Three Essays on the Economic Effects of Violent Conflicts and Culture

by

Christoph Eder

Mag., Johannes Kepler Universität, 2008

a Thesis submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

in the
Department of Economics
Faculty of Arts and Social Sciences

SIMON FRASER UNIVERSITY
Summer 2014

APPROVAL

Name: **Christoph Eder**

Degree: **Doctor of Philosophy (Economics)**

Title of Thesis: **Three Essays on the Economic Effects of Violent Conflicts and Culture**

Examining Committee:

Chair: **Fernando Aragon**
Assistant Professor, Department of Economics

Anke Kessler
Senior Supervisor
Professor, Department of Economics

Simon Woodcock
Supervisor
Associate Professor, Department of Economics

Christiaan Muris
Supervisor
Assistant Professor, Department of Economics

Dominique Gross
Internal Examiner
Professor, School of Public Policy

Nathan Nunn
External Examiner
Professor, Department of Economics
Harvard University

Date Defended/Approved: May 15, 2014

Abstract

Chapters 1 and 2 of my thesis deal with long-term economic effects of violent conflicts. In the first chapter, I use World War II casualties suffered in Austrian municipalities as a natural experiment for human casualties and find a significant negative causal effect of human losses on economic activity today. As I demonstrate, the likely channel through which the effect persisted over time is through its impact on the structural composition of the work force. Specifically, greater human losses increased the fraction of employment in manufacturing at the expense of agriculture until the 1970s and services from then onwards. A simple model shows that structural change can translate a lower labor share in agricultural production into less participation of service sector growth at a later time.

In the next chapter, I identify a channel through which the disadvantage of displacement during a violent conflict might be carried over to the next generation. In particular, I show that displaced parents spend significantly less on the education of their children years later. A decomposition of the causal effect shows that differences in income and the stock of durable goods can at most explain one third of the finding. Some evidence points towards increased uncertainty about the future of displaced parents and hence reduced spending on non-vital expenditure positions.

The final chaper revisits the paper by Algan & Cahuc (AER, 2010) in which they find that inherited trust has a large impact on GDP per capita. First, I show that the estimates presented in Algan & Cahuc might be biased due to a difference between the lag structure of inherited trust and initial income in their econometric specification. Next, I focus on their robustness checks, where I replicate their results and document that most of their robustness checks fail when a programming error and data problems are corrected. I conclude that their results should be considered with great care.

Acknowledgments

Special thanks goes to Anke Kessler, who supported me throughout the years with helpful discussion, patient guidance, and encouragement. She never shied away from communicating her honest judgment and was always prepared to contribute with her expertise.

The first chapter benefited also from detailed discussion with Fernando Aragón, Michele Battisti, Chris Muris, Simon Woodcock, and participants of various SFU seminars and the JKU lunch time seminar.

Chapter 2 improved greatly from comments of Gustavo Bobonis, Alex Karaivanov, Krishna Pendakur, and conference participants at the Atlantic Summer Conference on Public Policy Research, the Development Conference of the German Economic Association, the conference of the European Association of Labour Economists, seminar participants at Johannes Kepler University, and three anonymous referees.

Michele Battisti, Geoffrey Dunbar, Brian Krauth, Chris Muris, Krishna Pendakur and participants at the SFU Brown Bag seminar contributed to the final chapter through valuable comments.

All remaining errors are mine.

Contents

List of Tables

List of Figures

Chapter 1

Missing Men: World War II Casualties and Structural Change

1.1 Introduction

World War II (WWII) caused enormous losses of human life in many countries, with an estimated death count of over 60 million people (Beevor 2012). What is the effect on long-term economic development of fatalities in an event like this? Even though WWII was the largest shock to human population in history, we know relatively little of its economic impact over time in highly affected areas. This is at least partly due to the difficulty in answering this question credibly via a country level analysis: decisions to enter the war, being attacked, the extent of the damage, or the degree of resistance are likely correlated with country characteristics to which subsequent economic performance is endogenous. Hence, such an estimate of the consequences of war casualties could not be interpreted in a causal way.

In this paper I address the question of long-term effects of human loss of life due to WWII and shed light on the underlying channel of persistence in the post-war period. To this end, I exploit WWII as a natural experiment for military war casualties at the municipality level in Austria. The municipalities I consider share the same history and institutions and exhibit a large degree of ethnic, linguistic, cultural, geographic and economic homogeneity. Moreover, conscription and location of deployment of Austrian men was based on birth cohort rather than municipality characteristics. Thus, the relative number of casualties that a municipality suffered during the war was arguably random.

To conduct the analysis, I employ a novel data set on war casualties that I have collected. I find a strong and robust negative long-term effect of casualties on subsequent economic output, measured by the total wage bill, in the affected municipalities today. Within-district estimates rule out many political and geographic confounding factors. The results prove robust to the inclusion of various control variables. Falsification tests show that high-casualty municipalities did not experience different economic development before the war, strengthening the causal interpretation of the result. To understand the mechanism through which war casualties cause the difference in the output measure, I next take a closer look at the various determinants of economic activity at the municipality level. I find that the primary difference between high- and low-casualty municipalities today is the number and density of firms and a higher number of in-commuters. Interestingly, war casualties only affect the firm environment in the service sector, however, not in the manufacturing sector.

As a second step, I identify a possible channel of persistence, using municipality-level census data since the 1950s. In particular, an analysis of the sector composition of the labor market reveals that the transition from agriculture to the service sector followed differential paths depending on the share of dead WWII combatants. After the war more people are employed in the manufacturing sector in high-casualty municipalities, while fewer work in agriculture until the 1970s and fewer in the service sector since then. The initial sorting to manufacturing in highly affected municipalities is consistent with labor shortage and hence productivity reduction experienced in agriculture due to the perished male household heads.

I develop a simple model of structural change where WWII casualties reduce productivity in agriculture. High-casualty municipalities sort into manufacturing while low-casualty municipalities are in agriculture to exploit the productivity advantage. Income growth through an increase in total factor productivity shifts consumer demand away from agricultural products towards service sector goods. Marginal transition costs to switch between sectors ensure that the labor demand in services is satisfied from the agricultural sector. The low-casualty municipalities are therefore over-represented in the service sector. When the service sector grew substantially since the 1970s, people already working in the service sector established firms in their home municipalities, and hence the output difference today. To my knowledge, a labor shortage in agriculture and a consequently different path of structural change has not been identified as a link between violent conflict and long-term economic development.

The findings in this paper relate to a broader literature on the long-term effects of historic events. This literature was recently surveyed by Nunn (2013), who addresses a wide array of historic events and mechanisms underlying historic persistence. This paper is most closely related to a contribution by Acemoglu et al. (2011) on the effects of the Holocaust in Russia on subsequent economic and political development. They document a negative relationship between the reduction of the Jewish population and subsequent population growth, wages and political development. However, Acemoglu et al. (2011) focus on the reduction of a narrow subgroup of the population, a feature I extend upon by looking at a lack of men of all social classes.

A number of other papers study the persistent effects of forced population movements through channels different from structural change. Nunn (2008) focuses on slave extraction in Africa and shows a large and negative effect on today's economic development. Nunn and Wantchekon (2011) show that persistence can emerge through a culture of mistrust generated through the slave trades. Dell (2010) studies the mining *mita*, an extensive forced labor system established in the 16th century in Peru, and finds negative long-term effects through fewer formations of *haciendas*, large land holdings with an attached labor force, which in turn were responsible for providing education and the road network. A number of contributions have identified labor shortages in agriculture as the driving force behind adjustment processes to historic events. Chaney and Hornbeck (2013) find increased per capita output due to the adaptation of different agricultural technology in response to the 1609 expulsion of Moriscos from Spain. A more recent event, the 1927 Great Mississippi Flood, caused black out-migration and subsequent increased capital intensity in agriculture in affected areas (Hornbeck, Naidu (2014)). Conversely, I find that in times of rapid structural change labor shortage in agriculture can have a detrimental effect on long-run development.

Finally, the literature on the long-run consequences of wars include Davis and Weinstein (2002 and 2008), Brakman, Garretsen, and Schramm (2004), and Miguel and Roland (2011). These papers analyze bombings of Japanese or German cities in WWII and bombings during the Vietnam War, respectively, and find no long-term effects on a range of outcome measures. My findings stand in contrast to the results of this line of research, which suggests the relative importance of human versus physical capital.

The paper is organized as follows: Section 2 gives an overview of the history of Austria in WWII and Austrian municipalities in general. In Section 3 the identification strategy and the data are described, while Section 4 gives the main empirical results on the relationship between output and WWII casualties and performs robustness checks and falsification tests. The next section explores the determinants of the estimated difference in output, Section 6 looks at the process of divergence since the 1950s, while Section 7 concludes.

1.2 World War II and Austrian Municipalities

1.2.1 Austria and World War II

The annexation of Austria through Germany in March 1938, generally referred to as *Anschluss*, meant the temporary disappearance of Austria as a nation state. The union with Germany, which was preparing for war, brought both positive and negative economic shocks, but led to a great reduction of unemployment in Austria (Thalmann 1954, p.500).

When WWII broke out in the summer of 1939, Austrian forces were already fully integrated in the German army (Morawek, Neugebauer 1989, p.43). On June 15, 1938 compulsory military service was introduced and 95,000 Austrian men were drafted before the outbreak of war. Austrian soldiers were recruited into two regional divisions, which were employed in all major attacks throughout the war side by side with German forces. Over time, additional cohorts were drafted at a less than yearly intervals (Morawek, Neugebauer 1989, p.43f). Few cases of refusal to be drafted are known among the Austrian army (Morawek, Neugebauer 1989, p.49).

Over the course of the war, an estimated 1.2 million Austrian men were drafted, of which about 250,000 died (Hagspiel 1995, p.329). More than 100,000 returned with severe injuries, and estimates of Austrian prisoners of war are between 500,000 and 600,000 (Vocelka 2010, p.110; Hagspiel 1995, p.329). About 24,300 civilians died due to air raids (Bukey 2000, p.227). Nevertheless, the population of Austria increased from 6.7 million to 7 million between 1938 and 1945, mostly due to increased birth rates before and during the first years of the war (Hagspiel 1995, p.57f).

Just how large the demographic impact of the two world wars is, shows Figure 1.1, which depicts cohort sizes of males (solid line) and females (dashed line) by birth cohort in 1910, 1934, 1951 and 2001. The horizontal axis measures the birth year, while on the vertical axis the number of people by birth cohort is shown. The grey bars mark the years of the first and second world war. The panel for 1910 on the top left shows almost identical profiles of men and women, while in 1934 the effect of WW1 is clearly visible. The fallen men of the cohorts 1870-1900 and the reduction in births during the war left its marks. After WWII there are additional cohorts of men missing from the cohorts 1900-1930 and the sharp upward and then downward spikes in births during the war years are visible. The data from the 2001 census still show a large difference between men and women for older cohorts.

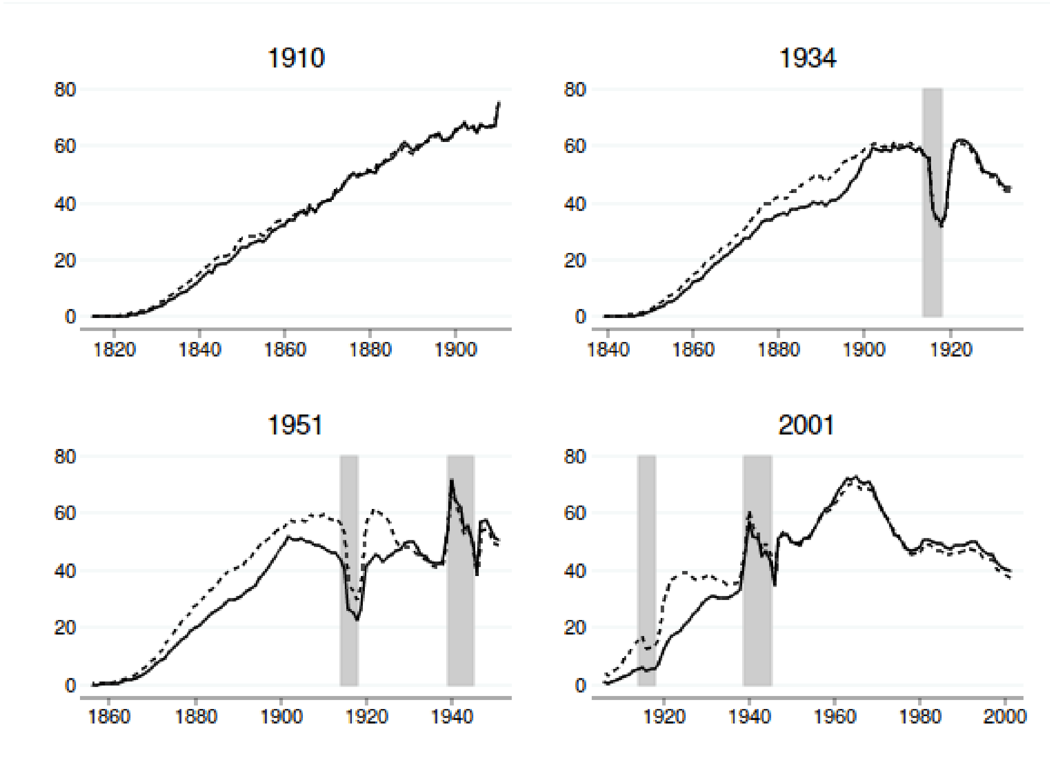
Already before the end of the war, Austria was reinstated as a sovereign state. However, Austria was occupied until 1955 by the four Allied forces, who split up the territory along provincial borders. The main problems of that time were food shortages and lack of intact physical capital. The Austrian government was required to implement a series of long-term investment projects. By 1950, the Austrian GDP reached twice its 1946 level and exceeded the 1937 level by about 10% (Thalmann 1954, p.503f). The Austrian economy grew at an average annual rate of around 2.5% in the following decades.

1.2.2 Austrian Municipalities

Austria's current population is 8.4 million. It is a federal democracy organized in its nine provinces, each with their own parliament. The provinces are subdivided into 98 districts, which do not have any elected body. The lowest administrative units are the 2,357 municipalities. Municipality sizes vary greatly from 53 people in Gramais, Tyrol, to Vienna with a population of 1.7 million.

Originally founded in 1849 as an administrative unit to replace the feudal system, the municipalities have always been the government offices with the most contact to citizens. The municipalities have an elected mayor and municipal council, which is elected every 5 or 6 years, depending on the province. The responsibilities of a municipal government include land use planning, energy and water supply, provision of schools and homes for the elderly etc. (Bauer, Paleczny, Schulmeister 1977, p.65f).

