w}{\alpha(1 - q\sigma)} \equiv \overline{h}.$$

Hence, there exists a value of $h \equiv \overline{h}$ above which $g_0^* = 1$ under these conditions.

A benevolent ruler's expected payoff from $g_1 = 1$ when $e^* = 1$ and $c^* = 1$ if and only if $\theta = 0$ is $w + \alpha h - \alpha \mathbb{E}(c(m_\theta|\hat{\theta})|\theta = 1) - \max\{1 - R^I, 0\}$. His reservation payoff is w. Suppose $g_1 = 1$. Then

$$w + \alpha h - \alpha \mathbb{E}(c(m_\theta|\hat{\theta})|\theta = 1) - \max\{1 - R^I, 0\} \geq w$$
$$\Leftrightarrow h \geq \frac{\max\{1 - R^I, 0\} + \alpha \mathbb{E}(c(m_\theta|\hat{\theta})|\theta = 1)}{\alpha} \equiv \underline{h}.$$

Hence, there exists a value of $h \equiv \underline{h}$ above which $g_1^* = 1$ under these conditions.

By inspection, $\overline{h} > \frac{1+\alpha}{\alpha} \geq \underline{h}$ for all $R^I \geq 0$. Hence, when $m_0 = m_1 = 0$, $g_1^* = 1$ is necessary for $g_0^* = 1$ when $e^* = 1$ and $c^* = 1$ if and only if $\theta = 0$.

Confirming enrollment ($e^* = 1$): By inspection, $e^* = 1$ if $h \geq \underline{h}$. Hence, $h \geq \overline{h} > \underline{h}$ is necessary and sufficient for a developmental schooling equilibrium [DSE] when $m_0 = m_1 = 0$.

(ii) $\frac{\partial \overline{h}}{\partial q} = \frac{\sigma(1+w)}{\alpha(1-q\sigma)^2} > 0$, $\frac{\partial \overline{h}}{\partial \sigma} = \frac{q(1+w)}{\alpha(1-q\sigma)^2} > 0$, $\frac{\partial \overline{h}}{\partial w} = \frac{q\sigma}{\alpha(1-q\sigma)} > 0$, and $\frac{\partial \overline{h}}{\partial \alpha} = -\frac{1+q\sigma w}{\alpha^2(1-q\sigma)} < 0$. The derivative of $\underline{h}$ with respect to α is $-(1 - R^I)\alpha^{-2} < 0$ if $1 > R^I$ and 0 otherwise, while the derivative of $\underline{h}$ with respect to R^I is $-\alpha^{-1} < 0$ if $1 > R^I$ and 0 otherwise. □

Proof of Proposition 2

Proof. (i) *Political club participation ($c^* = 1$)*: Now let education policy m_θ be anything. Suppose the citizen observes $g_\theta = 1$ and chooses $e = 1$.

Suppose the citizen then observes $\hat{\theta} = 0$, such that her expected payoff from $c = 1$ is

$$\hat{\pi}(0|0)q\sigma(w + \alpha h - \alpha) + \hat{\pi}(1|0)(w + \alpha h - \alpha - \max\{1 - R^I, 0\}).$$

Her payoff from $c = 0$ is $\hat{\pi}(1|0)(w + \alpha h - \max\{1 - R^I, 0\})$. Fix $m_1 = 0$, such that $\hat{\pi}(1|0) = 0$. Hence, $c^* = 1$ for all m_0 when $\hat{\theta} = 0$ and $m_1 = 0$.

Now suppose the citizen observes $\hat{\theta} = 1$. A citizen's payoff from $c = 1$ is

$$\hat{\pi}(0|1)q\sigma(w + \alpha h - \alpha) + \hat{\pi}(1|1)(w + \alpha h - \alpha - \max\{1 - R^I, 0\}).$$

Her payoff from $c = 0$ is $\hat{\pi}(1|1)(w + \alpha h - \max\{1 - R^I, 0\})$. Again fix $m_1 = 0$, but now suppose $m_0 = 1$. There exists a threshold of $Pr(\hat{\theta} = 1|\theta = 0, m_0 = 1) = \gamma \equiv \overline{\gamma}$, where $c^* = 0$

when both $\hat{\theta} = 1$ and $m_1 = 0$ if and only if

$$\gamma < \frac{\pi\alpha}{(1-\pi)q\sigma(w + \alpha h - \alpha)} \equiv \overline{\gamma}.$$

Hence, $c^* = 1$ if and only if $\hat{\theta} = 0$ when $g_0 = g_1 = 1$, $m_1 = 0$ and $m_0 = 1$, and $e = 1$, conditional upon $\gamma < \overline{\gamma}$.

(ii) *Education policy ($m_0^* = 1$)*: Suppose $g_0 = 1$, $e^* = 1$, and $c^* = 1$ if and only if $\hat{\theta} = 0$. Again fix $m_1 = 0$. Suppose $m_0 = 1$. A predatory ruler's expected payoff from $g_0 = 1$ when $m_0 = 1$, conditional upon $\gamma < \overline{\gamma}$, is

$$R^I + \gamma(w + \alpha h) + (1-\gamma)(1-q\sigma)(w + \alpha h - \alpha) - 1,$$

which is increasing in γ. Hence, $m_0^* = 1$ if $m_1 = 0$, $g_0 = 1$, $e^* = 1$, and $c^* = 1$ if and only if $\hat{\theta} = 0$.

Benevolent rulers never propagandize ($m_1^* = 0$): Suppose $g_0 = g_1 = 1$. Suppose $m_1 = 0$. The citizen's expected payoff from $e = 1$ when $m_0 = 0$, or when $m_0 = 1$ but $\gamma < \overline{\gamma}$, such that $c^* = 1$ if and only if $\hat{\theta} = 0$, is

$$Pr(\hat{\theta} = 0)[\hat{\pi}(0|0)q\sigma(w + \alpha h - \alpha) + \hat{\pi}(1|0)(w + \alpha h - \alpha - \max\{1 - R^I, 0\})] + Pr(\hat{\theta} = 1)\hat{\pi}(1|1)(w + \alpha h - \max\{1 - R^I, 0\})$$

$Pr(\hat{\theta} = 0|\theta = 1) = 0$ if and only if $m_1 = 0$. Deviating to $m_1 = 1$, such that $Pr(\hat{\theta} = 0|\theta = 1) > 0$, would increase the probability that the citizen chooses $c^* = 1$ and bears the cost α when $\theta = 1$, decreasing her expected payoffs. Hence, $m_1^* = 0$ when $g_0 = g_1 = 1$ and either $m_0 = 0$ or $m_0 = 1$ but $\gamma < \overline{\gamma}$, such that $c^* = 1$ if and only if $\hat{\theta} = 0$.

Now suppose $\gamma \geq \overline{\gamma}$. Recall that $m_\theta = 0$ is set by default, where $m_\theta = 1$ only if it is strictly improving. If a benevolent ruler decides to stay at $m_1 = 0$, then he knows that $c^* = 1$ for all $\hat{\theta}$ if $m_0 = 1$, or $c^* = 1$ if and only if $\theta = 0$ if $m_0 = 0$. But conditional upon $m_1 = 0$, a citizen will always condition her choice of c on $m_0 = 0$ as well, since she knows a predatory ruler cannot improve his payoffs by setting $m_0 = 1$ when $m_1 = 0$. Thus $m_1 = 0$ always yields a benevolent ruler's best possible payoff, in which $c^* \neq 1$ when $\theta = 1$.

Hence, $m_1^* = 0$ in all schooling equilibrium [SE].

Public education provision ($g_0^* = 1$): Suppose $g_0 = 1$. Recall that $m_0^* = 1$ if $g_0 = 1$ when $\gamma < \overline{\gamma}$. A predatory ruler's payoffs from $g_0 = 0$ are $R^I + w$. Hence, he prefers to invest in public education ($g_0^* = 1$) when $\gamma < \overline{\gamma}$ if and only if

$$R^I + \gamma(w + \alpha h) + (1-\gamma)(1-q\sigma)(w + \alpha h - \alpha) - 1 \geq R^I + w$$

$$\Leftrightarrow \gamma \geq 1 - \frac{\alpha h - 1}{q\sigma(w + \alpha h) + (1 - q\sigma)\alpha} \equiv \underline{\gamma}.$$

Suppose $g_1 = 1$. Recall that $\gamma \leq 1$. If $\gamma \geq \underline{\gamma}$, then $\underline{\gamma} \leq 1$. $\underline{\gamma} \leq 1$ if and only if $h \geq \alpha^{-1}$.

$m_1^* = 0$, so $\mathbb{E}(c(m_\theta|\hat{\theta})|\theta = 1) = 0$, such that $\alpha^{-1} \geq \underline{h}$, by Proposition 1.1. Thus, if $h \geq \alpha^{-1}$, then $h \geq \underline{h}$. Hence, $\gamma \geq \underline{\gamma}$ is sufficient for $h \geq \underline{h}$. By Proposition 1.1, $g_1^* = 1$ if and only if $h \geq \underline{h}$. Hence, $g_1^* = 1$ is necessary for $g_0^* = 1$ when $m_1^* = 0$, $m_0^* = 1$, $e^* = 1$, and $c^* = 1$ if and only if $\hat{\theta} = 0$. Hence, $\gamma \in [\underline{\gamma}, \overline{\gamma})$ must hold in a biased schooling equilibrium [BSE].

Confirming enrollment ($e^* = 1$): Suppose $e = 1$. Following Proposition 1.1, suppose $c^* = 1$ if and only if $\hat{\theta} = 0$ and that $g_0 = g_1 = 1$, $m_0 = 1$, and $m_1 = 0$. Then the citizen's expected payoff from $e = 1$ is

$$Pr(\hat{\theta} = 0)\hat{\pi}(0|0)q\sigma(w + \alpha h - \alpha) + Pr(\hat{\theta} = 1)\hat{\pi}(1|1)(w + \alpha h - \max\{1 - R^I, 0\})$$

if and only if $g_0 = g_1 = 1$, for which $h \geq \underline{h}$ is a necessary condition. Furthermore, by inspection, $h \geq \underline{h}$ implies that

$$Pr(\hat{\theta} = 0)\hat{\pi}(0|0)q\sigma(w + \alpha h - \alpha) + Pr(\hat{\theta} = 1)\hat{\pi}(1|1)(w + \alpha h - \max\{1 - R^I, 0\}) \geq \pi w.$$

Hence, $e^* = 1$ if and only if $g_1 = 1$ when $g_0 = 1$, $m_0 = 1$, $m_1 = 0$, and $c^* = 1$ if and only if $\hat{\theta} = 0$. Hence, $\gamma \in [\underline{\gamma}, \overline{\gamma})$ is necessary and sufficient for a BSE.

(iii) Hence, when $\gamma \in [\underline{\gamma}, \overline{\gamma})$, with probability $(1 - \pi)\gamma$, $c^* = 0$ when $\theta = 0$. To choose $c^* = 0$ when $\theta = 0$ is suboptimal *ex post*: if $c^* = 0$ when $\theta = 0$, $\tau_0^* = \tau_0^{PI}$.

(iv) $\frac{\partial \overline{\gamma}}{\partial q} = -\frac{\pi\alpha}{(1-\pi)q^2\sigma(w+\alpha h-\alpha)} < 0$, $\frac{\partial \overline{\gamma}}{\partial \sigma} = -\frac{\pi\alpha}{(1-\pi)q\sigma^2(w+\alpha h-\alpha)} < 0$, $\frac{\partial \overline{\gamma}}{\partial h} = -\frac{\pi\alpha^2}{(1-\pi)q\sigma(w+\alpha h-\alpha)^2} < 0$, $\frac{\partial \overline{\gamma}}{\partial w} = -\frac{\pi\alpha}{(1-\pi)q\sigma(w+\alpha h-\alpha)^2} < 0$, $\frac{\partial \overline{\gamma}}{\partial \alpha} = \frac{\pi w}{(1-\pi)q\sigma(w+\alpha h-\alpha)^2} > 0$, $\frac{\partial \overline{\gamma}}{\partial \pi} = \frac{\alpha}{(1-\pi)^2q\sigma(w+\alpha h-\alpha)} > 0$, $\frac{\partial \underline{\gamma}}{\partial h} = -\frac{\alpha(q\sigma(w+1)+(1-q\sigma)\alpha)}{(q\sigma(w+\alpha h)+(1-q\sigma)\alpha)^2} < 0$, $\frac{\partial \underline{\gamma}}{\partial \alpha} = -\frac{1+q\sigma(wh+h-1)}{(q\sigma(w+\alpha h)+(1-q\sigma)\alpha)^2} < 0$, $\frac{\partial \underline{\gamma}}{\partial q} = \frac{\sigma(\alpha h-1)(w+\alpha h-\alpha)}{(q\sigma(w+\alpha h)+(1-q\sigma)\alpha)^2} > 0$, $\frac{\partial \underline{\gamma}}{\partial \sigma} = \frac{q(\alpha h-1)(w+\alpha h-\alpha)}{(q\sigma(w+\alpha h)+(1-q\sigma)\alpha)^2} > 0$, and $\frac{\partial \underline{\gamma}}{\partial w} = \frac{q\sigma(\alpha h-1)}{(q\sigma(w+\alpha h)+(1-q\sigma)\alpha)^2} > 0$, the last three of which hold only when $h > \alpha^{-1}$. □

Proof of Remark 1

Proof. (i) By Propositions 1.1 and 2.1-2, $h \geq \underline{h}$ is necessary and sufficient for $g_1^* = 1$ and therefore $e^* = 1$, and $\gamma < \overline{\gamma}$ is necessary and sufficient for $c^* = 1$ if and only if $\hat{\theta} = 0$ when $m_0 = 1$ and $m_1 = 0$, such that $m_0^* = 1$ when $g_0 = 1$. However, by Proposition 2.2, $\gamma < \underline{\gamma}$ implies $g_0^* = 0$, inducing a weak non-schooling equilibrium [WNE].

(ii) By Proposition 2.4, $\underline{\gamma}$ is strictly decreasing in h. $\underline{\gamma}|_{h=\overline{h}} = 0$. Hence, $\max\{\underline{\gamma}, 0\} = 0$ when $h \geq \overline{h}$. $\gamma > 0$. Hence, $\gamma \not< \underline{\gamma}$ for all $h \geq \overline{h}$. □

Proof of Remark 2

Proof. By Remark 1, $\gamma < \underline{\gamma}$ is sufficient for $h < \overline{h}$, and $\gamma < \underline{\gamma}$ and $\gamma < \overline{\gamma}$ are necessary and sufficient for a WNE. Hence, $h \in [\underline{h}, \overline{h})$ and $\gamma < \underline{\gamma} < \overline{\gamma}$ are sufficient for a WNE. However, if $\gamma = \hat{\gamma} \in [\underline{\gamma}, \overline{\gamma})$, then a predatory ruler would prefer $g_0^* = 1$ and set $m_0^* = 1$, inducing a BSE. Moreover, because $\gamma > 0$, $\underline{\gamma} > 0$ if $\gamma < \underline{\gamma}$. Thus when $\gamma < \overline{\gamma}$, a citizen's expected payoffs from $e = 1$ are *strictly* greater in a BSE than a WNE. Hence, when $h \in [\underline{h}, \overline{h})$ and $\gamma < \underline{\gamma} < \overline{\gamma}$, $\gamma = \hat{\gamma}$ would strictly improve the citizen's expected payoffs. □

Proof of Proposition 3

Proof. (i) Suppose $g_0 = g_1 = 1$, $m_0 = 1$, $m_1 = 0$, $e = 1$, and $c = 1$ for all $\hat{\theta}$. Then by Proposition 2, $c^* = 1$ for all $\hat{\theta}$ if and only if $\gamma \geq \overline{\gamma}$. When $\gamma \geq \overline{\gamma}$, a predatory ruler's payoff when $g_0 = 1$, $m_0 = 1$, $e^* = 1$, and $c^* = 1$ is $R^I + (1 - q\sigma)(w + \alpha h - \alpha) - 1$. His reservation payoff is $R^I + w$. By Proposition 1, it must be the case that $h \geq \overline{h}$ if $g_0^* = 1$. By Proposition 1, $\overline{h} > \underline{h}$, such that if $g_0^* = 1$, then $g_1^* = 1$, and if $g_1^* = 1$, then $e^* = 1$.

However, recall that a ruler never deviates from $m_\theta = 0$ if he is indifferent. Suppose $m_0 = 0$. By Proposition 1.1, the citizen chooses $c^* = 1$ only if $\theta = 0$. Then a predatory ruler's payoff when $g_0 = 1$ but $m_0 = 0$ is still $R^I + (1 - q\sigma)(w + \alpha h - \alpha) - 1$. Hence, when $\gamma \geq \overline{\gamma}$ and $h > \underline{h}$, $m_0^* = 0$, where $g_0^* = g_1^* = 1$, $m_1^* = 0$, $e^* = 1$, and $c^* = 1$ if and only if $\hat{\theta} = 0$. Hence, $h \geq \overline{h}$ and $\gamma \geq \overline{\gamma}$ are necessary and sufficient for a DSE when $m_0 \in \{0, 1\}$.

(ii) Now let $h \in [\underline{h}, \overline{h})$, but suppose a predatory ruler chooses $g_0 = 1$ and $m_0 = 1$. Then $c^* = 1$ for all $\hat{\theta}$. However, since his payoff from $m_0 = 1$ is the same as his payoff from $m_0 = 0$ when $\gamma \geq \overline{\gamma}$, $m_0^* \neq 1$. Suppose he instead sets $m_0 = 0$. Then $c^* = 1$ if and only if $\hat{\theta} = 0$. But by Proposition 1.1, $g_0^* \neq 1$ when $m_0 = 0$ and $h < \overline{h}$. Hence, when $\gamma \geq \overline{\gamma}$ but $h \in [\underline{h}, \overline{h})$, $g_1^* = 1$ but $g_0^* = 0$ and $m_0^* = 0$, where $m_1^* = 0$, $e^* = 1$ if and only if $g_\theta = 1$ is observed, and $c^* = 1$ whenever $e^* = 1$. Hence, $h \in [\underline{h}, \overline{h})$ and $\gamma \geq \overline{\gamma}$ are necessary and sufficient for a repressive non-schooling equilibrium [RNE] whenever $m_0 \in \{0, 1\}$.

(iii) By Proposition 2.4, $\overline{\gamma}$ is strictly decreasing in h. Finally, $\overline{\gamma}|_{h=\overline{h}} = \frac{\alpha\pi(1-q\sigma)}{(1-\pi)q\sigma(1+w)} > 0$. □

Proof of Remark 3

Proof. By Proposition 3.1, if $\gamma \geq \overline{\gamma} > \underline{\gamma}$, $m_0^* = 0$, such that a predatory ruler's payoff from $g_0 = 1$ is $R^I + (1 - q\sigma)(w + \alpha h - \alpha) - 1$. His payoff from $g_0 = 0$ is $R^I + w$. By Proposition 2, $\gamma \in [\underline{\gamma}, \overline{\gamma})$ is necessary and sufficient for a BSE, where his payoff is $R^I + \gamma(w + \alpha h) + (1 - \gamma)(1 - q\sigma)(w + \alpha h - \alpha) - 1$. The latter payoff is always strictly greater than either of the first two when $\gamma > \max\{\underline{\gamma}, 0\}$, conditional upon $\gamma < \overline{\gamma}$.