Figure 1.1: Male and Female Population by Cohort in Austria



Birth cohorts are measured on the horizontal axis and the cohort size is on the vertical axis. The headline of each diagram reports the time of measurement. The male population is drawn with the solid line, while the dashed line marks the female population. Data come from the Austrian population censuses.

1.3 Identification and Data

This section will shortly introduce the econometric framework employed and discuss the necessary assumptions to estimate the causal effect of WWII casualties on economic outcomes. A description of the data and the employed sample selection follows.

1.3.1 Identification

In general, the estimated equation is of the following form:

$$\log Y_{i,j} = \beta S_{i,j}^{WW2} + X_{i,j}\delta + \mu_j + \varepsilon_{i,j}, \tag{1.1}$$

where $Y_{i,j}$ is the outcome variable in municipality i and district j. $S_{i,j}^{WW2}$ is the measure of WWII casualties, $X_{i,j}$ is a set of municipality-level covariates which includes a dummy variable of the market status of the municipality in 1945, the log. of population in 1939, the share of the population employed in agriculture in 1934, and the share of men in 1934. μ_j is a set of district fixed effects and $\varepsilon_{i,j}$ is an error term.

The main problem in this general specification is a possible correlation of $S_{i,j}^{WW2}$ and $\varepsilon_{i,j}$ that would bias the OLS estimate of β in (1.1). An example for this is the Allied advance in Austria towards the end of the war. Air raids and the progress of ground forces claimed civilian lives in many municipalities, while these actions possibly also had effects on future economic growth. Consider Allied air bombings that killed civilians and destroyed physical capital at the same time or the well documented harshness of Russian forces in the East on civilians, who at the same time dismantled machinery to be transported to the Soviet Union (Bukey 2000, p.216f; Hagspiel 1995, p.91f).

A way around this problem is the use of a subset of casualties in the variable $S_{i,j}^{WW2}$ that are actually orthogonal to the error term $\varepsilon_{i,j}$. In this study, war casualties are measured as the the number of men who served in the military and did not return from war. These soldiers either died in battle, as prisoners of war in a detention camp, of a disease, or due to an accident. Importantly, they did not die in their home town, but rather, as faraway as the Soviet Union, France or Africa.

Austrian men were drafted by cohort and employed alongside German soldiers in all major battles. Initially, men could get an exemption from military service if they could prove their importance in agriculture or industrial production. However, by the time the attack against the Soviet Union started in the summer of 1941, most exemptions were canceled and additional older and younger cohorts were recruited (Morawek, Neugebauer 1989, p.46f). At that time millions of prisoners of war had been brought into Germany and Austria to reduce labor shortages in agriculture and industrial production (Hagspiel 1995, p.66f). Finally, with the formation of the *Volkssturm* in September 1944 all men of age 16 to 60 years were drafted and deployed in the defense against approaching Allied forces. In February 1945 the cohort of 1929 was drafted (Hagspiel 1995, p.84f). Many men found their death in these last months of WWII.

The employed variable of WWII casualties is defined as:

$$S_{i,j}^{WW2} = \frac{DEAD\_SOLDIERS_{i,j}}{POPULATION\_1939_{i,j}}, \tag{1.2}$$

where $DEAD\_SOLDIERS_{i,j}$ measures the number of dead soldiers with municipality of residence i in district j before the war and $POPULATION\_1939_{i,j}$ is the total population of the municipality as counted in the census of May 1939.

In conclusion, I argue that almost the entire male population of certain birth years was drafted to serve in WWII. The municipality of residence did not play a role in the drafting process and location of deployment and consequently the probability of death is uncorrelated with pre-war municipality characteristics. I come back to this claim empirically after the data description.

1.3.2 Data

The data for this paper come mostly from two sources. The number of dead WWII combatants was collected from war memorials in each municipality. The outcome data are provided by *Statistik Austria*, the federal statistical agency, in various publications.

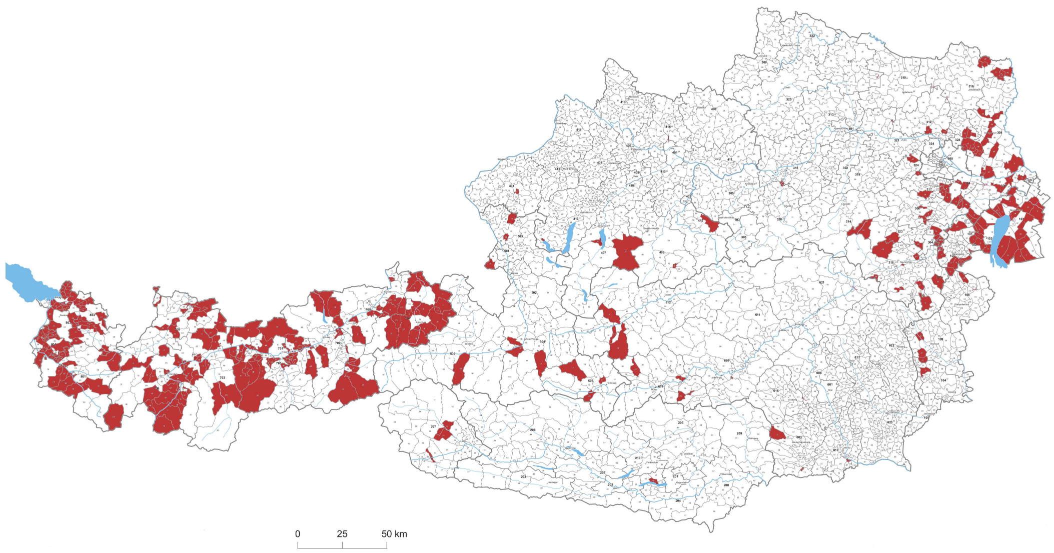
WWII Casualties

As described above, the war casualty variable is measured as the share of the municipality's 1939 population who served in the military and did not return from war. To the best of my knowledge, a death registry for soldiers of Nazi-Germany is not available with the location of residence before the war.

Given the lack of a data set of war casualties, I take another avenue to get data. Most municipalities have constructed a memorial to remember their dead and missing soldiers, of which most list the names of the dead and sometimes also date/location of birth and/or death. I use the number of names from these war memorials as the source of the casualty data. Various websites for genealogist provide photographs and transcripts of these memorials from many places in Austria.[1]

[1] The websites I have used are `www.denkmalprojekt.org` and `www.kriegerdenkmal.co.at`

Figure 1.2: Sample Municipalities



The municipalities in the sample are mostly from four provinces: Tyrol (38%), Vorarlberg (21%), Niederösterreich (20%), and Burgenland (12%).

Two issues arise from this approach. First, I need to ensure that civilian casualties are not included in the list of names on the memorial for the reason discussed above. The memorials are usually divided in fallen or dead, missing and sometimes civilian casualties. I never include the civilian casualties and do not use the data from a memorial if there is a female name among the list of dead, as this indicates that civilian casualties are included. The war memorials should be a complete list of all dead and missing soldiers from a municipality, because relatives of non-returning soldiers had an interest in ensuring that their fathers, sons and husbands were honored on the memorial. Often names were added to the end of the list, when missing soldiers did not return from war captivity after a number of years. In other cases names were erased when the person unexpectedly returned. In one instance a note was added to a name when the person returned in 1958 from imprisonment.

The second issue with these data is how to link a war memorial to a municipality and avoid double counting. The municipalities are subdivided into one or more localities, which can range from a cluster of houses to a city. In most cases, the largest locality has the same name as the municipality, so it is not entirely clear if the municipality or the locality is meant by a memorial. There are also memorials that refer to the area that belongs to a rectorate of the church. To reduce measurement error in the casualty rate variable, I

Table 1.1: Descriptive Statistics

	Obs.	Mean	Std. Dev.	Min.	Max.
	(1)	(2)	(3)	(4)	(5)
Share Dead Soldiers WWII	300	0.058	0.016	0.027	0.120
Number Dead Soldiers WWII	300	66.9	57.6	5	483
Population in 1939	300	1,162.6	995.4	70	8,773
Population in 2011	300	2,146.5	2435.6	53	21,181

make three restrictions. First, I restrict attention to municipalities whose borders have not changed since 1934, since I do not know if the war memorial corresponds to the old borders of a municipality or the one after the merging. Second, I drop municipalities with city status in 1945 since those often have several memorials with a different number of names on it. Finally, I focus on municipalities that have only one locality, to ensure that the soldier casualty data correspond to the whole municipality. These restrictions reduce the number of municipalities to 300 out of 2,357 available by January 1, 2013.[2]

This sample selection leave me with soldier casualty data for 300 municipalities, which include about 20,100 dead soldiers. Table 1.1 describes the war casualty data. An average of 5.8% of the 1939 population died as soldiers, with a standard deviation of 1.6%. The total number of deaths ranges between 5 and 483, with a mean of 67 soldiers. The population size of these mostly small municipalities is on average 1,163 in the year 1939 and by 2011, the average municipality in the sample has grown to 2,147 people.

Figure 1.2 shows the geographic distribution of the sample municipalities in a map. The western provinces of Tyrol and Vorarlberg make up almost 60% of the sample while two western provinces (Niederösterreich and Burgenland) account for another 32%. The other provinces account for the rest.

[2]Although the census was in 2011, the reported results are based on municipality borders on January 1, 2013.

Economic Outcomes

The outcome variables come from detailed information on Austrian municipalities from *Statistik Austria*. The available data were constructed from the censuses between 1934 and 2001 and various other registers and surveys. Of particular interest to this project are variables about the economic activity in a municipality. Table 1.2 describes the outcome variables, while Table A.1 in the Appendix describes some additional control variables. Figure A.1 in the Appendix shows partial correlations between the main variables used in the baseline regressions.

Total Wage Bill
Of particular interest to this project is a measure of the total wage bill in a municipality. The *Kommunalsteuer* is a tax on the sum of gross salaries and wages that a firm pays to its employees and workers (including the part of social security benefits the employee/worker pays) and has a uniform rate of 3% all over Austria. The salaries and wages of all employees of a firm located within a municipality are subject to this tax. Public administration is not included, but firms owned by the government are. Given the received tax within a year, in this case 2011, one can calculate the total wage bill for each municipality. Under the assumption of a constant capital-labor productivity across municipalities and no or uniform tax evasion per municipality, the total wage bill can be interpreted as an output measure of each municipality. The total wage bill calculated in this way for all of Austria constitutes about 31% of GDP in 2011. The correlation coefficient between the total wage bill of each province and the GDP by province is 0.998 and 0.982 when measured in per capita terms. I use the total wage bill as an approximation of total output produced within the borders of a municipality.

Census Data
Population censuses since WWII are collected in 10 year intervals. The data include the industry of employment, labor market participation, commuting patterns etc. While the outcome measures since 2001 come from the *Statistik Austria* website, the older variables are taken from print copies of the census results.

The variables on commuting pattern are the number of people commuting in and out of the municipality. The share of out-commuters indicate that about 73% of the working population do not find work within their municipality of residence. There are on average

Table 1.2: Descriptive Statistics

	Obs.	Mean	Std. Dev.	Min.	Max.
	(1)	(2)	(3)	(4)	(5)
Total Wage Bill in 2011 (in EUR 1,000)	300	18106.111	30754.054	66.667	187133.33
Total Wage Bill pc in 2011 (in EUR 1,000)	300	6.973	7.573	0.212	55.134
Number of Firms 2011	300	151.983	181.157	2	1475
Number of Firms - Manufacturing Sector	300	24.003	29.006	0	309
Number of Firms - Service Sector	300	127.980	154.745	2	1166
Firm Density 2011 (per 1,000 inhabit.)	300	70.334	34.968	20.444	321.429
Share of In-Commuters 2011	300	0.375	0.374	0.000	2.978
Share of Out-Commuters 2011	300	0.728	0.127	0.168	0.900
Total Participation Rate 2011	300	0.493	0.036	0.321	0.586
Share of Employed in Agriculture 1934	300	0.524	0.234	0.022	0.932
Share of Employed in Agriculture 1951	300	0.520	0.236	0.029	0.974
Share of Employed in Agriculture 1961	300	0.386	0.220	0.018	0.875
Share of Employed in Agriculture 1971	300	0.206	0.166	0.006	0.797
Share of Employed in Agriculture 1981	300	0.118	0.110	0.000	0.603
Share of Employed in Agriculture 1991	300	0.081	0.078	0.007	0.443
Share of Employed in Agriculture 2001	300	0.055	0.061	0.000	0.510
Share of Employed in Agriculture 2011	300	0.049	0.047	0.005	0.357
Share of Employed in Manufact. Sec. 1934	300	0.251	0.163	0.000	0.764
Share of Employed in Manufact. Sec. 1951	300	0.329	0.178	0.026	0.861
Share of Employed in Manufact. Sec. 1961	300	0.369	0.176	0.032	0.806
Share of Employed in Manufact. Sec. 1971	300	0.424	0.155	0.042	0.782
Share of Employed in Manufact. Sec. 1981	300	0.413	0.130	0.074	0.730
Share of Employed in Manufact. Sec. 1991	300	0.370	0.109	0.067	0.637
Share of Employed in Manufact. Sec. 2001	300	0.299	0.090	0.053	0.498
Share of Employed in Manufact. Sec. 2011	300	0.251	0.077	0.000	0.441
Share of Employed in Service Sector 1934	300	0.131	0.084	0.019	0.688
Share of Employed in Service Sector 1951	300	0.151	0.088	0.000	0.447
Share of Employed in Service Sector 1961	300	0.236	0.114	0.025	0.755
Share of Employed in Service Sector 1971	300	0.353	0.132	0.107	0.846
Share of Employed in Service Sector 1981	300	0.469	0.123	0.180	0.881
Share of Employed in Service Sector 1991	300	0.549	0.108	0.264	0.903
Share of Employed in Service Sector 2001	300	0.642	0.095	0.406	0.899
Share of Employed in Service Sector 2011	300	0.644	0.084	0.452	0.879

Sector shares in 1934 report the economic affiliation of the resident population, not the working population. There is a share of 9.6% reporting no economic affiliation. The share of of in-/out-commuters are the number of commuters in each group divided the working population in the municipality.

38% of in-commuters relative to the working population residing in a municipality.[3] Total participation rate in 2010 is 49.3% and ranges between 32 and 59% in various municipalities.

The population censuses also report the number of people working in agriculture, manufacturing, and the service sector. As in most industrialized countries, the share of people working in agriculture decreased sharply over the period 1934 to 2010, while the service sector increased in size at the same time period. The relative size of the manufacturing sector remained relatively stable, peaking in 1971 at 42% of total employment.