A ruler's payoff in a BSE is strictly increasing in γ for all $\gamma < \overline{\gamma}$. $\gamma \leq 1$. If $\gamma \geq \overline{\gamma}$, then it must be the case that $\overline{\gamma} \leq 1$. Hence, the payoff-maximizing $\gamma \equiv \tilde{\gamma}^* = \lim_{\overline{\varepsilon} \to 0} \overline{\gamma} - \overline{\varepsilon} < 1$.

It follows from this and Remark 2 that this is also weakly payoff-improving for the citizen when $h \in [\underline{h}, \overline{h})$. □

Proof of Example 1

Proof. Let $h \geq \underline{h}$. By Propositions 1.1 and 2.2, $e^* = 1$, $m_1^* = 0$, and $g_1^* = 1$. Fix $\gamma = \overline{\gamma}|_{h=\overline{h}}$.

Suppose $h = \overline{h}$, such that $\gamma = \overline{\gamma}$. By Proposition 2.1, $c^* = 1$ for all $\hat{\theta}$ if $m_0^* = 1$. Otherwise $c^* = 1$ if and only if $\hat{\theta} = 0$. Then by Proposition 3.1, $m_0^* = 0$, such that a predatory ruler's payoff from $g_0 = 1$ is $R^I + (1 - q\sigma)(w + \alpha\overline{h} - \alpha) - 1$.

Now suppose $h = \overline{h} - \eta < \overline{h}$. By Proposition 2.4, $\gamma < \overline{\gamma}|_{h=\overline{h}-\eta}$. Now $m_0^* = 1$ whenever $g_0 = 1$, such that a predatory ruler's payoff from $g_0 = 1$ is

$$\begin{aligned} R^I + \overline{\gamma}|_{h=\overline{h}}(w + \alpha(\overline{h} - \eta)) + (1 - \overline{\gamma}|_{h=\overline{h}})(1 - q\sigma)(w + \alpha(\overline{h} - \eta) - \alpha) - 1 \\ = R^I + \frac{\alpha\pi(1 - q\sigma)(w + \alpha\overline{h} - \alpha\eta)}{(1 - \pi)q\sigma(1 + w)} + (1 - q\sigma)(w + \alpha\overline{h} - \alpha\eta - \alpha) \\ - \frac{\alpha\pi(1 - q\sigma)^2(w + \alpha\overline{h} - \alpha\eta - \alpha)}{(1 - \pi)q\sigma(1 + w)} - 1. \end{aligned}$$

Plugging in $\overline{h}$, the latter is greater if and only if $\eta < \frac{\pi(q\sigma(1+w)+(1-q\sigma)\alpha)}{(1-q\sigma)q\sigma(\pi\alpha+(1-\pi)(1+w))} \equiv \hat{\eta} > 0$. □

Proof of Proposition 4

Proof. Preliminary information: Optimal taxes and rents $(\tau_{0t}^*, \tau_{1t}^*, R_{0t}^*, R_{1t}^*)$, and in turn r_t^*, are the same as in the basic model for all t. When $\theta = 1$, this follows from the citizen's myopia and the fact that if $h \geq \underline{h}$ in period 1, then $h_2 > \underline{h}$. When $\theta = 0$ and $t = 2$, this follows from Lemma 1. When $\theta = 0$ and $t = 1$, a predatory ruler's tax choice τ_{01} must account for the possibility that if $r_1 = 1$, then he survives into period 2 with probability q. However, if he survives, his $t = 2$ payoffs are the same as if $r_1 = 0$. Let his second period payoffs be V_2. A predatory ruler's expected payoff when $e_1 = 1$ and $c_1 = 1$ if he sets the highest possible τ_{01} at which $r_1^* = 0$ is $R^I + (1 - q\sigma)(w + \alpha h - \alpha) - 1 + \beta V_2$.

$$\begin{aligned} (1 - q\sigma)(w + \alpha h - \alpha) + \beta V_2 > (1 - q)[\sigma(w + \alpha h - \alpha) + \beta V_2] \\ \Leftrightarrow (1 - \sigma)(w + \alpha h - \alpha) + q\beta V_2 > 0. \end{aligned}$$

Sustaining a BSE into period 2 ($c_2^* = 1$ if and only if $\hat{\theta}_2 = 0$): Let $h \geq \underline{h}$. By Propositions 1.1 and 2.2, $e_t^* = 1$, $m_{1t}^{A*} = 0$, and $g_{1t}^* = 1$ for $t = 1$ and therefore all t. A benevolent ruler cannot benefit from weakening the productivity of education, such that $m_1^{B*} = 0$.

Suppose period 1 is a BSE: Let $\gamma < \overline{\gamma}(h_1|m_0^B = 0)$. $h_1 = h$ if $m_0^B = 0$, such that, by Propositions 2.1 and 2.4, $c_1^* = 1$ if and only if $\hat{\theta}_1 = 0$ for all m_{01}^A and m_0^B.

Suppose $m_{01}^A = 1$ and $m_0^B = 0$. Without unproductive education, period 2 is a DSE: Let $\gamma \geq \overline{\gamma}(h_2|m_0^B = 0)$, where by Proposition 2.1,

$$\overline{\gamma}(h_2|m_0^B = 0) \equiv \frac{\pi_2\alpha}{(1 - \pi_2)q\sigma(w + \alpha h + \alpha(1 - \delta)h - \alpha)},$$

such that $c_2^* = 1$ for all $\hat{\theta}_2$ and thus $m_{02}^{A*} = 0$ if there is no unproductive education ($m_0^B = 0$).

Now instead suppose that $m_0^B = 1$, such that $h_1 = (1 - \rho)h$ and $h_2 = h + (1 - \delta)(1 - \rho)h$. Suppose that $m_{02}^A = 1$ but $c_2 = 0$ when $\hat{\theta}_2 = 1$. By Proposition 2.1, this is true if and only if

$$\gamma < \frac{\pi_2\alpha}{(1 - \pi_2)q\sigma(w + \alpha h + \alpha(1 - \delta)(1 - \rho)h - \alpha)} \equiv \overline{\gamma}(h_2|m_0^B = 1).$$

Rearranging, $c_2^* = 1$ if and only if $\hat{\theta}_2 = 0$ when $m_{02}^A = 1$ and $m_0^B = 1$ if and only if

$$\rho > 1 - \frac{\pi_2\alpha - \gamma(1-\pi_2)q\sigma(w+\alpha h-\alpha)}{\gamma(1-\pi_2)q\sigma\alpha(1-\delta)h} \equiv \underline{\rho}.$$

$\rho > \underline{\rho}$ implies $\gamma < \overline{\gamma}(h_1|m_0^B = 1)$, since $\rho > \underline{\rho}$ implies $\gamma < \overline{\gamma}(h_2|m_0^B = 1)$, and $h_2 > h_1$ implies $\overline{\gamma}(h_2|m_0^B = 1) < \overline{\gamma}(h_1|m_0^B = 1)$ for all $\pi_2 < \pi_1$ by Proposition 2.4.

$\rho > \underline{\rho}$ also implies $\gamma < \overline{\gamma}(h_1|m_0^B = 0)$, since $\rho > \underline{\rho}$ implies $\underline{\rho} < 1$ for all $\rho < 1$, and $\underline{\rho} < 1$ implies $\gamma < \overline{\gamma}(h_1|m_0^B = 0)$. Finally, $\gamma \geq \overline{\gamma}(h_2|m_0^B = 0)$ implies $\underline{\rho} \geq 0$.

Unproductive propaganda ($m_{01}^{A*} = m_{02}^{A*} = m_0^{B*} = 1$)? Suppose $m_{01}^A = m_{02}^A = m_0^B = 1$. By Propositions 2.1 and 3.1, a predatory ruler's expected payoff in period 1 from $m_0^B = 1$, conditional upon $\gamma < \overline{\gamma}(h_1|m_0^B = 0)$, $\gamma \geq \overline{\gamma}(h_2|m_0^B = 0)$, and $\rho > \underline{\rho}$, is

$$\begin{aligned} &R^I + \gamma(w+\alpha(1-\rho)h) + (1-\gamma)(1-q\sigma)(w+\alpha(1-\rho)h-\alpha) - 1 + \beta[R^I + \gamma(w+\alpha h \\ &\quad + \alpha(1-\delta)(1-\rho)h) + (1-\gamma)(1-q\sigma)(w+\alpha h+\alpha(1-\delta)(1-\rho)h-\alpha) - 1]. \end{aligned} \tag{C.1}$$

His expected payoff from $m_0^B = 0$ in period 1 is

$$R^I + \gamma(w+\alpha h) + (1-\gamma)(1-q\sigma)(w+\alpha h-\alpha) - 1 + \beta(R^I + (1-q\sigma)(w+\alpha h+\alpha(1-\delta)h-\alpha) - 1). \tag{C.2}$$

The former is greater than the latter if and only if

$$\rho < \frac{\beta\gamma(q\sigma(w+\alpha(2-\delta)h) + (1-q\sigma)\alpha)}{\alpha h(1+(1-\delta)\beta)(1-(1-\gamma)q\sigma)} \equiv \overline{\rho} > 0.$$

Hence, when $\gamma < \overline{\gamma}(h_1|m_0^B = 0)$ but $\gamma \geq \overline{\gamma}(h_2|m_0^B = 0)$, $\rho \in (\underline{\rho}, \overline{\rho})$ is a necessary and sufficient condition for $c_t^* = 1$ if and only if $\hat{\theta}_t = 0$ and $m_{0t}^{A*} = m_0^{B*} = 1$ for all t, when $g_{0t} = g_{1t}^* = e_t^* = 1$ for all t, and $m_{1t}^{\gamma^*} = m_1^{B*} = 0$ for all t.

Finally, a predatory ruler's expected payoff from $g_{01} = g_{02} = 0$ is

$$(1+\beta)(R^I + w). \tag{C.3}$$

His expected payoff from $g_{01} = 0$ and $g_{02} = 1$, conditional upon $\gamma < \overline{\gamma}(h_2 = h)$, is

$$R^I + w + \beta(R^I + \gamma(w+\alpha h) + (1-\gamma)(1-q\sigma)(w+\alpha h-\alpha) - 1). \tag{C.4}$$

When $\gamma < \overline{\gamma}(h_2|m_0^B = 0)$, then unproductive education is unneeded. Finally, recall that if $\gamma \not< \overline{\gamma}(h_1|m_0^B = 0)$, then $\underline{\rho} \geq 1$, such that $\rho \not> \underline{\rho}$. Hence, when $\gamma < \overline{\gamma}(h_1|m_0^B = 0)$, $\gamma \geq \overline{\gamma}(h_2|m_0^B = 0)$, $h \geq \underline{h}$, and (C.1) exceeds the greater of (C.3) and (C.4), $\rho \in (\underline{\rho}, \overline{\rho})$ is necessary and sufficient for an unproductive biased schooling equilibrium [UBSE] in which human capital accumulates, yet $m_{0t}^{A*} = 1$ and $c_t^* = 1$ if and only if $\hat{\theta}_t = 0$ for all t.

Finally, to confirm that this equilibrium space is nontrivial, set $w = 0$, $\beta = 1$, $h = \overline{h}$, and $\gamma = \overline{\gamma}(h_2|m_0^B = 0)$. Then $\underline{\rho} = 0$. Then since $\overline{\rho} > 0$, $\overline{\rho} > \underline{\rho}$. Moreover, (C.2) exceeds (C.4), and (C.4) exceeds (C.3), such that (C.1) exceeds (C.2), (C.3), and (C.4) whenever $\rho < \overline{\rho}$.

(ii) $\frac{\partial \rho}{\partial \gamma} = \frac{\pi_2}{\gamma^2(1-\pi_2)q\sigma(1-\delta)h} > 0$, $\frac{\partial \rho}{\partial h} = \frac{\pi_2\alpha-\gamma(1-\pi_2)q\sigma(w-\alpha)}{\gamma(1-\pi_2)q\sigma\alpha(1-\delta)h^2} > 0$,
$\frac{\partial \overline{\rho}}{\partial \gamma} = \frac{\beta(1-q\sigma)(q\sigma(w+\alpha(2-\delta)h)+(1-q\sigma)\alpha)}{\alpha h(1+(1-\delta)\beta)(1-(1-\gamma)q\sigma)^2} > 0$, and $\frac{\partial \overline{\rho}}{\partial h} = -\frac{\beta\gamma(q\sigma w+(1-q\sigma)\alpha)}{\alpha h^2(1+(1-\delta)\beta)(1-(1-\gamma)q\sigma)} < 0$. □

Discussion: the basic model with two citizens

A natural reframing of the model that is closer in spirit to Edmond (2013) is to have it assume two citizens, $i = a, b$, each from an *ex ante* different social group. Citizens meet if they both enroll in a public school, and, after observing a public curriculum $\hat{\theta}$, form a political club when $c_a = c_b = 1$. In this formulation, propaganda prevents political mobilization by undermining coordination, i.e. preventing a political network from being formed between different social groups. For instance,

	$c_b = 1$	$c_b = 0$
$c_a = 1$	$\hat{\pi}(0\|\hat{\theta})q\sigma(w+\alpha h-\alpha)$ $+\hat{\pi}(1\|\hat{\theta})(w+\alpha h-\alpha-\tau_1^{BE})$, $\hat{\pi}(0\|\hat{\theta})q\sigma(w+\alpha h-\alpha)$ $+\hat{\pi}(1\|\hat{\theta})(w+\alpha h-\alpha-\tau_1^{BE})$	$\hat{\pi}(1\|\hat{\theta})(w+\alpha h-\alpha-\tau_1^{BE})$, $\hat{\pi}(1\|\hat{\theta})(w+\alpha h-\tau_1^{BE})$
$c_b = 0$	$\hat{\pi}(1\|\hat{\theta})(w+\alpha h-\tau_1^{BE})$, $\hat{\pi}(1\|\hat{\theta})(w+\alpha h-\alpha-\tau_1^{BE})$	$\hat{\pi}(1\|\hat{\theta})(w+\alpha h-\tau_1^{BE})$, $\hat{\pi}(1\|\hat{\theta})(w+\alpha h-\tau_1^{BE})$

gives the payoff matrix for each citizen's political club decision when there are two citizens. As in the basic set up, $Pr(\hat{\theta} = 0|\theta = 1) = 0$, since a benevolent ruler never uses propaganda. In turn, $\hat{\pi}(1|0) = 0$ and $\hat{\pi}(0|0) = 1$. Hence, when citizens observe $\hat{\theta} = 0$, the only Nash equilibrium at that stage is $c_a^* = c_b^* = 1$.

Thus a predatory ruler on one hand wants to use propaganda because it will lower the probability that citizens observe $\hat{\theta} = 0$ and raise the probability that they observe $\hat{\theta} = 1$, at which point there are potentially two equilibria: one with $c_a^* = c_b^* = 0$ and one with $c_a^* = c_b^* = 1$. However, propaganda (i.e. high γ) also increases $\hat{\pi}(0|1)$ and decreases $\hat{\pi}(1|1)$, to the extent that the $c_a^* = c_b^* = 0$ equilibrium may disappear from the $\hat{\theta} = 1$ case as well. Hence, the non-monotonic effect of propaganda.

Parameter key

Summary of variables and parameters		
Parameter	Interpretation	Description
$\theta \in \{0,1\}$	The ruler's type	The ruler's preference for policies that maximize social welfare versus his own rents.
$\pi \in (0,1)$	The citizen's trust level	The citizen's prior belief that the ruler will maximize social welfare (i.e. $Pr(\theta = 1)$).
$g_\theta \in \{0,1\}$	Public education provision	The ruler's choice of whether to invest in public education.
$e \in \{0,1\}$	Enrollment in school	The citizen's choice of whether to attend public school.
$w > 0$	Non-productive income	–
$h > 1$	Productivity effect of education	A measure of productivity growth under education.
$\alpha > 0$	Income effect of human capital	–
$\tau_\theta \geq 0$	Ruler's choice of tax	–
$R^I \geq 0$	Initial rents*	The ruler's rents at the start of the game.
$R_\theta \geq 0$	Final rents	The ruler's rents at the end of the game.
$c \in \{0,1\}$	Political club participation	The citizen's choice of whether to join a political club while enrolled in school.
$r \in \{0,1\}$	Mounting of a political challenge	The citizen's choice to contest a ruler's choices of τ_θ, conditional upon $c = 1$.
$q \in (0,1)$	Political power	The probability that the citizen wins a political challenge.
$\sigma \in (0,1)$	Rule of law	The extent to which political challenges are resolved nonviolently (i.e. are non-destructive).
$\hat{\theta} \in \{0,1\}$	Curriculum	Informs upon and mirrors the state and structure of the political environment, as defined by the ruler's type. A *ruler-favorable* curriculum is $\hat{\theta} = 1$, regardless of θ.
$m_\theta \in \{0,1\}$	Ruler's education policy	The ruler's choice of whether to embed propaganda within (i.e. bias) the education system. *Neutral* policy is $m_\theta = 0$. *Biased* policy is $m_\theta = 1$.
$\gamma \in (0,1]$	State of the propaganda technology	The extent to which propaganda is successfully transmitted under a biased education policy. Higher γ implies stronger technology.
$\rho \in (0,1)$	Unproductive education	The extent to which propaganda decreases the productivity of education in the extension model's first period.

*Model later assumes $R^I \geq \max\{1 - (1-q)\sigma(w + \alpha h - \alpha), 0\}$.