Firm Census Data
Another source of data is the firm census (*Arbeitsstättenzählung*) conducted in 2011, which report the number of non-agricultural firms by firm size and industry. The average number of firms per municipality in the sample is 152 in 2001, and the mean number of firms per 1,000 inhabitants is 70. There are on average four times as many service sector firms as there are manufacturing firms.

1.3.3 WWII Casualties and Pre-War Municipality Characteristics

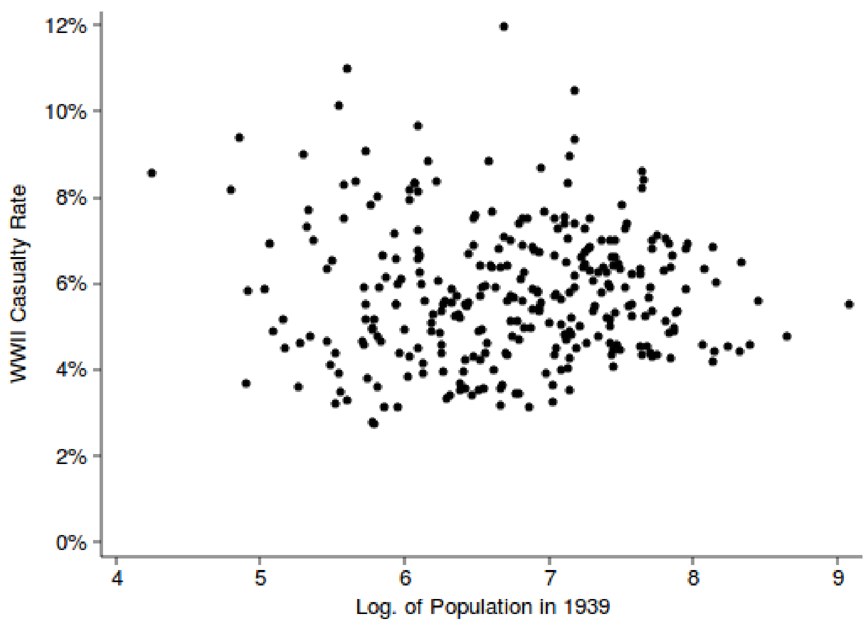
Given the municipality data, I am now able to test the relationship between the available pre-war municipality characteristics and the WWII casualty rate. If municipality characteristics influence the casualty rate, we would suspect that the identification strategy is not valid as the same municipality characteristic could influence post-war economic development and create a spurious correlation. Table 1.3 shows the regression result of the share of dead soldiers of the 1939 population on various municipality variables. District fixed effects are always included. The result shows only one explanatory variable with a point estimate significantly different from zero: the population in 1939. If this relationship would be an artefact of different population characteristics in larger municipalities and hence different mortality rates in the military, that relationship should also imply significant effects of the sector shares in column (2) and (7) on the casualty rate. However, this is not the case.

An alternative explanation is sample selection. Figure 1.3 shows the WWII casualty rate plotted against the log. of population in 1939. The variance of the casualty rate decreases

[3] Note that this figure number can exceed 100% by construction. Within Austria the share of in- and out-commuters should be about equal, but since cities are dropped in the sample, there is a higher share of out-commuters.

with population size as expected. While for small municipalities with a high casualty rate there are a number of observations, this is not true for small municipalities with a low casualty rate. Municipalities with relatively few WWII casualties maybe did not erect a war memorial and hence are not present in my sample. This would generate a negative correlation of the WWII casualty variable with the population size in 1939. This would only be a problem in the following sections if the model is misspecified with respect to the the population size in 1939 and hence the WWII casualty variable is correlated to the error term. However, the main results hardly change when a squared polynomial of the population size in 1939 are included in the robustness checks in Table 1.5. Additional polynomials do not change the results either (not shown).

Figure 1.3: WWII Casualties and Population Size in 1939



The pre-war or time invariant municipality characteristics in the regression include the elevation of the municipality, employment shares in agriculture, manufacturing, and services, the market status of the municipality, the share of the male population, a set of political variables from the last pre-war federal election, and the share of the Jewish population. Table 1.3 shows that none of these variables shows a significant correlation with the WWII casualty rate. A test for joint significance of these ten variables in column 7 results in a F-value of 1.61 and a p-value of 0.104. The separate regressions of the casualty rate on these

Table 1.3: WWII Casualties and Pre-War Municipality Characteristics

	WWII Casualty Rate of 1939 Population						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log. Population in 1939	-0.005***	-0.004**	-0.004***	-0.004***	-0.005***	-0.004***	-0.005**
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Log. of Elevation	-0.006						-0.005
	(0.004)						(0.004)
Share in Agriculture in 1934		-0.008					0.002
		(0.018)					(0.018)
Share in Manufacturing in 1934		-0.002					0.000
		(0.020)					(0.019)
Share in Services in 1934		-0.033					-0.027
		(0.025)					(0.025)
Market Status Dummy in 1945			0.002				0.003
			(0.003)				(0.003)
Share of Male Population in 1934				0.007			0.003
				(0.042)			(0.042)
Vote Share of NSDAP in 1930					0.063		0.067
					(0.047)		(0.046)
Vote Share of Social Democrats in 1930					0.013		0.017
					(0.011)		(0.013)
Vote Share of Christian Democrats in 1930					0.007		0.005
					(0.009)		(0.010)
Share of Jewish Population in 1934						-0.108	-0.097
						(0.113)	(0.099)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. Observations	300	300	300	300	300	300	300
R-squared	0.25	0.25	0.24	0.24	0.25	0.24	0.27

Significance: $*$: 10% $**$: 5% $***$: 1%
Standard errors in parenthesis.
Sector shares in 1934 report the economic affiliation of the resident population. The employment category not included in the regression is the share of the population reporting no economic affiliation with an average of 9.6%.

variables controlling for 1939 population in columns 1-6 do not produce significant results either.

Even though the finding shows that the WWII casualty rate is not driven by municipality characteristics, this does not establish that the measure of WWII casualties is orthogonal to the error term in equation (1.1). Unobservable characteristics could still drive WWII casualties and post-war economic development. Robustness checks and falsification tests after the main results elaborate on this issue.

1.4 WWII Casualties and Economic Output

This section describes the relationship between economic outcomes and casualties in WWII. The main interest of this paper lies in whether economic activity in a municipality today is affected by the number of dead soldiers during WWII. Economic activity is measured by the total wage bill paid within the borders of a municipality. The total wage bill can be interpreted as an approximate output measure for each municipality in the year 2010.

Figure 1.4: Total Wage Bill per Capita and WWII Casualties

(a) Unconditional Relationship (b) Within-District Relationship

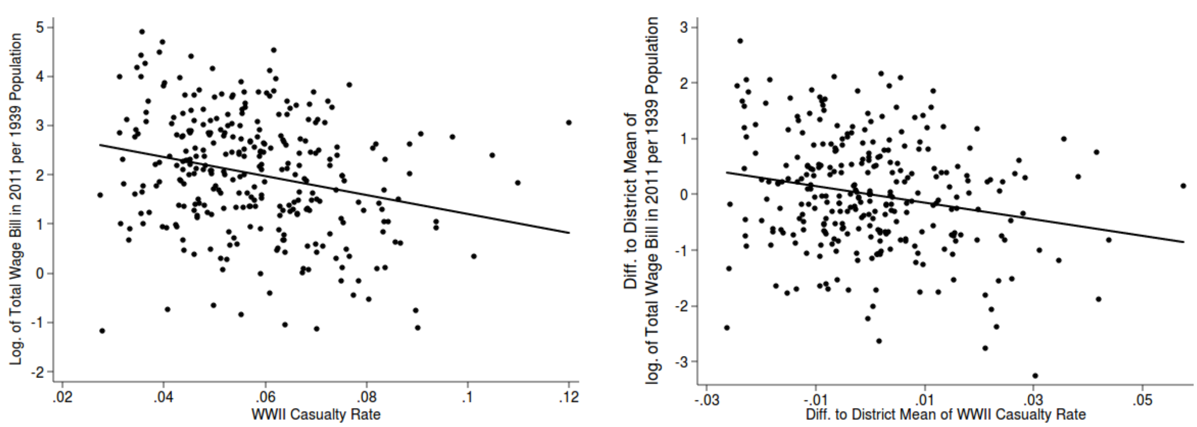


Figure 1.4 shows a scatter plot of the log. of the total wage bill per capita and the WWII casualty rate. The total wage bill is divided by the 1939 population so that the graph captures the total effect of WWII casualties on the output level today, including its effect on subsequent population growth. However, slower population growth in high-casualty

municipalities does not drive the negative relationship found here because when the total wage bill divided by population in 2011 is used instead, the graph is almost indistinguishable from the one shown. Panel A shows the unconditional correlation with a clear negative relationship. Deviations from district means are used in Panel B, which shows a very similar picture.

The set of regressions reported in Table 1.4 shows a strong negative relationship between WWII casualties and total municipality output in 2011. Column 1 reports a rather large unconditional correlation. The point estimates decrease slightly by controlling for the population size in 1939 in column 2. The effect decreases further as municipality level control variables for the market status of a municipality in 1945, the share of employment in agriculture in 1934, and the share of males in the population in 1934 are included.

In columns 4 and 5 geographic fixed effects, either for the province or the district, are added which reduces the point estimate further. In the eight provinces of the sample there were 98 districts in total in 2010, out of which 43 are represented in the sample. District fixed effects take out much of the geographic variation, administrative differences and other unobservable characteristics within Austria. In what follows, all results are based on within-district variation.

Column 5 is the preferred specification. The estimated effect is -9.62 and is significantly different from zero with a p-value of 0.032. A one percentage point increase in the share of dead soldiers is associated with about a 9.5% reduction of the total wage bill more than 60 years later. Note that the standard deviation of the share of dead soldiers is 1.6%, which implies that the effect of on the municipality wage bill is far from trivial.

In column 6, I weight the observations by population size in 1939 to place higher importance on more populous municipalities. The point estimate remains almost unchanged. Column 7 uses an adjusted version of the output variable to account for the fact that wages paid in agriculture and the public sector are not subject to the *Kommunalsteuer*, which is the basis for the calculated wage bill. A lower wage bill measure could therefore be due to a higher share of employees in agriculture or the public sector. However, the estimated effect decreases only sightly compared to the preferred specification in column 6 and remains significant at the 5% level.

Table 1.4: Total Wage Bill by Municipality in 2011

	Log. of Total Wage Bill in 2011								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
WWII Casualty Rate of 1939 Population	-20.31***	-19.02***	-14.71***	-10.46**	-9.62**	-9.76**	-9.03**	-8.24**	-10.62*
	(6.286)	(4.858)	(4.315)	(4.394)	(4.453)	(4.305)	(4.446)	(4.046)	(5.453)
WWII Casualty Rate* Russian Occupation Zone									3.99
									(8.916)
Log. Population in 1939		1.22***	0.90***	1.11***	1.12***	1.13***	1.10***		1.11***
		(0.079)	(0.092)	(0.093)	(0.097)	(0.103)	(0.096)		(0.098)
Log. Population in 2011								1.23***	
								(0.081)	
Market Status Dummy in 1945			-0.20	0.15	0.20	0.15	0.19	0.12	0.19
			(0.167)	(0.160)	(0.169)	(0.153)	(0.167)	(0.148)	(0.169)
Share in Agriculture in 1934			-2.34***	-2.35***	-2.11***	-1.87***	-1.99***	-1.20***	-2.12***
			(0.303)	(0.297)	(0.337)	(0.325)	(0.333)	(0.315)	(0.339)
Share of Male Population in 1934			-0.97	2.20	0.96	1.12	1.15	3.66	1.00
			(3.118)	(3.115)	(3.106)	(3.264)	(3.077)	(2.772)	(3.096)
Province FE				Yes					
District FE					Yes	Yes	Yes	Yes	Yes
Population 1939 Weights					Yes				
Adjusted Total Wage Bill						Yes			
No. Observations	300	300	300	300	300	300	300	300	300
R-squared	0.04	0.45	0.54	0.63	0.70	0.79	0.69	0.77	0.70
Adj. R-squared	0.04	0.45	0.54	0.62	0.65	0.75	0.64	0.73	0.65

Significance: $*: 10\%$ $**: 5\%$ $***: 1\%$

Standard errors in parenthesis.

The adjusted output takes into account that people employed in agriculture and publicly employed are not subject to the tax on salaries, which is the basis for the output measure. The adjusted variable is the log. of total output as defined before divided by the share employed outside of agriculture and the public sector in 2011.

The next column (8) tests an obvious explanation for the large negative effect found above: a different population growth pattern since WWII. If a higher casualty rate translates into lower population growth in the affected municipalities, it would not be a surprise if those municipalities have a lower total wage bill in 2011. The estimated effect reduces to -8.24, but remains significant at the 5% level. This shows that a changed population growth pattern is not the driving force behind the results.

One could argue that the Russian occupation in the east of Austria has left a mark on the economic landscape and the estimated effect is entirely driven by those eastern districts. Column 9 shows that this was not the case. The interaction of the dummy variable for the Russian occupation zone with the war casualty measure shows no significantly different pattern than in the rest of Austria. The point estimate for the base group (central and western Austria) increases slightly to -10.62 and looses slightly in significance (p-value: 0.053).

In summary, the negative estimated effects of WWII casualties are large. A one percentage point increase in the casualty rate of a municipality reduces the total wage bill in 2011 by 9.5% in the preferred specification. I interpret the total wage bill as an approximate output measure, that is, at least at the provincial level, highly correlated with GDP. When, for purely illustrative purposes, these numbers are interpreted as a reduction in production without spillovers to other municipalities, a counterfactual of no deaths during WWII would increase output by 55.6%. The reduction in output per capita would be slightly lower, but still large.

A more realistic interpretation takes into account that WWII casualties influenced the development of commercial hubs in which economic activity is clustered. In an extreme case of this interpretation, the estimated effect does not translate into an aggregate reduction of output at all, but a clustering of economic activity into certain municipalities.

1.4.1 Robustness Checks

Before exploring the channels through which the established effect could emerge, I first run a series of robustness checks and falsification tests to confirm that the effect of casualties on the total wage bill is a causal one. This section tries to rule out a number of observed municipality characteristics to drive both, WWII casualties and post-war economic development.