Tables

Table C.1: Summary statistics (I: model 1, unbalanced)

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
PressConstraints	2448	4.710	2.253	0.5	9.4
Crisis	2448	0.061	0.240	0	1
Education expenditures (% Gov)	1457	14.832	4.750	4.771	37.687
lnEducation	1457	2.646	0.324	1.562	3.629
GDP per capita ($ 2010)	2442	11976.79	17214.42	186.919	91593.633
lnGDPpc	2442	8.301	1.589	5.231	11.425
Gov't expenditures (% GDP)	2315	15.839	5.819	2.804	69.543
lnGovExpend	2315	2.695	0.377	1.031	4.242
Military expenditures (% GDP)	2256	2.170	2.295	0	39.607
lnMilitary	2223	0.464	0.973	-6.908	3.679
$lnMilitary_2$	2256	1.021	0.473	0	3.704
Polity	2439	4.173	6.140	-10	10
Country	136	–	–	–	–
Year	18	–	–	1997	2014

Table C.1: Summary statistics (II: model 1, balanced)

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
PressConstraints	1548	4.396	2.306	0.5	9.4
Crisis	1548	0.056	0.229	0	1
Education expenditures (% gov)	1157	14.837	4.388	4.771	37.687
lnEducation	1157	2.654	0.297	1.562	3.629
GDP per capita ($ 2010)	1545	14114.48	18503.57	205.07	91593.633
lnGDPpc	1545	8.511	1.600	5.323	11.425
Gov't expenditures (% GDP)	1516	15.776	5.174	4.157	31.58
lnGovExpend	1516	2.699	0.358	1.425	3.453
Military expenditures (% GDP)	1464	2.045	1.881	0.001	16.157
lnMilitary	1464	0.376	1.072	-6.908	2.782
$lnMilitary_2$	1464	0.996	0.448	0.001	2.842
Polity	1540	5.022	6.021	-10	10
Country	86	–	–	–	–
Year	18	–	–	1997	2014

Table C.1: Summary statistics (III: model 2)

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Age_t	77056	26.345	5.995	18	43
$PressConstraints_{b=18}$	58758	4.291	1.957	0.5	9.9
$PressConstraints_{b=22}$	57985	4.235	1.972	0.5	9.9
$EducationAttainment$ (EA)	74000	4.157	2.044	0	7
Education completion age	65335	20.356	7.612	1	99
Employed	76384	0.524	0.499	0	1
Immigrant	36337	0.033	0.180	0	1
$IncomeLevel_t$	71618	3.784	2.257	0	9
International	27283	0.010	0.100	0	1
Noncitizen	28673	0.020	0.141	0	1
Number of children	76341	0.917	1.367	0	8
Participation (active)	69854	1.463	1.272	0	4
Participation (passive)	74846	2.578	1.673	0	6
Participation (demonstrate)	71748	0.675	0.712	0	2
Participation (petition)	71286	0.797	0.770	0	2
Participation (importance)	75743	1.301	0.978	0	3
Participation (interest)	75947	1.272	0.946	0	3
Polity_t	77056	6.735	3.862	-6	10
Relationship status	76837	1.272	1.387	0	3
Sex	77023	0.514	0.500	0	1
Student_t	76384	0.187	0.390	0	1
Country	70	–	–	–	–
Year	20	–	–	1994	2014

Table C.2: FH score as a propaganda proxy (I: correlations)

Dependent variable:	Academic constraints					
PressConstraints	0.086***	–	0.072***	0.077***	0.078***	0.073***
	(0.012)		(0.024)	(0.024)	(0.026)	(0.026)
Polity	–	−0.028***	−0.006	−0.005	−0.021	−0.015
		(0.006)	(0.010)	(0.011)	(0.016)	(0.011)
Observations	195	165	165	162	129	136
Exclude if missing Polity?	No	No	No	Yes	No	No
Exclude autocracies > 4?	No	No	No	No	Yes	No
Exclude "crisis" countries?	No	No	No	No	Yes	No
Table 3.1 countries only?	No	No	No	No	No	Yes
R^2	0.20	0.13	0.17	0.18	0.23	0.26

Robust standard errors are reported in parentheses, with *** denoting significance at the 1% level. Independent variables are the FH press constraint score (0 to 10) and the Polity IV POLITY2 score (-10 to 10), both averaged over 2011 to 2016. Countries missing POLITY2 scores for any of those years are excluded in column 4. Dependent variable is an academic constraints indicator, which equals 1 if a country's scholars have ever faced imprisonment, prosecution, or travel restrictions in the Scholars at Risk (SAR) Network's Incident Index from 2011 to October 2017. A country is considered a "crisis" country if it received a -66, -77, or -88 POLITY score from 2011 to 2016. These are excluded from column 5, since countries in crisis may have weak states that are unable to enact or enforce education policy. Countries with Polity IV Autocracy scores over 5 are also omitted in column 5, since strong autocracies tend to be less transparent and therefore have fewer reported incidents.

Table C.2: FH score as a propaganda proxy (II: summary statistics)

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
AcademicConstraints	195	0.287	0.454	0	1
PressConstraints	195	4.795	2.357	0.95	9.7
Polity	165	4.075	6.087	-10	10

Table C.3: Model 1 with alternative military transformation

Dependent variable:	Degree of press constraints and media censorship			
	(1a)	(1b)	(1c)	(1d)
lnEducation	2.028***	1.881**	2.542***	2.522***
	(0.778)	(0.780)	(0.929)	(0.911)
$lnEdu \times lnGDPpc$	−0.268***	−0.251**	−0.312***	−0.309***
	(0.098)	(0.098)	(0.117)	(0.115)
lnGDPpc	0.339	0.274	0.418	0.398
	(0.332)	(0.335)	(0.409)	(0.394)
Polity	−0.111***	−0.104***	−0.104***	−0.098***
	(0.022)	(0.023)	(0.025)	(0.026)
Crisis	–	0.198*	–	0.167
		(0.108)		(0.123)
$lnMilitary_2$	−0.047	−0.005	0.051	0.091
	(0.154)	(0.175)	(0.178)	(0.202)
lnGovExpend	–	−0.043	–	−0.200
		(0.210)		(0.236)
Country fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Balanced panel?	No	No	Yes	Yes
Mean imputation?	No	No	No	No
Observations	1379	1356	1101	1092
Countries	133	132	85	85
Adj. R^2	0.23	0.22	0.22	0.21

Standard errors are clustered by country, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. The variable $lnMilitary_2$ corrects for deletion or exacerbation of zero or near-zero military expenditure values post-logarithm by adding 1 to each value pre-logarithm.

Table C.4: Model 1, correlates of missing education data

Dependent variable:	Correlates of missing education data						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
lnGDPpc	−0.032***	–	–	–	–	−0.001	−0.204
	(0.006)					(0.008)	(0.146)
Polity	–	−0.011***	–	–	–	−0.009***	−0.004
		(0.002)				(0.002)	(0.006)
Crisis	–	–	0.151***	–	–	−0.060	−0.024
			(0.042)			(0.051)	(0.056)
lnMilitary	–	–	–	0.031***	–	0.022*	−0.002
				(0.011)		(0.012)	(0.059)
lnGovExpend	–	–	–	–	−0.045*	−0.043	−0.063
					(0.027)	(0.034)	(0.115)
Country fixed effects	No	No	No	No	No	No	Yes
Year fixed effects	No	No	No	No	No	No	Yes
Observations	2442	2439	2439	2223	2315	2134	2134
Adj. R^2	0.01	0.02	0.00	0.03	0.00	0.02	0.19

Robust standard errors reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table C.5: Model 1 with categorical polity controls ($= \{-10, -9, ..., 10\}$)

Dependent variable:	Degree of press constraints and media censorship						
	(1a)	(1b)	(1c)	(1d)	(1e)	(2)	(3)
lnEducation	1.784**	1.723***	1.729***	1.699***	1.493**	1.642**	2.207***
	(0.880)	(0.576)	(0.575)	(0.593)	(0.589)	(0.636)	(0.687)
$lnEdu \times lnGDPpc$	−0.234**	−0.219***	−0.220***	−0.219***	−0.193**	−0.208***	−0.269***
	(0.111)	(0.074)	(0.074)	(0.076)	(0.075)	(0.078)	(0.089)
lnGDPpc	0.198	0.312	0.310	0.342	0.310	0.336	0.518
	(0.367)	(0.298)	(0.298)	(0.293)	(0.295)	(0.267)	(0.338)
Crisis	–	–	0.044	0.087	0.073	0.035	0.050
			(0.104)	(0.109)	(0.109)	(0.146)	(0.124)
lnMilitary	–	–	–	−0.048	−0.049	−0.028	0.002
				(0.094)	(0.106)	(0.083)	(0.116)
lnGovExpend	–	–	–	–	0.100	−0.076	−0.035
					(0.199)	(0.163)	(0.229)
Polity controls?	No	Yes	Yes	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Balanced panel?	No	No	No	No	No	No	Yes
Mean imputation?	No	No	No	No	No	Yes	No
Observations	1455	1452	1452	1366	1343	2134	1092
Countries	136	136	136	132	131	132	85
Adj. R^2	0.08	0.31	0.31	0.32	0.31	0.33	0.31

Standard errors are clustered by country, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. Column 2 applies a country mean imputation procedure to the education expenditure data and controls for a dummy, which takes a value of 1 when a value is missing.

Table C.6: Model 1 with standardized explanatory variables

Dependent variable:	Proxy for propaganda: press constraints and media censorship			
	(1a)	(1b)	(1c)	(1d)
lnEducation	−0.050	−0.055	−0.061	−0.056
	(0.043)	(0.044)	(0.046)	(0.047)
$lnEdu \times lnGDPpc$	−0.124***	−0.125***	−0.133***	−0.124**
	(0.047)	(0.047)	(0.050)	(0.050)
lnGDPpc	−0.583	−0.621	−0.629	−0.625
	(0.450)	(0.449)	(0.455)	(0.466)
Polity	−0.666***	−0.645***	−0.654***	−0.637***
	(0.122)	(0.121)	(0.136)	(0.144)
Crisis	–	0.176*	0.189*	0.194*
		(0.100)	(0.106)	(0.108)
lnMilitary	–	–	−0.025	0.008
			(0.099)	(0.115)
lnGovExpend	–	–	–	−0.017
				(0.081)
$lnEdu + lnEdu \times lnGDP_{min}$	0.190**	0.187**	0.196**	0.183*
	(0.088)	(0.085)	(0.091)	(0.094)
$lnEdu + lnEdu \times lnGDP_{10\%}$	0.118*	0.114*	0.118*	0.110
	(0.065)	(0.063)	(0.067)	(0.070)
$lnEdu + lnEdu \times lnGDP_{90\%}$	−0.231**	−0.238**	−0.256**	−0.237**
	(0.092)	(0.093)	(0.098)	(0.099)
Country fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Balanced panel?	No	No	Yes	Yes
Mean imputation?	No	No	No	No
Observations	1452	1452	1366	1343
Countries	136	136	132	131
Adj. R^2	0.22	0.23	0.23	0.22

Standard errors are clustered by country, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. Coefficients and standard errors for all non-binary explanatory variables are standardized with mean zero. Rows labeled $lnEdu + lnEdu \times lnGDP$ report the combined effect (and standard errors) of *lnEducation* and $lnEdu \times lnGDPpc$ at different levels (sample minimum, 10% and 90%) within the sample distribution of *lnGDPpc*.