Table 1.5: Total Wage Bill by Municipality in 2011 - Robustness Checks

	Log. of Total Wage Bill in 2011					
	(1)	(2)	(3)	(4)	(5)	(6)
WWII Casualty Rate of 1939 Population	-9.53**	-8.95**	-9.35**	-9.60**	-11.22***	-10.10**
	(4.438)	(4.381)	(4.522)	(4.462)	(4.292)	(4.244)
Log. Population in 1939	1.88**	1.11***	1.11***	1.12***	1.08***	1.37
	(0.912)	(0.096)	(0.100)	(0.097)	(0.097)	(0.918)
Market Status Dummy in 1945	0.24	0.18	0.16	0.20	0.26	0.18
	(0.178)	(0.176)	(0.177)	(0.171)	(0.162)	(0.184)
Share in Agriculture in 1934	-2.12***	0.67	-2.63***	-2.10***	-1.63***	0.99
	(0.340)	(1.449)	(0.425)	(0.344)	(0.371)	(1.266)
Share of Male Population in 1934	1.19	1.16	2.18	0.96	1.76	3.06
	(3.131)	(3.276)	(3.201)	(3.113)	(2.907)	(3.171)
Log. of Population in 1939 squared	-0.06					-0.02
	(0.069)					(0.069)
Share in Manufacturing in 1934		2.99*				3.03**
		(1.559)				(1.362)
Share in Services in 1934		4.45*				5.47***
		(2.267)				(2.044)
Vote Share of NSDAP in 1930			3.52			4.18
			(3.395)			(2.645)
Vote Share of Social Democrats in 1930			-1.08			-0.59
			(0.780)			(0.758)
Vote Share of Christian Democrats in 1930			0.04			0.41
			(0.612)			(0.577)
Share of Jewish Population in 1934				0.47		1.77
				(3.810)		(4.933)
Log. of Elevation					-1.03***	-1.18***
					(0.310)	(0.305)
District FE	Yes	Yes	Yes	Yes	Yes	Yes
No. Observations	300	300	300	300	300	300
R-squared	0.70	0.71	0.71	0.70	0.72	0.74
Adj. R-squared	0.65	0.65	0.65	0.65	0.67	0.67

Significance: $*$: 10% $**$: 5% $***$: 1%
Standard errors in parenthesis.

In Table 1.5, I include several additional control variables to the preferred specification to check if the estimated effect of WWII casualties disappears.

In column 1, the log. of population of 1939 is included in a non-linear fashion to allow for a more general population growth pattern. Column 2 includes the employment shares in 1934 for the three sectors agriculture, manufacturing, and services to take into account the pre-war industry composition, which might had an effect on the casualty pattern during WWII and post-war economic growth. The excluded group are the unemployed who never had work before. However, the estimated effect of WWII casualties remains of the same qualitative level and significantly different from zero at the 5% level.

Another hypothesis concerns the political environment of the municipality. Increased support for the NSDAP, the party of Hitler, in a municipality could act as a confounding factor. Resource allocation decisions could have been in favor of highly supporting municipalities, whose soldiers were also more fanatic in the battlefield. Support for the NSDAP is measured here as the vote share in the 1930 federal election, the last free election before the *Anschluss*. The other included parties are the Social Democratic Party and the Christian Democratic Party. As column 3 shows, a higher vote share does not change the effect of WWII casualties on the total wage bill.

Acemoglu et al. (2011) show that the reduction of the Jewish population during the Holocaust had long-lasting effects on Russian cities. In this spirit, I test the effect of the share of the Jewish population before the war. The share of Jews was very low in rural Austria in 1934 with an average of 0.2% in the sample municipalities. As can be seen in column 4, the inclusion of this variable does not change the main result.

The elevation of the municipality could plausibly have an effect on both the casualty rate through increased hiding possibilities and post-war economic development, for example through tourism. In column 5, I include the log. of elevation of the municipality and find that the effect of interest does increase and decrease the standard error at the same time.

In column 6, all the above variables are included. This leads to a small change of the estimated effect to -10.10. These findings show that the negative effect of WWII casualties on the approximate output measure is robust to the inclusion of observable municipality characteristics. Further indication of a causal relationship is presented in the next section that performs falsification tests.

1.4.2 Falsification Tests

I now focus on some falsification tests, which test the correlation of pre-war economic development with the WWII casualty measure. There is still the possibility that an omitted variable drove both economic development and WWII casualties. If this is the case, the omitted variable could have also driven pre-war economic development and hence a correlation between the WWII casualty rate and economic variables measured before the war should be existent.

Table 1.6: Falsification Tests

	Share in Agricultural Sector in 1934	Share in Manufacturing Sector in 1934	Share in Service Sector in 1934		Share Never Employed in 1934	Log. Number of Large Land Holdings in 1900	Log. Number of Factories in 1900	Log. of Population in 1939
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
WWII Casualty Rate of 1939 Population	0.15	0.45	-0.55*	-0.37	-0.25	-0.55	0.81	-1.16
	(0.666)	(0.472)	(0.286)	(0.252)	(0.159)	(3.020)	(1.349)	(0.763)
Log. Population in 1939	-0.17***	0.11***	0.04***	0.04***	-0.00	0.14*	0.16***	
	(0.017)	(0.013)	(0.007)	(0.007)	(0.004)	(0.078)	(0.051)	
Log. Population in 1900								0.95***
								(0.031)
Market Status Dummy in 1945	0.01	-0.05*	0.03**	0.04***	0.01*	-0.06	-0.02	-0.04
	(0.035)	(0.027)	(0.015)	(0.012)	(0.007)	(0.209)	(0.135)	(0.039)
Share in Agriculture in 1934					-0.15***	0.09	-0.57***	-0.54***
					(0.020)	(0.231)	(0.145)	(0.087)
Share of Male Population in 1934	1.14**	-0.69**	-0.29	-0.36*	0.15	1.87	2.36**	0.21
	(0.484)	(0.308)	(0.198)	(0.193)	(0.184)	(1.630)	(0.970)	(0.563)
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Outliers Dropped				Yes				
No. Observations	300	300	300	298	300	242	242	300
R-squared	0.58	0.51	0.38	0.44	0.59	0.58	0.57	0.96

Significance: $*$: 10% $**$: 5% $***$: 1%

Standard errors in parenthesis.

The sector employment shares in 1934 measure the affiliation of the resident population with each sector. Each persons in a household who has never worked, is affiliated to the sector of employment of the household head. The share of is the share of household heads including his/her family who has never been employed relative to the total population in 1934, which underestimates the unemployment rate at that time.

Unfortunately, there are very limited information on pre-war development available to me. The LHS variables I use are the share of the resident population affiliated to agriculture, manufacturing, and services, the share of residents who never worked before (a proxy for unemployment with an average of 9.1% in 1934)[4], as well as the population growth between 1900 and 1934. Additionally there is a small set of economic variables from the year 1900: the number of large land holdings and the number of factories in each municipality.

The results are presented in Table 1.6, with the columns 1-4 testing the relationship between employment shares in 1934 and the WWII casualty variable. The share in the service sector in 1934 is significantly correlated with the casualty rate at the 10% level, but dropping two massively outlying observations, the point estimate shrinks and turns insignificant.[5] Considering also regression 2 of the robustness checks in Table 1.5 with the full set of sector shares as control variables included, where the estimated effect on the output measure in 2011 decreases only slightly, the negative relationship found here does not seem to drive the effect on the total wage bill.

Column 5 uses the share of people who never worked before in 1934, which I use as a proxy for unemployment. However, there is no significant correlation with the war casualty measure. Columns 6 and 7 tests the relationship the number of large land holdings and factories, both measured in 1900.[6] There is also no correlation between the 1900 economic landscape and the variable of interest. A proxy for pre-war economic development is population growth between 1900 and 1939. But as column 8 shows there is again no significant relationship with the war casualty variable.

In summary, this section established a strong negative and robust relationship between the output generated within the borders of a municipality and the share of the population of 1939 that died during WWII as soldiers. The robustness checks and falsification test suggest that this is a causal relationship and not a mere correlation. Causality implies that the human loss of life caused by the war has a significant impact on today's economic landscape. The question why this difference exists is what I explore in the following sections.

[4] Not working family members are affiliated in the sector of the household head. The 9.1% are therefore the share of the population where the household head has never been employed before.

[5] The municipalities are "Lech" and "Semmering", both known for large ski resorts.

[6] Observations are lost since only municipalities with unchanged borders since 1900 can be used. Also the province of Burgenland was not included in the Austrian census in 1900.

1.5 Where Does the Effect Come From?

In this section I try to identify the determinants of the output difference. To this end, I look at contemporary measures of inputs into a hypothetical production function to identify why high-casualty municipalities produce less output today.

1.5.1 Population, Employment, and Commuting Pattern

Table 1.7 reports the direct effect of WWII casualties on the population size in 2011 in column 1. The effect is insignificant, which confirms the finding in Table 1.4 with the control variable for population in 2011. Column 2 shows that there is also no effect on the working population of a municipality, while columns 3 and 4 find no difference in male and female employment rate. These findings all suggest, that WWII casualties did not affect aggregate statistics of the resident population.

Table 1.7: Population, Work Force, and Commuting Pattern in 2011

	Log. Total Pop.	Log. Working Pop.	Employment Rate Male	Employment Rate Female	Log. Employed in Munic.	Log. Out-Commuters	Log. In-Commuters
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
WWII Casualty Rate of 1939 Population	-0.76	-0.63	0.10	0.05	-7.13**	1.20	-8.72**
	(1.318)	(1.390)	(0.154)	(0.181)	(2.794)	(1.821)	(4.272)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. Observations	300	300	300	300	300	300	299
R-squared	0.91	0.90	0.28	0.40	0.80	0.85	0.72

Significance: ∗ : 10% ∗∗ : 5% ∗ ∗ ∗ : 1%
Standard errors in parenthesis.
Control variables include the log. population in 1939, an indicator variable for market status in 1945, the share of population in agriculture in 1934 and the share of the male population in 1934.

If there is no difference in the resident population, could it be a difference of people working in specific municipalities, but living in another one? Column 5 tests this hypothesis, which uses the number of people employed within the borders of a municipality as the dependent variable. The result shows that the number of employed people within the borders of a municipality is vastly reduced by WWII casualties. A one percentage point increase in the share of dead soldiers reduces the number of jobs in a municipality by 7.13%, which is of a similar magnitude than the total effect on the total wage bill.

If there are more people working within the borders of a municipality, there could be two reasons: ceteris paribus, there could be more in-commuters or fewer out-commuters. Columns 6 and 7 focus on commuting streams, where the effect on in-commuters clearly dominates the effect on out-commuters. In section A.4 in the Appendix, I estimate a gravity model with municipality fixed effects, which gives more precise estimates. Those results clearly show that the large effect of WWII casualties on total employment in a municipality is due to a difference in the number of in-commuters.

1.5.2 Number and Density of Firms

There are more people working within municipalities with fewer WWII casualties. A natural next step is to ask where they work. I use data from the *Arbeitsstättenzählung* of 2011, a firm census conducted by the national statistical agency, to explore the firm landscape by municipality. The scope of the survey are non-agricultural employers and it contains data on the number of employers by industry and size, grouped into intervals of number of people employed.

Table 1.8: Number, Density, and Size of Firms in 2011

	Log. of Number of Firms		Log. of Number of Firms per 1,000 Inhabitants		Size of Firms by Number of Workers (grouped)		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
WWII Casualty Rate of 1939 Population	-5.74*** (2.044)	-5.48** (2.120)	-4.98*** (1.677)	-4.57*** (1.513)	-0.40** (0.191)	-0.97 (0.618)	-0.65** (0.310)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Population 1939 Weights		Yes		Yes			
Ordered Probit						Yes	
Interval Regression							Yes
No. Observations	300	300	300	300	53,536	53,536	53,536
(Pseudo) R-squared	0.84	0.88	0.29	0.42	0.02	0.02	.

Significance: $*$: 10% $**$: 5% $***$: 1%
Robust standard errors in parenthesis.
Control variables include the log. population in 1939, an indicator variable for market status in 1945, the share of population in agriculture in 1934 and the share of the male population in 1934.

In a set of regressions, the outcome variables are the log. of the total number of firms of any size, the log. of firm density, defined as the total number of firms by 1,000 inhabitants in 2001, and measures of firm size. The results in Table 1.8, columns 1-4, show a negative relationship between WWII casualties and the total number of firms and the firm density in the weighted and unweighted regressions. The number for firms is reduced by 4 to 6% if WWII casualties increases by one percentage point. This number is smaller than the effect on total output and the number of employed people in a municipality, but complimentary to those findings, as more in-commuters need to be employed somewhere.

However, a larger number of firms is not the only way to employ more people, as larger firms could produce the same result. I find some evidence that the size of firms, measured by the number of employees, is correlated with WWII casualties as can be seen in the regression of columns 5-7. The firm size is grouped into intervals of 0-4, 5-19, 20-99, 100-250, and more than 250 employees. A regression using OLS on the bins (numbered 1 to 5) in column 5, ordered probit in column 6, and an interval regression model with frequency weights by number of firms in the interval in the last column, produce negative point estimates with two of them significantly different from zero at the five percent level.

The data in the firm census of 2011 groups them into a number of industries. Dividing them into manufacturing and services gives further insights. The results in Table 1.9 show clearly that the effect on the number and density of firms stems entirely from the service sector. The manufacturing sector produces point estimates close to zero, while the effect of WWII casualties on the number of service sector firms is very similar to the effect on the total number of firms. Neither in the manufacturing nor the service sector an effect on firm size is found (not shown in paper).

Table 1.9: Firms by Sector in 2011

	Log. of Firms by Sector			Log. of Firms by Sector per 1,000 Inhabitants		
	Manufacturing	Service	Service	Manufacturing	Service	Service
	(1)	(2)	(3)	(4)	(5)	(6)
WWII Casualty Rate of 1939 Population	-1.70 (2.289)	-5.83*** (2.080)	-5.80*** (2.157)	-2.37 (1.922)	-5.34*** (1.885)	-5.02*** (1.614)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Population 1939 Weights			Yes			Yes
No. Observations	300	300	300	294	300	300
R-squared	0.81	0.83	0.87	0.31	0.27	0.42

Significance: $*$: 10% $**$: 5% $***$: 1%
Standard errors in parenthesis.
Control variables include the log. population in 1939, an indicator variable for market status in 1945, the share of population in agriculture in 1934 and the share of the male population in 1934.Manufacturing Sector includes mining, manufacturing, energy and water supply, and construction industries. Service Sector category is made up of commerce, tourism, IT and communication, banking and insurance, real estate, public administration, education, health industry, and personal services.

The results in Table 1.9 are surprising insofar as the manufacturing sector does not exhibit any difference in high-casualty and low-casualty municipalities. In contrast, the service sector does account for all of the difference found earlier. Considering the relatively small share of the service sector in the years following WWII is not likely that the difference set in right after the war and remained constant. Instead, it seems more likely that a dynamic effect has been at play and the effect on output emerged at a later time, which will be explored next.