Table C.7: Model 2 with separated variables

	Press constraints and political participation							
Dependent variable:	Interest		Importance		Demonstration		Petition	
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
$EducationAttainment$	0.069***	0.116***	0.026	0.054**	0.045**	0.071***	0.071***	0.061***
	(0.021)	(0.023)	(0.027)	(0.021)	(0.022)	(0.020)	(0.018)	(0.019)
$EA \times Constraints_b$	−0.005*	−0.035***	−0.002	−0.017**	−0.001	−0.015***	−0.004*	−0.006
	(0.003)	(0.007)	(0.003)	(0.007)	(0.002)	(0.005)	(0.002)	(0.006)
$EA \times Constraints_b^2$	–	0.003***	–	0.002**	–	0.002***	–	0.0004
		(0.001)		(0.001)		(0.001)		(0.001)
$EA\times$ Polity$_t$	0.001	0.002	0.003	0.003**	0.002	0.001	0.001	0.002*
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
$PressConstraints_b$	0.002	0.157***	0.003	0.088*	0.009	0.076***	0.011	0.007
	(0.014)	(0.035)	(0.016)	(0.046)	(0.010)	(0.026)	(0.011)	(0.032)
$PressConstraints_b^2$	–	−0.015***	–	−0.008	–	−0.007**	–	0.0003
		(0.004)		(0.005)		(0.003)		(0.003)
$Polity_t$	0.007	0.014	−0.009	0.003	0.022*	0.025**	0.006	0.016
	(0.015)	(0.012)	(0.019)	(0.011)	(0.012)	(0.010)	(0.016)	(0.011)
Age$_t$	0.024**	0.023**	0.033***	0.027***	0.004	−0.001	0.006	0.007
	(0.011)	(0.010)	(0.013)	(0.010)	(0.011)	(0.009)	(0.009)	(0.008)
Age$_t^2$	−0.0002	−0.0002	−0.0004**	−0.0003*	0.00005	0.0001	0.00003	0.00001
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0001)
$IncomeLevel_t$	0.021***	0.018***	0.011***	0.008***	−0.002	−0.004	0.004	0.006***
	(0.004)	(0.003)	(0.004)	(0.003)	(0.005)	(0.004)	(0.003)	(0.002)
Student$_t$	–	0.091***	–	0.063***	–	0.081***	–	0.103***
		(0.020)		(0.014)		(0.020)		(0.019)
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Non-students only?	Yes	No	Yes	No	Yes	No	Yes	No
Exclude immigrants?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exclude non-citizens?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	32787	50123	32651	49952	31221	47637	30962	47229
Countries	65	70	65	70	65	70	65	70
Adj. R^2	0.11	0.10	0.10	0.09	0.13	0.13	0.26	0.26

Standard errors are clustered by country, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. Base year (b) is derived from the survey year (t), minus the respondent's age (Age_t), plus 18. All regressions control for sex, relationship status, number of children, and employment status.

Table C.8: Model 2 with alternative base age ($b = 22$)

	Press constraints and political participation					
Dependent variable:	Passive participation			Active participation		
	(1a)	(1b)	(1c)	(2a)	(2b)	(2c)
$EducationAttainment$	0.066	0.151***	0.149***	0.123***	0.152***	0.148***
	(0.044)	(0.045)	(0.045)	(0.035)	(0.035)	(0.032)
$EA \times PressConstraints_b$	−0.003	−0.054***	−0.042**	−0.004	−0.021**	−0.021**
	(0.006)	(0.016)	(0.017)	(0.004)	(0.010)	(0.009)
$EA \times PressConstraints_b^2$	–	0.006***	0.004**	–	0.002*	0.002*
		(0.002)	(0.002)		(0.001)	(0.001)
$EA \times \text{Polity}_t$	0.005*	0.006**	0.004*	0.001	0.002	0.001
	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)
$PressConstraints_b$	0.034	0.234***	0.208**	0.016	0.087*	0.080*
	(0.023)	(0.078)	(0.081)	(0.015)	(0.050)	(0.042)
$PressConstraints_b^2$	–	−0.023***	−0.018**	–	−0.008	−0.008*
		(0.008)	(0.008)		(0.005)	(0.004)
Polity_t	−0.018	−0.019	0.012	0.026	0.025	0.041**
	(0.024)	(0.024)	(0.023)	(0.024)	(0.024)	(0.019)
Age_t	0.065***	0.062**	0.052**	0.031*	0.031*	0.021
	(0.025)	(0.024)	(0.023)	(0.017)	(0.017)	(0.017)
Age_t^2	−0.001**	−0.001*	−0.001	−0.0003	−0.0003	−0.0001
	(0.0004)	(0.0004)	(0.0004)	(0.0003)	(0.0003)	(0.0003)
$IncomeLevel_t$	0.034***	0.035***	0.029***	−0.001	−0.001	0.001
	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.005)
Student_t	–	–	0.138***	–	–	0.158***
			(0.051)			(0.048)
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Non-students only?	Yes	Yes	No	Yes	Yes	No
Exclude immigrants?	Yes	Yes	Yes	Yes	Yes	Yes
Exclude noncitizens?	Yes	Yes	Yes	Yes	Yes	Yes
Base age	22	22	22	22	22	22
Observations	38167	38167	49263	35951	35951	46336
Countries	65	65	70	65	65	70
Adj. R^2	0.13	0.13	0.12	0.23	0.23	0.23

Standard errors are clustered by country, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. Base year (b) is derived from the survey year (t), minus the respondent's age (Age_t), plus 22. All regressions control for sex, relationship status, number of children, and employment status.

Table C.9: Model 2, evaluating immigrants and noncitizens

	Is migration correlated with propaganda or educational attainment?							
Dependent variable:	Press constraints				Educational attainment			
	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Immigrant	–	-0.036 (0.026)	–	–	–	-0.071 (0.108)	–	–
Non-citizen	–	–	-0.025 (0.035)	–	–	–	-0.358^{**} (0.157)	–
International	–	–	–	-0.062 (0.042)	–	–	–	-0.448^{**} (0.172)
Polity_t	-0.216^{***} (0.024)	-0.141^{***} (0.017)	-0.137^{***} (0.017)	-0.136^{***} (0.017)	-0.335^{***} (0.012)	-0.183^{***} (0.024)	-0.178^{***} (0.030)	-0.173^{***} (0.030)
Age_t	-0.047 (0.050)	-0.199^{**} (0.085)	-0.222^{**} (0.091)	-0.227^{**} (0.095)	0.390^{***} (0.034)	0.411^{***} (0.035)	0.394^{***} (0.041)	0.387^{***} (0.043)
Age_t^2	0.001 (0.001)	0.004^{**} (0.002)	0.004^{**} (0.002)	0.004^{**} (0.002)	-0.006^{***} (0.001)	-0.006^{***} (0.001)	-0.006^{***} (0.001)	-0.006^{***} (0.001)
$IncomeLevel_t$	0.003 (0.002)	0.003 (0.003)	0.002 (0.003)	0.003 (0.003)	0.190^{***} (0.010)	0.159^{***} (0.011)	0.164^{***} (0.012)	0.164^{***} (0.013)
Female	-0.0003 (0.011)	-0.019 (0.015)	-0.024 (0.018)	-0.024 (0.018)	0.146^{***} (0.052)	0.169^{***} (0.059)	0.164^{**} (0.069)	0.154^{**} (0.072)
Relationship Status	0.004 (0.006)	0.013 (0.009)	0.013 (0.009)	0.016^{*} (0.009)	-0.106^{***} (0.025)	-0.045^{**} (0.018)	-0.030 (0.019)	-0.032 (0.020)
Number of Children	-0.001 (0.012)	-0.007 (0.019)	-0.006 (0.020)	-0.008 (0.020)	-0.220^{***} (0.035)	-0.269^{***} (0.029)	-0.280^{***} (0.032)	-0.274^{***} (0.033)
Employed	0.014 (0.020)	0.019 (0.028)	0.018 (0.030)	0.017 (0.031)	0.468^{***} (0.066)	0.480^{***} (0.061)	0.507^{***} (0.070)	0.504^{***} (0.071)
Student	-0.020 (0.020)	-0.070^{*} (0.036)	-0.112^{**} (0.044)	-0.116^{**} (0.046)	1.256^{***} (0.120)	1.339^{***} (0.115)	1.327^{***} (0.152)	1.290^{***} (0.153)
Country-year F.E.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exclude immigrants?	Yes	No	No	No	Yes	No	No	No
Exclude non-citizens?	Yes	No	No	No	Yes	No	No	No
Base age (b)	18	18	18	18	N/A	N/A	N/A	N/A
Observations	52551	26565	23624	22674	66327	33026	26698	25543
Countries	70	53	39	37	70	53	39	37
Adj. R^2	0.87	0.87	0.87	0.85	0.31	0.28	0.29	0.28

Standard errors are clustered by country, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. International is a dummy that takes a value of 1 only if immigrant and non-citizen both take values of 1.