In summary, the difference in the total wage bill between high-casualty and low-casualty municipalities can likely be explained by the number of people being employed in the municipalities, which in turn stems from a higher number and density of firms located in the low-casualty municipalities. More workers and firms also produce more output. Moreover, the difference in firms can be attributed to service sector firms, while there is no difference in the manufacturing sector. A closer look at the sector composition of employment can give further insights into the channel of persistence.

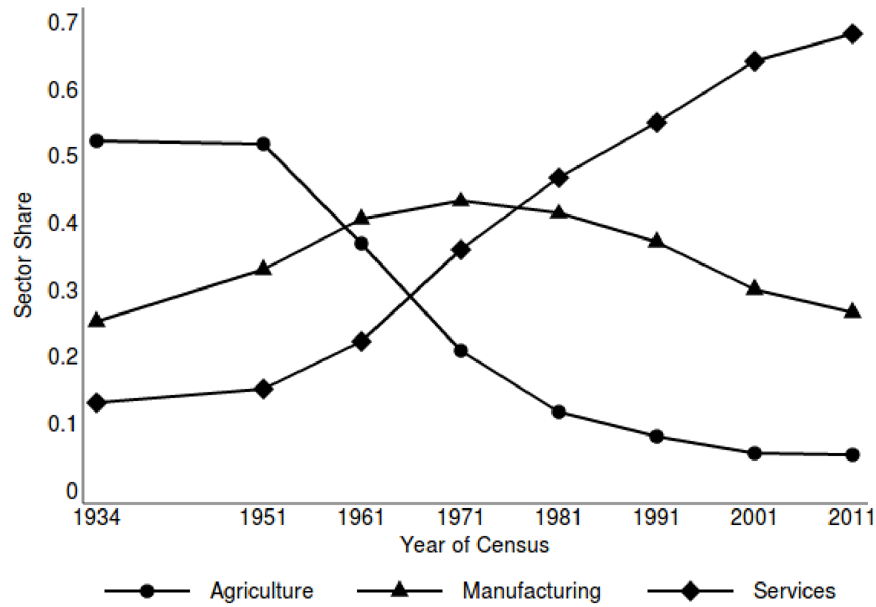
1.6 The Process of Divergence

To better understand the process of divergence between municipalities with high and low human losses during WWII, I now turn to a dynamic investigation.

The sector composition of the labor market has changed dramatically since the 1950s in Austria. Figure 1.5 plots the share of workers employed in agriculture, manufacturing, and the service sector for the municipalities of the sample.[7] In 1951, more than half the population was working in the agricultural sector, while only 15% was working in the service sector. By 2010, the share of workers in agriculture slumped to 6%, while services grew to 68%. The share of the manufacturing sector follows a hump shaped path, peaking in 1971 at 43%. As we will see, the differences in transition patterns between high-casualty and

[7] The numbers are taken from the Austrian population censuses, where the working population was divided into several industries. The number of reported industries increased every decade, but were put into those three broad categories to be comparable over time. In the sector classification used in 2001, the sector "Agriculture" includes agriculture, forestry and fishing, while "Manufacturing Sector" includes mining, manufacturing, energy and water supply, and the construction industry. The "Service Sector" category is made up of commerce, tourism, IT and communication, banking and insurance, real estate, public administration, education, health industry, and personal services. The numbers from the census of 1934 divides the resident population into sectors, not the working population, and the unknown classification was ignored.

Figure 1.5: Evolution of Sector Shares in the Sample Municipalities



low-casualty municipalities are a probable channel that produced the difference found in the current economic landscape.[8]

1.6.1 Estimation

Sector shares of the labor market are correlated between sectors (as they always sum to one) and over time. I therefore employ a GLS model that takes these correlations into account. All the error terms within a municipality are allowed to be correlated with each other, but not between municipalities. I also use the pre-war sector share of the outcome variable as a control variable to make the point estimates between sectors more comparable. This also explains why the point estimates do not sum exactly to zero for each year. District-Year-Sector fixed effects are included.

[8]Appendix A.5 discusses another channel, the influx of foreigners and sorting into high-casualty municipalities to replace the missing workforce. However, there is no change in the share of foreign population in 1951, 1971, and 2001 due to WWII casualties.

Table 1.10: Sector Shares

PANEL A	Share Working in Agricultural Sector in						
	1951	1961	1971	1981	1991	2001	2011
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
WWII Casualty Rate	-0.71**	-0.44	-0.79**	-0.08	-0.23	0.02	0.14
of 1939 Population	(0.276)	(0.337)	(0.376)	(0.271)	(0.193)	(0.167)	(0.126)
PANEL B	**Share Working in Manufacturing Sector in**						
	1951	1961	1971	1981	1991	2001	2011
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
WWII Casualty Rate	0.39	0.74**	1.10***	0.90***	1.06***	0.71***	0.55**
of 1939 Population	(0.259)	(0.323)	(0.379)	(0.336)	(0.307)	(0.270)	(0.226)
PANEL C	**Share Working in Service Sector in**						
	1951	1961	1971	1981	1991	2001	2011
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
WWII Casualty Rate	0.30	-0.22	-0.04	-0.75**	-0.77**	-0.70**	-0.28
of 1939 Population	(0.195)	(0.256)	(0.340)	(0.321)	(0.313)	(0.279)	(0.205)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District-Year-Sector Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. Observations	300	300	300	300	300	300	300

Significance: $*$: 10% $**$: 5% $***$: 1%
Standard errors incorporate correlation between sectors and over time.
The reported estimates are based on a GLS model that allows for within municipality correlation of the error terms between sectors and over time. District-Year-Sector fixed effects are always included. The usual control variables are included, except the sector share of 1934 of the outcome variable is used instead of the share of the agricultural sector in 1934.

The result of the regressions paint an interesting picture of the differential patterns of divergence between high-casualty and low-casualty municipalities during the post-war period, shown in Table 1.10. In Panel A the size of the agricultural sector is the dependent variable. There is a large negative relation between WWII casualties and the share working in agriculture in 1951 and 1971. Afterwards no sizable difference can be detected.

Conversely, Panel B shows that the share of workers in the manufacturing sector is significantly higher in high-casualty municipalities in all the decades except the 1950s. Interestingly, the increased share of the manufacturing sector comes at the expense of the agricultural sector until 1971 (not significant in 1961) and at the expense of the service sector thereafter. The size of the effect is not trivial, as a one percentage point increase in

WWII casualties results in a one percentage point increase in the share of manufacturing sector in 1971. Panel C investigates the service sector and finds that highly affected municipalities had a lower share of workers employed in the service sector from 1981 onwards.

In summary, residents of high-casualty municipalities are more likely to find work in the manufacturing sector and conversely less likely to be in agriculture until 1971 and less likely to be in services since the 1981 census.

These results suggest that people in high-casualty municipalities sorted initially from the agricultural sector into the manufacturing sector. Bellou and Cardia (2013) document increased employment in the manufacturing sector in 1960 for men and women in the US in states with higher mobilization rates during WWII. This is consistent with my findings as higher mobilization rates likely translate into more casualties.

Labor shortage can explain the initial sorting from agriculture to manufacturing. Given the technology used in agriculture at that time, men had a comparative advantage in agriculture, while women were comparatively more productive in manufacturing. A household affected by a war casualty was less productive and had to give up the farm and likely ended up in manufacturing, as this was the next largest sector after the war.

The explanation above suggests that there were more manufacturing sector firms in high-casualty municipalities in the decades after the war and there was initially no difference in the number of service sector firms. Since the 1980s, the difference in the manufacturing sector should have disappeared, but there should be now a negative difference in the number of service sector firms.

This is exactly the pattern found in Table 1.11 that looks at the number of firms by sector since 1973, when the first first round of the firm census was conducted. While there is a significant positive effect of WWII casualties on the number of manufacturing firms in 1973, that effect disappears and even shows negative point estimates in 2011. The effect on the number of service sector firms is close to zero and not significant in 1973, but the point estimates increase in magnitude and statistical significance until now.

The model in the next section explores the dynamics of a labor shortage in agriculture in combination with structural change that is able to explain the pattern found in the data.

Table 1.11: Number of Firms by Sector

PANEL A	Log. of Number of Firms in Manufacturing Sector				
	1973	1981	1991	2001	2011
	(1)	(2)	(3)	(4)	(5)
WWII Casualty Rate	5.29***	1.98	0.73	-0.10	-1.95
of 1939 Population	(1.79)	(1.88)	(1.93)	(1.83)	(1.85)

PANEL B	Log. of Number of Firms in Service Sector				
	1973	1981	1991	2001	2011
	(1)	(2)	(3)	(4)	(5)
WWII Casualty Rate	-0.92	-2.75*	-3.41**	-3.68**	-5.05***
of 1939 Population	(1.69)	(1.91)	(1.97)	(1.86)	(1.79)
Control Variables	Yes	Yes	Yes	Yes	Yes
District-Year-Sector Fixed Effects	Yes	Yes	Yes	Yes	Yes
No. Observations	300	300	300	300	300

Significance: $*$: 10% $**$: 5% $***$: 1%
Standard errors incorporate correlation between sectors and over time.
The reported estimates are based on a GLS model that allows for within municipality correlation of the error terms between sectors and over time. District-Year-Sector fixed effects are always included. The usual control variables are included.

1.6.2 A Model of Structural Change and WW2 Casualties

The regression results in Table 1.10 pose the question how a initial reduction of employment in agriculture in high-casualty municipalities can translate to lower service sector employment decades later. One prominent feature of the Austrian economy at that time was structural change as already shown in Figure 1.5. The following model of structural change illustrates a way to make sense of the regression results.

Structural change is modeled as a demand side phenomenon following Kongsamut, Rebelo, Xie (2001), so that growing income lets consumers reduce the fraction of income spent on agricultural goods and increase their relative spending on service sector goods. The casualty rate enters the model only through its negative impact on agricultural productivity. Municipalities are small and trade is performed at the national level, so that prices are determined at the aggregate.

Household Decision

Preferences of the representative household in a municipality are given by the utility function

$$u\left(c_a, c_m, c_s\right) = \left(c_a - \delta\right)^{\alpha}\left(c_m\right)^{\beta}\left(c_s + \delta\right)^{\gamma}.$$

Without loss of generality let $\alpha + \beta + \gamma = 1$. The parameter δ can be thought of as a subsistence level of consumption, so that up to a certain threshold no services are consumed. Households provide one unit of labor inelastically and receive a wage w in return. Let there be a measure one of households.

Production Decision

The production technology is linear with respect to labor input L. The WWII casualty rate enters the model only through its impact on agricultural productivity. A higher casualty rate reduces agricultural productivity as fewer men are available, who are relatively more productive in agriculture than women. Let the casualty rate in municipality i be d_i, which is measured as deviations from the mean, such that $E\left[d\right] = 0$. Let $F(d)$ be the cdf of d. Total factor productivity (TFP) in all sectors is described by the parameter B and the production functions are:

$$\begin{aligned} y_{i,a} &= B(1 - d_i)L_{i,a} \\ y_{i,m} &= BL_{i,m} \\ y_{i,s} &= BL_{i,s} \end{aligned}$$

Price Determination

The service sector good is used as the numeraire good. Then the first order conditions for profit maximization ($w = p_a B(1 - d) = p_m B = B$), market clearing conditions, and the definition of d such that $E\left[d\right] = 0$ imply that $w = B$ and $p_a = p_m = p_s = 1$.[9]

[9] I also assume that $Var\left[d\right]$ is small. Since there will be sorting of high-productivity (low-casualty) municipalities into the agricultural sector, the average productivity in that sector will increase leading to a lower market clearing price than $p_a = 1$. The market clearing price can not be solved for without distributional assumptions on d. I ignore this price adjustment under that premise that $Var\left[d\right]$ is small and therefore the price change is small.

Sector Shares

To connect the model with the regression results, I now look at the sectors of employment. Below a critical TFP level B^* the demand for service sector goods is zero. The sector employment shares are therefore divided into two segments.

1. If $B \leq B^* = \frac{\delta}{\gamma}$ and therefore $c_s = 0$, there exists a critical casualty rate $d_1^* = F^{-1}(1-\beta)$ such that

$$\begin{aligned} L_a = 1 \quad & \text{if} \quad d \leq d_1^* \\ L_a = 0 \quad & \text{if} \quad d > d_1^*. \end{aligned}$$

2. If $B > B^* = \frac{\delta}{\gamma}$ and therefore $c_s > 0$, there exists a function $d_2^*(B) = F^{-1}\left(\alpha + \frac{\delta}{B}\right)$ such that

$$\begin{aligned} L_a = 1 \quad & \text{if} \quad d \leq d_2^*(B) \\ L_a = 0 \quad & \text{if} \quad d > d_2^*(B). \end{aligned}$$

The model predicts initial sorting of low-casualty municipalities into the agricultural sector to exploit a productivity advantage.

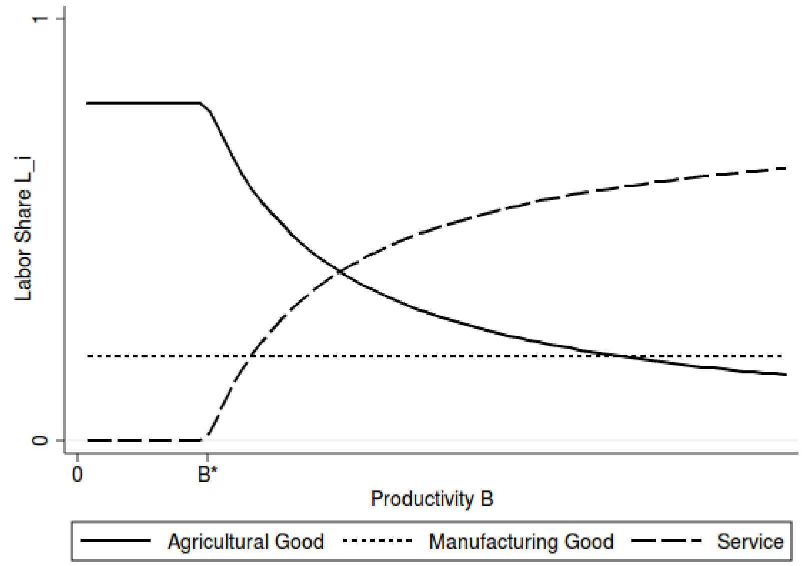
Productivity Growth and Sectoral Adjustment

Over time TFP increases which causes household income and demand for service sector goods to grow. Labor input shifts from the agricultural sector towards the service sector. Figure 1.6 illustrates an example of sector shares of total employment over a range of productivity levels.