Table C.9: Model 2, evaluating immigrants and noncitizens (II)

	Does migration affect the mapping from schooling to participation?					
Dependent variable:	Passive participation					
Immigrant	−0.102* (0.059)	−0.018 (0.145)	–	–	–	–
Noncitizen	–	–	−0.210*** (0.070)	−0.388*** (0.138)	–	
International	–	–	–	–	−0.305** (0.114)	−0.496** (0.204)
EducationAttainment	0.074*** (0.014)	0.075*** (0.015)	0.074*** (0.017)	0.073*** (0.017)	0.067*** (0.016)	0.067*** (0.016)
EA×Immigrant	–	−0.018 (0.024)	–	–	–	–
EA×Noncitizen	–	–	–	0.043 (0.030)	–	
EA×International	–	–	–	–	–	0.045 (0.040)
$IncomeLevel_t$	0.026*** (0.007)	0.026*** (0.007)	0.034*** (0.009)	0.034*** (0.009)	0.034*** (0.009)	0.034*** (0.009)
Polity_t	−0.282*** (0.011)	−0.282*** (0.011)	−0.276*** (0.013)	−0.276*** (0.013)	−0.275*** (0.013)	−0.275*** (0.013)
Age_t	0.064*** (0.019)	0.064*** (0.019)	0.049** (0.023)	0.049** (0.023)	0.050** (0.024)	0.050** (0.024)
Age_t^2	−0.001*** (0.0003)	−0.001*** (0.0003)	−0.001 (0.0004)	−0.001 (0.0004)	−0.001* (0.0004)	−0.001* (0.0004)
Sex	−0.289*** (0.036)	−0.289*** (0.036)	−0.303*** (0.042)	−0.303*** (0.042)	−0.300*** (0.044)	−0.300*** (0.044)
Relationship Status	0.013 (0.012)	0.013 (0.012)	0.020 (0.014)	0.020 (0.014)	0.021 (0.014)	0.020 (0.014)
Number of Children	0.0004 (0.012)	0.0004 (0.012)	0.001 (0.012)	0.002 (0.012)	0.001 (0.013)	0.001 (0.013)
Employed	0.006 (0.028)	0.006 (0.028)	−0.008 (0.035)	−0.008 (0.035)	−0.011 (0.035)	−0.011 (0.035)
Student	0.176*** (0.045)	0.176*** (0.045)	0.138** (0.054)	0.138** (0.054)	0.133** (0.057)	0.133** (0.057)
Country-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	32326	32326	26374	26374	25237	25237
Countries	53	53	39	39	37	37
Adj. R^2	0.10	0.10	0.10	0.10	0.10	0.10

Standard errors are clustered by country, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. International is a dummy with a value of 1 only if immigrant and noncitizen both take values of 1.

Table C.9: Model 2, evaluating immigrants and noncitizens (III)

	Does migration affect the mapping from schooling to participation?					
Dependent variable:	Active participation					
Immigrant	−0.175** (0.068)	−0.108 (0.136)	–	–	–	–
Noncitizen	–	–	−0.268*** (0.086)	−0.370** (0.175)	–	
International	–	–	–	–	−0.299** (0.125)	−0.387* (0.206)
$EducationAttainment$	0.108*** (0.010)	0.108*** (0.010)	0.110*** (0.011)	0.109*** (0.011)	0.108*** (0.011)	0.108*** (0.011)
$EA\times$Immigrant	–	−0.015 (0.022)	–	–	–	–
$EA\times$Noncitizen	–	–	–	0.025 (0.027)	–	
$EA\times$International	–	–	–	–	–	0.021 (0.032)
$IncomeLevel_t$	0.005 (0.006)	0.005 (0.006)	0.001 (0.008)	0.001 (0.008)	0.002 (0.008)	0.002 (0.008)
Polity$_t$	−0.050*** (0.009)	−0.050*** (0.009)	−0.049*** (0.009)	−0.048*** (0.009)	−0.048*** (0.010)	−0.048*** (0.009)
Age$_t$	0.004 (0.012)	0.004 (0.012)	−0.007 (0.012)	−0.007 (0.012)	−0.008 (0.013)	−0.008 (0.013)
Age$_t^2$	0.0001 (0.0002)	0.0001 (0.0002)	0.0002 (0.0002)	0.0002 (0.0002)	0.0003 (0.0002)	0.0003 (0.0002)
Sex	−0.094*** (0.025)	−0.094*** (0.025)	−0.067** (0.029)	−0.067** (0.029)	−0.076** (0.029)	−0.076** (0.029)
Relationship Status	−0.027*** (0.007)	−0.027*** (0.007)	−0.021*** (0.008)	−0.022*** (0.008)	−0.022** (0.008)	−0.022** (0.008)
Number of Children	0.008 (0.007)	0.008 (0.007)	0.004 (0.007)	0.004 (0.007)	0.004 (0.007)	0.004 (0.007)
Employed	0.102*** (0.024)	0.102*** (0.024)	0.123*** (0.028)	0.123*** (0.028)	0.121*** (0.029)	0.121*** (0.029)
Student	0.169*** (0.050)	0.169*** (0.050)	0.174*** (0.062)	0.174*** (0.062)	0.165** (0.062)	0.165** (0.062)
Country-year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	30100	30100	25530	25530	24407	24407
Countries	52	52	39	39	37	37
Adj. R^2	0.22	0.22	0.23	0.23	0.20	0.20

Standard errors are clustered by country, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. International is a dummy with a value of 1 only if immigrant and noncitizen both take values of 1.

Table C.10: Model 2 with standardized explanatory variables

	Press constraints and political participation							
Dependent variable:	Passive participation				Active participation			
	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)	(4a)	(4b)
$EducationAttainment$	0.183***	0.180***	0.138***	0.134***	0.225***	0.223***	0.212***	0.198***
	(0.024)	(0.024)	(0.029)	(0.024)	(0.013)	(0.013)	(0.017)	(0.017)
$EA \times Constraints_b$	−0.048**	−0.032	−0.040*	−0.044**	−0.034***	−0.023	−0.025	−0.018
	(0.020)	(0.022)	(0.021)	(0.020)	(0.013)	(0.017)	(0.017)	(0.016)
$EA \times Constraints_b^2$	–	–	0.043***	0.038***	–	–	0.012	0.014*
			(0.015)	(0.011)			(0.009)	(0.008)
$EA \times \text{Polity}_t$	–	0.030	0.033	0.035**	–	0.022	0.023	0.023
		(0.024)	(0.023)	(0.017)		(0.026)	(0.026)	(0.019)
$PressConstraints_b$	−0.056	−0.052	−0.029	0.010	−0.003	−0.001	−0.008	0.019
	(0.035)	(0.035)	(0.037)	(0.038)	(0.018)	(0.017)	(0.023)	(0.021)
$PressConstraints_b^2$	–	–	−0.015	−0.012	–	–	−0.007	0.009
			(0.021)	(0.016)			(0.012)	(0.009)
Polity_t	0.043	0.047	0.055	0.167*	0.138	0.141	0.145	0.202***
	(0.107)	(0.106)	(0.108)	(0.084)	(0.103)	(0.104)	(0.103)	(0.072)
Age_t	0.133***	0.135***	0.133***	0.130***	0.078***	0.079***	0.079***	0.070***
	(0.025)	(0.025)	(0.025)	(0.023)	(0.013)	(0.013)	(0.013)	(0.015)
Age_t^2	−0.023*	−0.023*	−0.022*	−0.022*	0.003	0.003	0.003	0.004
	(0.013)	(0.013)	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)
$IncomeLevel_t$	0.071***	0.070***	0.071***	0.056***	0.005	0.004	0.005	0.006
	(0.015)	(0.015)	(0.015)	(0.012)	(0.015)	(0.015)	(0.015)	(0.011)
Student_t	–	–	–	0.158***	–	–	–	0.192***
				(0.030)				(0.034)
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Non-students only?	Yes	Yes	Yes	No	Yes	Yes	Yes	No
Exclude immigrants?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exclude non-citizens?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	32432	32432	32432	49552	30546	30546	30546	46492
Countries	65	65	65	70	65	65	65	70
Adj. R^2	0.13	0.13	0.13	0.11	0.23	0.23	0.23	0.22

Standard errors are clustered by country, with ***, **, and * denoting significance at the 1%, 5%, and 10% levels, respectively. Base year (b) is derived from the survey year (t), minus the respondent's age (Age_t), plus 18. All regressions control for sex, relationship status, number of children, and employment status. Coefficients and standard errors for all non-binary explanatory variables are standardized with mean zero.

Variable descriptions

- **PressConstraints**: from Freedom House's (FH) Freedom of the Press index (1997-2014), rescaled to be from 0 to 10.

- **PressConstraints$_b$**: same as above, with a countries' value (1993-2014) assigned to a WVS respondent's base age (i.e. 18 or 22). This was done by

  1. Determining the amount of years that have passed since a respondent was base age, given a respondent's age at time of survey (or removing from sample if missing an age).
  2. Subtracting that amount of time from the survey year to derive the year in which that respondent was base age.
  3. Assigning the country-year Freedom Score measures to those years (or removing from sample if pre-1993).

  Note: Since North Macedonia is missing a score for 1993 but otherwise had quite stable scores, its 1994 value was used for 1993.

- **Crisis**: derived from Polity IV's POLITY2 index; I give a value of 1 for any year it denoted a country as being engaged in "interregnum" or "transition"; a value of 1 for the respective years before and after any given "interregnum," "transition," or "foreign interruption" began; and a value of 1 for any year immediately preceding or following a change in its POLITY2 score of a magnitude larger than 5. Note that POLITY2 leaves years during a "foreign interruption" as missing values.