In a frictionless world, workers from manufacturing might switch to the service sector. However, marginal sector transition costs of ε ensure that the sectoral transfer of labor takes place between agriculture and the service sector only, as a worker in manufacturing has no benefit from a sector transfer but faces a cost of ε. These transition costs could be a lack of inter-generational mobility, sector specific human capital, or uncertainty concerning a sector transition.

The sector of employment is a deterministic function of the casualty rate d: high-casualty municipalities are in manufacturing, medium-casualty municipalities in the service sector,

Figure 1.6: Sector Shares



and low-casualty municipalities in the agricultural sector. This kind of sorting results in $E\left[d|L_m=1\right] > E\left[d|L_s=1\right]$ after an extended period of productivity growth: the pattern found in the data.

Labor shortage in agriculture in combination with structural change is therefore able to convert a negative shock of a municipality into a long-lasting effect. This finding stands in strong contrast to Chaney and Hornbeck (2013) and Hornbeck and Naidu (2014) who find positive effects on income per capita due to labor shortage in agriculture and hence adjusted production technology.

1.7 Conclusion

This paper exploits WWII as a natural experiment for war casualties at the municipality level in Austria. I find a large negative effect of war casualties on current economic activity, as measured by the total wage bill. The effect is far from trivial, as a one percentage point increase in the WWII casualty rate decreases the local output by about 9.5%. Robustness checks and falsification test suggest a causal relationship. The underlying determinants of this reduction in output are a smaller number and density of firms and a smaller work force in highly affected municipalities. However, this is only true for the service sector and not in the manufacturing sector.

In a dynamic investigation I identify a differential path of the sectoral composition of the labor market in the post-war period. Municipalities with a high casualty rate in WWII have a higher share of workers in the manufacturing sector. Until 1971 the reduction can be found in the agricultural sector and from then onwards in the service sector.

The proposed channel of persistence is a sorting into manufacturing jobs from agriculture in highly affected municipalities after the war, due to a labor shortage in agriculture and hence lower agricultural productivity. When the labor demand in agriculture decreased due to structural change in the 1970s, workers in low-casualty municipalities (with a higher agricultural share) were more likely to switch to the, then quickly growing, service sector.

Chapter 2

Displacement and Education of the Next Generation: Evidence from Bosnia and Herzegovina

2.1 Introduction

Violent conflict is a regular phenomenon in the developing world and its long-term consequences can be substantial through the destruction of human and physical capital, damage to infrastructure, and displacement. However, research on the economic consequences of wars and other violent events have only recently gained momentum as micro-level data from conflict areas become available. This is an important line of research as it indicates to policy makers the direct consequences of violent conflicts and which issues need to be dealt with once the conflict is over.

The findings in this literature so far offer interesting insights into peoples' lives in war-torn countries and the effects of exposure to conflict. However, displacement has seen relatively little attention, even though the UNHCR counts 10.4 million refugees and 36 million people of concern in 2009 in their Global Report (UNHCR, 2009). Unfortunately, the literature mostly deals with short-run effects of exposure to conflicts. In this paper, I ask if the disadvantage of displacement has the potential to spill-over to the next generation.

I use the ethnic division during the Bosnian War between 1992 and 1995 as a natural experiment for displacement. The war divided the ethnically mixed Bosnia and Herzegovina into two parts of homogenous ethnic make-up. The resulting displacement of part of the population enables me to uncover its causal effects. The identification strategy circumvents the problem of endogenously migrated households by using only households who moved across the front line during the war. I identify education as one channel through which displacement has potential long-term consequences for the next generation.

The short time span between the end of the war and the collection of the data set does not allow me to evaluate educational outcomes of children of displaced parents.[1] However, the education system of Bosnia and Herzegovina requires parents to provide textbooks, uniforms, school materials etc. to their children, which gives me the chance to look at inputs in educational production. In particular, I am interested in education expenditure of displaced parents on their children.

I find that displaced parents spend significantly less on the education of their children five years after the end of the war than comparable households that were not displaced. The estimates of the reduction in spending on education range between 20% and 30%. Considering that the average household in Bosnia and Herzegovina spends more than half a monthly household income on the education of a child per year, this is quantitatively a large difference. The finding is robust to a number of specifications and nearest-neighbor matching estimates confirm the finding. Displaced parents also spend significantly less on single expenditure positions like annual tuition in secondary school, textbooks and other school materials.

Recent experimental evidence suggests that students' test scores react positively to additional educational inputs within a given institutional setting (Das et al, 2013), confirming an intuition that has been challenged in the past.[2] As small an investment as $3 per student spent mostly on child's stationary, classroom materials, and practice books can increase test scores by 0.1 standard deviations in India (Das et al, 2013). As Hanushek (2003) points out, there is evidence that increased school quality has positive effects on labor market outcomes

[1] Grade per age of children in school is not an educational outcome variable as displaced children might have missed years of schooling while being on the run, which then does not indicate lower performance in school.

[2] As Das et al (2013) show, when parents expect to receive additional resources, they substitute out of their educational expenditure and therefore the total educational inputs available to students do not increase much. This substitution is likely responsible for no effects found in previous experimental studies.

and productivity. Given this evidence, the lack of educational inputs provided by parents can have a long-term negative impact on childrens' labor market outcomes.

Exploring the causal channels through which displacement influences education expenditure, I show that at most one third can be explained by differences in income and wealth levels. The employment status of parents also fails to explain a major part of the difference, so there is no evidence that displaced parents substitute school inputs for increased parental effort as found by Houtenville and Smith Conway (2008). Some evidence supports the hypothesis that increased uncertainty about the future is the main reason, why displaced parents spend less on the education of their children.

This paper is organized as follows: The next section reviews the literature, Section 2.3 discusses the background of the study in Bosnia and Herzegovina, as well as the data and the identification strategy, while Section 2.4 presents the main results. Section 2.5 discusses some channels through which displacement works and Section 2.6 concludes.

2.2 Literature Review

The availability of micro-data from a large number of developing countries has greatly increased the possibilities of researchers to investigate the effects of violent conflicts at the individual and household level. A large number of papers look at the effects of exposure to violence on school attainment (eg. Akresh, and de Walque, 2011; Swee, 2009; Shemyakina, 2011; Chamarbagwala, and Morán, 2011; Merrouche, 2006; Rodriguez, and Sanchez, 2009). Most of these studies find a negative effect for at least a subgroup of the population even after the end of the violent spell. Justino (2011) surveys this literature. However, the panel data approach by Pivovarova and Swee (2012) suggest that the findings could be driven by selection into victimization. In particular, Swee (2009) analyzes the case of Bosnia and Herzegovina and finds lower completion rates in secondary school. However, he argues the effect is driven by youth soldiering and should not have long-run consequences. Since displacement is a direct consequence of potential exposure to violence, my findings contribute to the understanding of why educational production is affected even after the end of the violence.

Exposure to violence also has impacts on peoples' preferences. Bellows and Miguel (2009) find that it increases civil participation, while Voors et al (2012) find that preferences towards risk, the discount rate, and altruistic behavior towards neighbors are affected.

The effects of displacement have also been studied in a number of papers. Fiala (2009) finds that displaced households have in general fewer assets and lower consumption quality. Only the previously poorest are marginally better off after displacement. Lower asset holdings of displaced households are also confirmed by Rahim, Jaimovich, and Ylönen (2013). Kondylis (2010) shows that displacement makes men more likely to be unemployed and women more likely to be out of the labor force in the case of Bosnia and Herzegovina.

However, in the long-run Sarvimäki, Uusitalo, and Jäntti (2009) find positive effects on income for displaced agricultural workers in Finland, which is confirmed by Bauer, Braun, and Kvasnicka (2013) for Germany. However, the latter find lower income for the rest of the displaced even for the second generation.

With the exception of Bauer, Braun, and Kvasnicka (2013), the above mentioned contributions focus on the effect on the generation of the displaced, however the negative consequences of displacement do not need to end there. This is where my findings contribute to the literature on displacement, as the consequences of lower investment in education of children will show its full effect only in the future.

However, the channel through which exposure to violence and displacement work is less studied than the total effects they have. But the knowledge of those channels is exactly what would be of great interest to policy makers as policies that affect the channels are likely easier to implement than policies that stop wars and displacement.

The destruction of schools during periods of violence often fails to explain the reduction in educational attainment as that would affect boys and girls in the same way, which is often not found in the data (Shemyakina, 2011). Swee (2009) argues that youth soldering has prevented older males from attending school, while Shemyakina (2011) mentions that girls were kept at home to avoid sexual assaults and harassment on their way to school. León (2012) finds evidence that the death of teachers and the health status of parents could be the causal mechanism. As already mentioned, in the case of Nepal, Pivovarova and Swee (2012) attribute the finding of reduced educational attainment in a cross-section to self-selection into victimization.

The positive long-run effects of displacement found in Sarvimäki, Uusitalo, and Jäntti (2009) and Bauer, Braun, and Kvasnicka (2013) are likely due to increased sector and geographic mobility of the displaced population.

2.3 Context, Data, and Identification

2.3.1 The Bosnian War

Bosnia and Herzegovina became independent in the Fall of 1991 after the breakdown of former Yugoslavia. The three major ethnic groups, Bosniaks, Serbs, and Croats, were struggling to gain power and eventually the conflict turned violent in April 1992. Initially all three ethnic groups were fighting each other, with the Serbs in control of the army of the former Yugoslavia (Silber & Little, 1996, p.222). In February 1994, Croats and Bosniaks reached a peace agreement and eventually joined forces against the Serbs. With air support from the NATO, Bosniaks and Croats were able to regain control of large areas and push the Serbs back. The Bosnian War ended in December 1995 with the Dayton Agreement, according to which Bosnia and Herzegovina was divided into two entities along the front line at the end of the war: the *Federation of Bosnia and Herzegovina* and the *Republika Srpska*. Now most Bosniaks and Croats live in the Federation of Bosnia and Herzegovina and most Serbs in the Republika Srpska. Those two entities are like separate states with their own administration and they currently cooperate only in a few areas. During the war about 100,000 - 110,000 people were killed and an estimated 1.3 - 1.8 mio. were displaced (the total population in 1991, the year of the last official census, was 4.38 mio.).

2.3.2 The Education System in Bosnia and Herzegovina

The education system in Bosnia and Herzegovina faces many challenges and changes these days. As of 2001, primary education lasts for 8 years, where during the first 4 years the entire material is taught by one teacher and in grades 5-8 each subject has its own teacher. Secondary education is divided into vocational training and gymnasium (more academically oriented), where curricula are taught in 3-5 year programs (UNESCO, 1996; UNESCO, 1997). In general, primary schools have no annual tuition, but textbooks, school materials, etc. still need to be payed for by the parents. Only few municipalities ensure that textbooks for disadvantaged are provided (OECD, 2006). Low or non-existent incomes, migration, and difficult post-war conditions are common reasons why parents are unable to be “active parents" and fail to provide school equipment, supplies, and textbooks (UNESCO, 1996).

The post-war financial situation for schools in Bosnia and Herzegovina was constrained, to say the least, as this paragraph from a report about the education system in the Republika Srpska illustrates (UNESCO, 1997, p.ii.):

> "Primary education is, in theory, free, and is financed from the government budget. In practice the government is often unable to pay salaries, and school repairs have often become the responsibility of the municipal authorities. At the secondary level the central government is expected to pay the salaries of personnel, and the municipality all other charges. It was reported that in December 1996 public sector employees, including teachers, had not been paid for 4 to 5 months. The education system today is largely dependent on financial sacrifices made by teachers and families. Textbooks, for example, are extremely expensive: an average primary school text costing DM 1-3.4 [DM = Convertible Mark] and a secondary one as much as DM 7."

Reports show that the curriculum for primary school is designed for a child equipped with 10 textbooks per grade. For most parents that seems hardly affordable, given that a qualified teacher earned in 1996 only 120 Convertible Mark per month (UNESCO, 1996) and unemployment is high.[3] In the years following the publication of these reports, some reforms concerning the curriculum took place and in 2004 primary school was extended to 9 years (Swee, 2010). International aid has certainly improved some issues, but it is unclear if this reduced the parents' financial burden of having children in school. A project report on the education reform in Bosnia and Herzegovina by the European Union from 2008 observes (EU, 2008): "The education reform process evolves at an uneven and slow pace."

2.3.3 Data

For this study I use household survey data from the "Living Standard Measurement Survey" (henceforth LSMS) (State Agency for Statistics of BiH et al) of Bosnia and Herzegovina. The data collection started in 2001 in 25 municipalities with about 5,400 households. From

[3] Detailed reports about conditions at schools during the academic year 2000-01 are not available.

2002-2004 about half of the households were reinterviewed each year to form a 4-year panel data set. The LSMS covers a wide range of topics. The different sections ask questions about housing, education, health, labor, credit, migration, and social assistance. There are also sections on consumption, household businesses, and agricultural activities. For most of this paper, I use the cross-sectional data from 2001. I do this for two reasons: First, the sample size is reduced in the panel data to half the number of households, and second, the 2002-2004 interviews cover a limited number of topics. I will go into more detail about data issues in the respective sections.

2.3.4 Identification

In order to seriously estimate an effect of displacement on household behavior, the treatment of displacement must be random, ie. ex ante, households in both, treatment and control group, do not differ in their characteristics, neither observable nor unobservable. Given that displacement is a form of migration, this is a strict requirement to satisfy. In this section I argue that the Bosnian War provides a rare, although sad, opportunity to study this by-product of violent conflict. I think of displacement of a version of migration, where the instinct of self-preservation dominates all economic considerations. The single most important push-factor is survival, while pull-factors of certain destinations do not pose a problem to the estimation of a causal effect, as municipality fixed effects are always included in the regressions.

My identification strategy, in a nutshell, is that in Bosnia and Herzegovina ethnicities were mixed before the war. The Bosnian War introduced a line of division along which two ethnically homogeneous territories emerged. The front line was not drawn by economic motives and people did not sort themselves into displacement but were forced by their instinct of self-preservation. The random course of the line of division and the absence of self-selection allow me to identify the effect of displacement on household behavior by comparing households that crossed the front line to households that did not move during the war.

At the time when Bosnia and Herzegovina was part of former Yugoslavia, the population was a mix of Bosniaks, Serbs, and Croats in most municipalities. The ethnic conflict in Bosnia and Herzegovina between 1992 and 1995 caused many people to leave their home and take refuge on the other side of the front line. Table 2.1 describes this homogenization of ethnicities in the two entities of Bosnia and Herzegovina. The share of Serbs in the Federation of Bosnia and Herzegovina shrank to 2.3% from 17.6%, while the share of Bosniaks (Croats) went down to 2.2% from 28.1% (1% from 9.2%). Serbs were leaving from the Bosnian/Croatian side of the front line to the Serb side and Bosniaks/Croats the other way round.