- **EducationAttainment (EA)**: from the World Values Survey's (WVS) X025 (1994-2014), meaning highest education level, where 0 =incomplete primary, 1 =complete primary, 2 = incomplete vocational secondary, 3 =complete vocational secondary, 4 =incomplete preparatory secondary, 5 =complete preparatory secondary, 6 =some university, 7 =full university.

- **Education completion age**: from WVS's X023 (1994-2014), meaning *expected* completion age, ranging from 1 to 99, with negative values treated as missing. See comment on "non-students only" below for relevance.

- **Education expenditures (% gov) and lnEducation**: from the World Bank's (WB) World Development Indicators (WBI) database (1997-2014). Logged for normality.

- **Employment and student status**: derived from WVS's X028 (1994-2014), meaning current employment status. Missing observations are treated as missing for all variables. Those who answered "full time," "part time," or "self-employed" given a value of 1 for the employed variable; those who answered "student" given a 1 for the student variable.

  - Some who identified as students marked a younger-than-current-age education completion age, while some who answered that they expect to complete their education at an older-than-current-age did not identity as students.

    - In pursuit of the most possible certainty, regressions marked yes to "non-students only" include only those who said they were not a student *and* that they were at the age of or older than their education completion age. Regressions marked no to "non-students only" are all-inclusive with respect to these variables.

- **GDP per capita ($ 2010) and lnGDPpc**: from the WB WDI (1997-2014). Logged for normality.

- **Gov't expenditures (% GDP) and lnGovExpend**: from the WB WDI database (1997-2014), measuring total government consumption expenditures. Logged for normality.

- **Immigrant, noncitizen, and international**: all derived from WVS (1994-2014).

    - *Immigrant*: derived from G017 ("born in this country"), G018 ("when came to country"), and G027A ("immigrant"). Anyone who answered yes on G017, anything on G018, and/or yes on G027A was given a value of 1 for the immigrant variable. Missing observations remained missing. A small number of observations marked that they were both an immigrant and born in their current country; I marked those as missing, given the indeterminacy.
    - *Noncitizen*: derived from G005 ("citizen of country") and G027B ("citizen"). Anyone who answered yes to either was given a value of 1 for the noncitizen variable. Missing observations remained missing.
    - *International*: designed to denote individuals who likely were not educated in their current country. Anyone given a 1 for both the immigrant and noncitizen variables was given a 1. Missing observations for either variable are considered missing for the international variable. However, in pursuit of most possible certainty, main regressions exclude all immigrants *and* noncitizens.

- **$\mathbf{IncomeLevel}_t$**: from WVS's X047 (1994-2014), denoting income deciles from 0 to 9 (i.e. low to high), with negative values treated as missing.

- **Military expenditures (% GDP), lnMilitary, and $\mathbf{lnMilitary}_2$**: : from the WB WDI database (1997-2014). Logged for normality. $lnMilitary_2$ corrects for exacerbation or deletion of zero or near-zero expenditure values post-logarithm by adding 1 prior.

- **Other individual-level data ($\mathbf{Age}_t$, sex, number of children, and relationship status)**: from WVS's X003, X001, X011, and X007 (1994-2014), respectively.

    - All missing or negative Age_t observations removed from sample, as it is a key variable. Because of the necessary lower bounds on observing propaganda and data limitations going further back in time, ages range from 18 to 43.
    - Individuals are given a 1 if female and 0 if male, with negative values treated as missing.

  - Number of children ranges from 0 to 8 (i.e. 8 or more), with negative values treated as missing.
  - A relationship variable was derived from X007, wherein individuals were given a 0 if they answered "single/never married"; a 1 if they answered "divorced," "separated," or "widowed"; a 2 if they answered "living together as married"; and a 3 if they answered "married." Negative values treated as missing.

- **Participation (active)**: composite measure of WVS's E025 and E027 (1994-2014), in which individuals are given a 0 if they answer that they "would never" sign a petition or attend a lawful demonstration, a 1 if they answer that they "might," and a 2 if they answer that they "have," for a total score ranging from 0 to 4. Negative or missing observations removed from sample.

- **Participation (passive)**: composite measure of WVS's A004 and E023 (1994-2014), in which individuals are given a 0 if they state that politics is "not at all" important or interesting, a 1 if they answer "not very" important or interesting, a 2 if they answer "somewhat" important or interesting, and 3 if they answer "very" important or interesting, for a total score ranging from 0 to 6. Negative or missing observations removed from sample.

- **Polity/Polity$_t$**: from Polity IV's POLITY2 index measuring net democracy, ranging from -10 to 10, with 10 being the most democratic (on net).

- **AcademicConstraints**: equals 1 if a country's scholars have ever faced imprisonment, prosecution, or travel restrictions in the Scholars at Risk (SAR) Network's Incident Index from 2011 through October 2017.

Countries: model 1 (1997-2014)*

Albania; Angola; **Argentina**; **Armenia**; Australia; **Austria**; **Azerbaijan**; Bahrain; **Bangladesh**; **Belarus**; **Belgium**; **Benin**; **Bhutan**; **Bolivia**; Botswana; **Brazil**; **Bulgaria**; Burkina Faso; **Burundi**; Cambodia; **Cameroon**; **Canada**; **Cape Verde**; **Central African Republic**; **Chad**; **Chile**; China; **Colombia**; Comoros; Costa Rica; Croatia; **Cyprus**; **Czech Republic**; Democratic Republic of the Congo; **Denmark**; **Djibouti**; Dominican Republic; Ecuador; Egypt; **El Salvador**; Eritrea; **Estonia**; Ethiopia; **Fiji**; **Finland**; **France**; Gabon; **Gambia**; **Georgia**; Germany; **Ghana**; Greece; Guatemala; **Guinea**; Guinea-Bissau; **Guyana**; **Hungary**; **India**; **Indonesia**; **Iran**; **Ireland**; **Israel**; **Italy**; **Jamaica**; **Japan**; Jordan; Kazakhstan; Kenya; Kuwait; **Kyrgyzstan**; **Laos**; **Latvia**; **Lebanon**; Lesotho; Liberia; Lithuania; **Madagascar**; Malawi; **Malaysia**; **Mali**; Mauritania; **Mauritius**; **Mexico**; **Moldova**; **Mongolia**; Morocco; Mozambique; Namibia; **Nepal**; **Netherlands**; **New Zealand**; Nicaragua; **Niger**; **Norway**; **Oman**; **Pakistan**; Panama; **Paraguay**; **Peru**; **Philippines**; **Poland**; **Portugal**; **Qatar**; Republic of Congo; **Romania**; Russia; Rwanda; **Saudi Arabia**; **Senegal**; **Sierra Leone**; **Singapore**; **Slovakia**; **Slovenia**; Solomon Islands; **South Africa**; **Spain**; Sri Lanka; **Swaziland**; **Sweden**; Switzerland; **Tajikistan**; Tanzania; **Thailand**; **Togo**; Trinidad and Tobago; **Tunisia**; Turkey; Uganda; **Ukraine**; **United Kingdom**; **United States**; Uruguay; Venezuela; Vietnam; Yemen; Zambia

Countries: model 2 (1994-2014)**

Albania (1998, 2002); Algeria (2002, 2013); Argentina (1995, 1999); Australia (1995, 2005, 2012); Bangladesh (1996, 2002); Armenia (1997, 2011); Brazil (2006, 2014); Bulgaria (1997, 2005); Burkina Faso (2007); Canada (2000, 2006); Chile (1996, 2000, 2006, 2011); Colombia (1998, 2012); Cyprus (2006, 2011); Czech Republic (1998); Dominican Republic (1996); Ecuador (2013); El Salvator (1999); Ethiopia (2007); Estonia (1996, 2011); Finland (1996, 2005); France (2006); Georgia (2009, 2014); Germany (1997, 2006, 2013); Ghana (2007, 2012); Guatemala (2004); Hungary (2009); India (1995, 2001, 2006, 2014); Indonesia (2001, 2006); Italy (2005); Japan (2000, 2005, 2010); Lebanon (2013); Kyrgyzstan (2003, 2011); Latvia (1996); Lithuania (1997); North Macedonia (1998); Malaysia (2006, 2012); Mali (2007); Mexico (1995, 1996, 2000, 2005, 2012); Moldova (1996, 2002, 2006); Netherlands (2006, 2012); New Zealand (1998, 2004, 2011); Nigeria (1995, 2000, 2011); Norway (1996, 2007); Pakistan (1997, 2001, 2012); Peru (1996, 2001, 2006, 2012); Philippines (2001, 2012); Poland (2005, 2012); Romania (1998, 2005, 2012); Russia (1995, 2006, 2011); Singapore (2002, 2012); Slovakia (1998); Slovenia (2005, 2011); South Africa (1996, 2001, 2006, 2013); South Korea (1996, 2001, 2005, 2010); Spain (1995, 2000, 2007, 2011); Sweden (1996, 2006, 2011); Switzerland (1996, 2007); Taiwan (1994, 1996, 2012); Tanzania (2001); Thailand (2007, 2013); Trinidad and Tobago (2006, 2011); Turkey (1996, 2001, 2007, 2011); Uganda (2001); Ukraine (1996, 2006, 2011); United Kingdom (1998, 2005); United States (1995, 1999, 2011); Uruguay (1996, 2006, 2011); Venezuela (1996, 2000); Yemen (2014); Zambia (2007)

*Countries in both unbalanced and balanced panel specifications appear in **bold**.
**There are no observations from 2008.