Table 2.1: Main Ethnic Groups in Bosnia and Herzegovina

Bosnia and Herzegovina (Overall)

	Bosniaks	Serbs	Croats
1991	43.5%	31.2%	17.4%
1996	46.1%	37.9%	14.6%

Federation of Bosnia and Herzegovina

	Bosniaks	Serbs	Croats
1991	52.3%	17.6%	21.9%
1996	72.5%	2.3%	22.8%

Republika Srpska

	Bosniaks	Serbs	Croats
1991	28.1%	55.4%	9.2%
1996	2.2%	96.8%	1.0%

Source: Official census in 1991 and unofficial census conducted by the UN in 1996.

During the war, Bosniaks and Croats in the Serb territory were at risk of being killed, what became to be known as "ethnic cleansing". A main goal of Serb forces was to create an ethnically homogeneous territory within Bosnia and Herzegovina. Serbs beyond the front line faced a similar fate and were abandoning whole villages within a few days (Silber & Little, 1996, p.358). Even after the Dayton Agreement was signed, displacement did not come to a halt. Several villages in the Federation of Bosnia and Herzegovina and suburbs of Sarajevo are reported to have been abandoned after the local Serbs realized they were trapped in Bosniak territory (Silber & Little, 1996, p.30). Thus displacement during and after the war produced for the most part ethnically homogeneous regions.

Of course displacement status is not a question in the LSMS and has to be inferred from the data. The data include the municipality of residence right before the war and the location of residence in 2001. When a person did not move at all, he/she enters the control groups of non-movers. If a person resided on the other side of the front line before the war than he/she does in 2001, the person is considered displaced. People who changed their location of residence but did not cross the front line are dropped as are people who returned to their pre-war residence. A household enters either the control or treatment group when both, the household head and his/her spouse, did not move or got displaced.

Note that the identification strategy implies the ethnicity of a person. A person living in the current Republika Srpska before the war and is now living in Federation of Bosnia and Herzegovina, the person is considered to be a displaced Bosniak. Conversely, a person, who has moved from the Federation of Bosnia and Herzegovina before the war to the Republika Srpska after the war, is considered to be a displaced Serb. However, information about the ethnicity of a person is not available in the first wave of the data set, but only in the smaller second wave. I can test this prediction of the ethnicity by the migration pattern only on this smaller sample. Table 2.2 reports the ethnicities of the groups of non-movers and displaced people. Out of 1,040 displaced individuals, there is one Croat and no Serb in the Federation of Bosnia and Herzegovina and one Bosniak and one Croat in the Republika Srpska. The ethnicity mix in the sample of non-movers is not as clear-cut, which probably originates in the presence of enclaves in both, the Republika Srpska and the Federation of Bosnia and Herzegovina. However, these enclaves do not pose a threat to my identification, because these people did not got "treated" by forced displacement, neither did they endogenously decide to migrate. This evidence is a strong argument in support of my selection strategy.[4]

Table 2.2: Displacement by Ethnic Groups

		After-War Entity of Residence	
		Federation of Bosnia and Herzegovina	Republika Srpska
Displaced	Bosniak	285	1
	Serb	0	747
	Croatian	1	1
	Other	5	0
Not Moved	Bosniak	2,847	59
	Serb	121	2,306
	Croatian	614	22
	Other	108	39

Ethnicity data are from the second wave of the the LSMS on Bosnia and Herzegovina, which only includes half of the household from the first wave. The number of observations is reported.

[4]Croats are hardly found as displaced people in the data, because households of Croats who found them self in Serb territory likely moved to the - then newly formed - Republic of Croatia (IDMC, 2009). I will therefore focus on Bosniaks and Serbs from now on.

The first main threat to the identifying assumptions would be a systematic course of the front line so that economically different areas were divided. If economically undeveloped areas were targeted and conquered during the war, the displaced households were poorer than their non-mover counterparts already before the war. Kondylis (2010, p.241f) discusses this possibility and concludes that the war was “determined more by geo-strategic motives rather than economic motives.” The Serb invasion followed the goal to connect the Serb stronghold around Banja Luka with the Serb nation. In addition, Kondylis (2010) provides evidence that pre-war educational attainment is uncorrelated with war casualties (and hence developed areas were not more or less contested) and there is no clear connection between the pre-war ethnic mix and war intensity.

The second key assumption implies that, ex ante, displaced households were not different to non-mover households. One possibility is that households of a certain type could have moved into areas, where they were especially exposed to the risk of displacement. This includes, for instance, a Bosniak family moved to Banja Luka (now the capital of the Republika Srpska) before the war so the household head can take a good position there. The data does not suggest evidence of sorting before the war in Bosnia and Herzegovina. 70% of the household heads in the data still lived in their municipality of birth just before the war. Considering the small size of the average municipality in Bosnia and Herzegovina of 373 km^2, this suggests that households generally do not exhibit high mobility. Moreover, the results of the paper are not sensitive to restricting the sample to households who still resided in their birth municipality just before the war. If the results are purely driven by pre-war sorting, the findings would vanish in such a selected sample.

Another potential problem is that households with certain characteristics were displaced, while others at the same location and of the same ethnicity did not move, and hence displacement was not a life-saving treatment as suggested earlier. In terms of observable characteristics displaced and non-mover households do not differ as the descriptive statistics in Table 2.3 show. While Serbs were more likely to be displaced, there is no significant difference in other household characteristics. There is also no difference of the distributions of the highest level of education achieved and the age of the household head using a Kolmogorov-Smirnov equality of distributions test.[5] The column "Regression" reports

[5]This is also true for the larger sample of all household heads and their spouses in the LSMS data set, not just the ones with children in school that is used here.

Table 2.3: Descriptive Statistics and Exogeneity of Displacement

	All	Not Moved	Displaced	Difference	Regression
	(1)	(2)	(3)	(4)	(5)
Displaced Family	0.189	0	1		
Republika Srpska	0.44	0.40	0.61	0.21***	
				(0.037)	
Number of Children of Family in School	1.81	1.81	1.81	0.00	0.09*
				(0.056)	(0.054)
Oldest Child	0.41	0.41	0.40	-0.01	-0.01
				(0.023)	(0.025)
Female Child	0.47	0.47	0.48	0.01	0.02
				(0.028)	(0.030)
Education of Household Head	10.30	10.34	10.10	-0.25	-0.22
				(0.279)	(0.305)
Age of Household Head	45.74	45.80	45.47	-0.33	-0.87
				(0.735)	(0.834)
Municipality FE					Yes
Observations	1,952	1,584	368	1,952	1,952

Significance: $*$: 10% $**$: 5% $***$: 1%
Standard errors clustered at the household level in parenthesis.
The column "Regression" reports the point estimate of a regression of the household characteristic on the displacement indicator and municipality fixed effects.

the point estimate of a regression of the household characteristic on the displacement indicator and municipality fixed effects. However, regional differences are not a problem as municipality fixed effects are included in all regressions.

Table 2.4 tests whether the pre-war household characteristics are jointly correlated with displacement. In a regression of the displacement indicator on the characteristics, only the dummy variable for the Serb territory is a good predictor of displacement. All the other variables are unable to significantly explain displacement. A joint test of the five household characteristics can not reject the hypothesis that displacement is a selective process. The results do not change when municipality fixed effects are included in column 2.

A final concern to identification is international migration/displacement. The Ministry for Human Rights and Refugees (MHHR, 2003) reports the number of international refugees as 1.2 million between 1992 and the end of the war, which is more than a quarter of the total population in Bosnia and Herzegovina before the war. About half of this group returned to their home country until 2003. This is potentially a threat to the identification strategy if the families who left Bosnia and Herzegovina during the war have a characteristic that is

Table 2.4: Regression Output: Displacement

	Dep. Variable: Displaced Family	
	(1)	(2)
Republika Srpska	0.136***	
	(0.0239)	
Number of Children of Family in School	-0.000	0.021
	(0.0150)	(0.0148)
Oldest Child	-0.011	-0.009
	(0.0154)	(0.0140)
Female Child	0.008	0.011
	(0.0080)	(0.0110)
Education of Household Head	-0.004	-0.004
	(0.0038)	(0.0036)
Age of Household Head	-0.002*	-0.002
	(0.0012)	(0.0012)
Municipality FE		Yes
Observations	1,952	1,952
R-squared	0.031	0.162
F-value	0.70	1.19
(p-value)	(0.622)	(0.310)

Significance: $*$: 10% $**$: 5% $***$: 1%
Standard errors clustered at the household level in parenthesis.
The F-value (p-value) corresponds to a test of the joint hypothesis that the effects of the five control variables (not the geographic control for the Republika Srpska) are zero.

different to the people who stayed in the country and the treatment and control groups are unequally affected.

For this purpose I compare the displaced Serbs in Bosnia and Herzegovina to the residents of the Republic of Serbia who lived in Bosnia and Herzegovina in 1991 using the LSMS of Serbia in 2002. Under the assumption that Serbs living in the future Federation of Bosnia and Herzegovina (the Bosniak territory) before that war had the choice between internal displacement and displacement/migration to the Republic of Serbia. I find that household heads and their spouses in the Republic of Serbia have on average 1.4 years more of education and are 4 years younger (results not shown).

These numbers suggest that the internally displaced are different from the externally displaced, which confirms the finding of Kondylis (2010) for migration to Western European countries. However, the fact that within Bosnia and Herzegovina the displaced and the non-movers are indistinguishable in terms of education and age (Table 2.3) indicates that my treatment and control groups were not unequally affected by internal displacement/migration and international displacement/migration is therefore not a problem for identification. The problem of international migration/displacement is common in the literature on conflicts, because micro data set usually restrict the sample to a country.

2.4 Main Results

The treatment effect of displacement on education expenditure is estimated by OLS of the estimating equation

$$y_{i,j} \quad = \quad \delta d_{i,j} + X_{i,j}\beta + \eta_j + \varepsilon_{i,j}, \tag{2.1}$$

where $y_{i,j}$ is the log. of education expenditure on child i in municipality j, $d_{i,j}$ is the displacement indicator, $X_{i,j}$ are exogenous control variables (ie. exogenous in the sense that they are not influenced by displacement), η_j is the municipality fixed effect, and $\varepsilon_{i,j}$ is an error term. The standard errors are clustered at the household level to account for intra-household correlation.

The selection issues discussed in the previous sections would be a problem if $d_{i,j}$ and $\varepsilon_{i,j}$ are correlated. However, the discussion showed that many issues can be ruled out. I also employ nearest neighbor matching as an alternative estimation method to check the robustness of the OLS findings.

2.4.1 Preliminary Evidence

In a first step, I plot the Kernel density function of education expenditure of displaced and non-mover households against each other. Panel (a) in Figure 2.1 shows the unconditional densities while Panel (b) depicts deviations from municipality means. It is clearly visible that the whole distribution of the displaced households is shifted to the left of the distribution of the non-mover households, both, unconditionally and in deviations from municipality means.

Figure 2.1: Difference in Education Expenditure of Displaced vs. Non-MoverHouseholds

(a) Unconditional Difference (b) Within-Municipality Difference

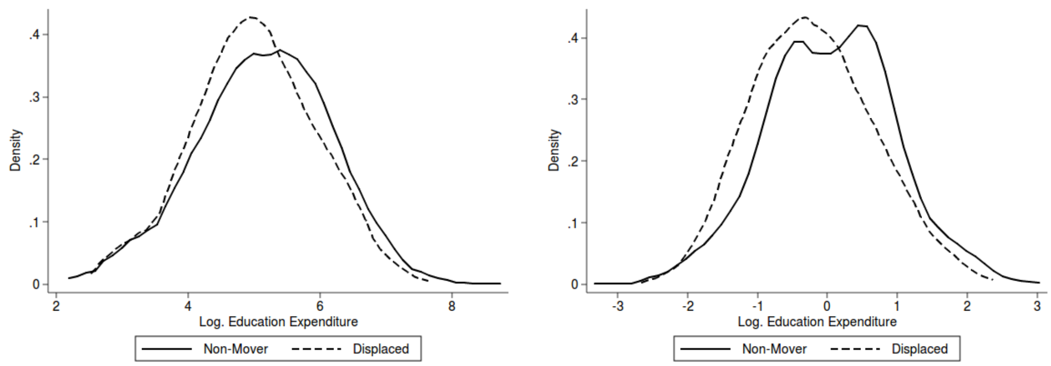


Table 2.5 reports descriptive statistics for the various expenditure groups per child in Convertible Mark, the local currency. In a regression of the expenditure category on the displacement indicator and municipality fixed effects, only few categories show a significant difference. A simple difference of means between non-movers and displaced households would not be informative due to regional variation in education expenditure, therefore the regression.

Education expenditure is not a trivial part of total expenditure for households in Bosnia and Herzegovina. The average annual expenditure on education per child is 267.1 Convertible Mark, while the average total household income per month in the data set is 481 Convertible Mark. This relates to 4.6% of an average total household income is spent on the education of each child. This a high price to pay for generally “free” education, which causes Mooney and French (2005) to suggest financial support for the education of children of displaced households.

In most specifications I use the sum of all expenditure classes, because of the group called “Total cost (not included in previous classes)”. This expenditure class forms a pool for expenses, that parents cannot classify or do not bother to split up into the exact groups. The problem is that the group “Total cost (not included in previous classes)” is negatively correlated with all other groups, which suggests that some households do not take the effort to split up the expenditures into the various classes and put everything into this group. Since dropping such households would reduce the sample size considerably, I use the sum of all groups in the main specifications to avoid the loss of many observations. In additional

Table 2.5: Descriptive Statistics: Education Expenditure by Classes

	Mean	Stand. Dev.	Regression
	(1)	(2)	(3)
Annual Tuition	12.1	(58.9)	-3.6 (2.39)
Special Tuition	2.1	(25.2)	1.6 (1.99)
Membership Fee for Parent's Association	0.7	(16.5)	-0.1 (0.22)
School Uniforms and other School Clothing	36.5	(89.4)	-7.7 (4.98)
Textbooks	35.3	(53.6)	-2.5 (4.51)
Other School Materials	31.5	(34.5)	-2.4 (2.55)
Food and Lodging	44.2	(104.0)	-16.3*** (5.7)
Other Costs	19.9	(71.9)	-13.6*** (4.27)
Total Costs (not included in previous classes)	84.9	(275.5)	-28.8** (14.62)
Expenditure on Education (sum of all groups)	267.1	(316.4)	-73.4*** (17.18)
Municipality FE			Yes
Observations			1,952

Significance: $*$: 10% $**$: 5% $***$: 1%
Standard errors clustered at the household level in parenthesis.
Values are denoted in *Convertible Mark*. The column "Regression" reports the coefficient of a regression of the expenditure category on the displacement indicator and municipality fixed effects. A direct comparison of means between non-movers and displaced households is not informative due to regional variation in education expenditure.

specifications, I restrict myself to a number of selected groups and drop the households that use the class "Total cost (not included in previous classes)".

2.4.2 Total Causal Effect

In this section, I present the estimation results for the total causal effect of displacement on education expenditure. The robustness checks show that the difference in spending on education holds across various specifications. An exact interpretation of dummy variables in semilogarithmic regressions is provided in the row "Transformation" following van Garderen and Shah (2002). However, the exact interpretation deviates from the regression coefficients only slightly and also significance is unchanged.

The main results are reported in Table 2.6: the findings indicate a highly significant and robust drop in education expenditure in all specifications. Quantitatively, the difference in education expenditure between displaced parents and parents that did not move during the war is in the ballpark of 25 to 30% depending on the specification. Column (1) reports the difference controlling only for the entity of residence. Including control variables for the children and parents (column (2)) hardly changes the estimated effect. The inclusion of municipality fixed effects increases the effect to a difference of almost 30%, which is identical between the two entities as column (4) shows.

Column (5) tests whether displaced parents discriminate between boys and girls, but shows that there is not significant difference. In column (6) I interact the displacement variable with the secondary school variable to test whether the difference in education expenditure originates in primary or secondary school. The point estimate of the interaction term is zero and hence displaced parents spend less in both primary and secondary school.

In column (7), I control for the grade of school the child is in. This variable might be influenced by displacement, as children of displaced parents might only attend lower grades and therefore education might be cheaper. Then the difference in education expenditure would not be a decision of the parents, but imposed by the system. However, when the grade of the school is included in the regression, the point estimate increases to a difference of almost 33%.

Table 2.7 reports the results of the nearest-neighbor matching. Matching is considered to reduce selection issues as individuals of ex ante comparable observable characteristics who differ only by treatment status are matched in pairs and compared to each other. The average treatment effect in column (1) is a difference in education expenditure of 33%, while the average treatment effect on the treated in column (2) is 44%. These numbers confirm the OLS results and are actually larger than those. These results are robust to an increase in the number of neighbors the treated observations are matched to. Column (3) includes only an indicator for the Republika Srpska and finds a difference of 22%, which is comparable to the OLS estimate.

The inclusion of regional control variables pose a violation of the assumptions of matching where the matching variables must be unaffected by the treatment. However, displacement affects the location of residence by definition. Column (4) tests the case without regional control variables and finds a similar significant result. Hence, if matching is better able to deal with selection into treatment than OLS and it produces similar results, selection does not seem to drive the main findings.

A closer look at some education expenditure groups is taken in Table 2.8. For this table, I restrict the sample to the 1,325 children, whose parents split up all their costs to the detailed expenditure groups and did not use the category “Total Costs (not included in previous classes)”. Including a child which zero reported expenditure on textbooks, for instance, and a single position in the group of unclassified expenditures would lead to unreasonable results in these regressions.

In primary school (grades 1-8) only few parents pay annual tuition and as regression (1) shows, there is no significant difference in spending. However, in secondary school, where areas of specialization are offered, there is a large and significant difference between children of displaced and non-mover parents. Regression (2) shows a reduction by about 80%. In terms of other school material, which includes notebooks, pencils, etc., there is a difference of 17.1%. The spending on textbooks in column (4) is conditional on positive spending on textbooks by anyone in the municipality, as in some municipalities textbooks are provided by the municipality or the federal government. The difference is still a significant 21.4%. Adding up these three groups, which seem to be especially important for the quality of education, a difference of 14.6% is estimated.

These results suggest that displaced parents restrict expenditures on the education of their children wherever they can, that is even in matters like the choice of the secondary school and the provision of study materials.

The specifications in Table 2.9 perform some robustness checks. If someone outside the household paid for education expenditures and that is the reason why parents spend less, the difference in education expenditure could be inconsequential for educational output. Das et al (2013) show that anticipated public supply of additional school inputs in India and Zambia is offset by an expenditure reduction of parents. Fortunately, the LSMS records the expenses paid from someone outside the household, however only as a total amount. Including these expenditures and running the baseline regressions with the new dependent variable in column (1) reduces the difference in education expenditure to 23.8% but remains highly significant. Restricting the sample to the households without any outside funds for education in column (2) shows a slightly increased difference of 30.6%. These findings imply that the reduced education expenditure is not driven by displaced households who are substituting for additional external funds.

In column (3) dummy variables for rural and mixed municipalities are interacted with the displacement dummy, while the base group are urban municipalities. The interaction terms produce positive, but insignificant point estimates, while the difference in the base group is estimated at 37%. When the sample is restricted to households that still lived in their municipality of birth just before the war, one can rule out that some type of household exposed themselves to a higher risk of displacement by migration before the war and now drives the main finding. This can be ruled out as the restricted sample, that did not migrate before the war, also spends 29.1% less on the education of their children.

To sum up, I find a strong negative relationship between displacement and the spending on the education of their children. The result is robust to a number of sample restrictions and the inclusion of various control variables.

Table 2.6: Regression Output: Education Expenditure I

	Log. of Education Expenditure						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Displaced Family	-0.250***	-0.240***	-0.294***	-0.299**	-0.346***	-0.299***	-0.329***
	(0.0699)	(0.0697)	(0.0670)	(0.1187)	(0.0763)	(0.0722)	(0.0663)
Displaced Family* Republika Srpska				0.009			
				(0.1434)			
Displaced Family* Female Child					0.111		
					(0.0977)		
Displaced Family* Secondary School						0.000	
						(0.1197)	
Secondary School						0.285***	
						(0.0558)	
Grade of School							0.074***
							(0.0074)
Republika Srpska	0.461***	0.455***					
	(0.0536)	(0.0540)					
Number in Children of Family in School		0.083*	0.030	0.030	0.030	0.021	-0.003
		(0.0430)	(0.0347)	(0.0347)	(0.0346)	(0.0345)	(0.0344)
Oldest Child		0.236***	0.246***	0.245***	0.244***	0.205***	0.153***
		(0.0426)	(0.0394)	(0.0394)	(0.0394)	(0.0402)	(0.0407)
Female Child		0.058	0.064	0.064	0.043	0.069*	0.065
		(0.0444)	(0.0412)	(0.0412)	(0.0470)	(0.0410)	(0.0402)
Education of Household Head		0.014*	0.021***	0.021***	0.021***	0.018**	0.016**
		(0.0084)	(0.0079)	(0.0080)	(0.0079)	(0.0079)	(0.0077)
Age of Household Head		0.008***	0.010***	0.010***	0.010***	0.007***	0.004
		(0.0029)	(0.0027)	(0.0027)	(0.0027)	(0.0027)	(0.0027)
Transformation	-22.29***	-21.56***	-25.64***	-26.39***	-29.47***	-26.03***	-28.21***
	(5.423)	(5.458)	(4.974)	(8.705)	(5.376)	(5.334)	(4.753)
Municipality FE			Yes	Yes	Yes	Yes	Yes
No. Observations	1,952	1,952	1,952	1,952	1,952	1,952	1,952
R-squared	0.054	0.071	0.201	0.201	0.202	0.215	0.244

Significance: $*$: 10% $**$: 5% $***$: 1%
Standard errors clustered at the household level in parenthesis.
The line "Transformation" reports approximate unbiased estimator of the percentage change of a dummy variable in a semilogarithmic regression (Kennedy, 1981) and its standard error following van Garderen and Shah (2002).

Table 2.7: Regression Output: Nearest-Neighbor Matching

	Log. of Education Expenditure			
	(1)	(2)	(3)	(4)
Displaced Family	-0.330***	-0.440***	-0.222***	-0.205***
	(0.0822)	(0.0772)	(0.0701)	(0.0653)
Average Treatment Effect	Yes		Yes	Yes
Average Treatment Effect on the Treated		Yes		
Municipality FE	Yes	Yes		
Control for Republika Srpska			Yes	
No. Observations	1,952	1,952	1,952	1,952

Significance: $*$: 10% $**$: 5% $***$: 1%
Standard errors in parenthesis.
Nearest-Neighbor-Matching is performed on the basis of the usual control variables (number of children in school in the family, indicator for the oldest child, indicator for a female child, the education of the HHH, and the age of the HHH). Every treated observation is matched to a single untreated one.

Table 2.8: Regression Output: Education Expenditure II

	Annual Tuition	Annual Tuition	Other School Materials	Textbooks	Important Groups
	(1)	(2)	(3)	(4)	(5)
Displaced Family	0.148	-0.805***	-0.171**	-0.214**	-0.146**
	(0.1371)	(0.2937)	(0.0759)	(0.0910)	(0.0617)
Primary School Only	Yes				
Secondary School Only		Yes			
Controls	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes
No. Observations	984	341	1,325	1,101	1,325
R-squared	0.078	0.278	0.164	0.724	0.428

Standard errors clustered at the household level in parenthesis.
Significance: $*$: 10% $**$: 5% $***$: 1%
Control variables include number of children in HH enrolled in school, education of HHH, age of HHH and dummy variables for being the oldest child and females. Only households with zero expenditure in the residual category ”Total Costs (not includedin previous columns)” are used in all regressions. In the regression ”Textbooks”, only municipalities with some positive expenditures were used. “Important Groups” is the sum of the previous three groups. All independent variables are in measured in logs.

Table 2.9: Regression Output: Education Expenditure III

	Log. of Education Expenditure			
	All Funds	Original	Original	Original
	(1)	(2)	(3)	(4)
Displaced Family	-0.238***	-0.306***	-0.370***	-0.291***
	(0.0723)	(0.0721)	(0.0971)	(0.0830)
Displaced Family * Rural Municipality			0.151	
			(0.1646)	
Displaced Family * Mixed Municipality			0.183	
			(0.1563)	
No Funds from Outside of Household		Yes		
Never Moved before War				Yes
Controls	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes
No. Observations	1,947	1,827	1,952	1,454
R-squared	0.170	0.202	0.202	0.184

Standard errors clustered at the household level in parenthesis.
Significance: $*$: 10% $**$: 5% $***$: 1%
Control variables include number of children in HH enrolled in school, education of HHH, age of HHH and dummy variables for being the oldest child and females. "All Funds" includes education expenditures paid by someone from outside the household. Column (2) restricts the sample to households that do not receive funds for education from outside the household. Column (4) restricts the sample to households who still lived in their municipality of birth before the war.

2.5 Channels

So far the paper has estimated the total causal effect of displacement on education expenditure, but has not offered a channel through which this effect works. However, it is of great interest to researchers and policy makers alike to understand the mechanisms through which displacement operates. I first start with some econometric considerations to guide the analysis.

2.5.1 Econometric Considerations

The estimation of the channel through which displacement works faces two issues: First, including a variable of a channel introduces an endogenous variable and hence OLS produces biased estimates. And second, the proper counterfactual of an evaluation of a channel changes the interpretation of the estimation results.

Evaluating the Indirect Effect

The following two equations describes the effect of an intermediary variable m_i (eg. income) on displacement and the data generating process of education expenditure. Municipality fixed effects are omitted for simplicity.

$$y_i = \delta_1 d_i + X_i\beta_1 + \rho_1 m_i + \theta_1 m_i d_i + \varepsilon_{1i} \tag{2.2}$$

$$m_i = \delta_2 d_i + X_i\beta_2 + \varepsilon_{2i} \tag{2.3}$$

In these equations, displacement has a direct effect δ_1 on education expenditure, changes the intermediary variable through δ_2, and might influence the effect of the intermediary variable on education expenditure through θ_1. Simply including the intermediary variable and its interatction with displacement introduces an endogeneity problem as $Cov(\varepsilon_{1i}, \varepsilon_{2i})$ might be nonzero. In the case of $Cov(\varepsilon_{1i}, \varepsilon_{2i}) \neq 0$, the OLS estimates in equation (2.2) will be biased (Angrist & Pischke, 2009, p.64). However, the bias can be signed: Using income as an expample, it is most likely that $Cov(\varepsilon_{1i}, \varepsilon_{2i}) > 0$, as high income households are more likely to value education more. This causes the estimate of ρ_1 to be biased upwards, ie. the indirect channel of displacement-on-income-on-education expenditure captures too much of the total effect and hence results in a downward bias of δ_1.

Counterfactual

The second problem in these regressions concerns the proper counterfactual. The question I want to answer with these regressions is: "Would the household spend less on education even if displacement would not have changed the intermediary variable?" In a usual regression of education expenditure on a displacement indicator, the intermediary variable, and its interaction with displacement, one has to do some calculations to get a precise answer to the question above and hypothesis testing gets more difficult. However, when the intermediary variable is redefined a single parameter delivers a sufficient answer.

The conditional expectation function is (control variables X_i omitted)

$$E\left[y_i|d_i, m_i\right] = \alpha + \delta d_i + \rho m_i + \theta m_i d_i.$$

Redefine the intermediary variable by subtracting the average of the intermediary variable of the non-displaced households:

$$\tilde{m}_i = m_i - E\left[m_i|d_i = 0\right]$$

Then the expected difference in education expenditure between displaced and non-movers, given no change of the intermediary variable by displacement is

$$\begin{aligned} E\left[y_i|d_i = 1, \tilde{m}_i = E\left[\tilde{m}_i|d_i = 0\right]\right] - E\left[y_i|d_i = 0\right] &= \delta + (\rho + \theta)\tilde{m}_i - \rho\tilde{m}_i \\ &= \delta \end{aligned}$$

since $E\left[\tilde{m}_i|d_i = 0\right] = 0$.

Therefore the transformation of the intermediary variable allows me to estimate the difference in education expenditure under the assumption of no change in the intermediary variable with the single parameter δ. All other estimators remain unchanged because variances and covariances do not change by subtracting a constant.

2.5.2 Income and the Stock of Durables Goods

One of the most natural explanations for the difference in education expenditure would be reduced income and wealth levels of displaced households. It is not surprising that displaced households have lower labor income and wealth than households that did not have to move during the war. The descriptive statistics in Table 2.10 document that displaced households experience a significant reduction in income and the stock of durable goods. Also the share reporting zero labor income of 36.1% is a lot higher than that of non-movers of 19.1%. Higher non-labor income (pensions and allowances) partly counterbalance the reduction in labor income. However, the largest difference is in the stock of durable goods, which can be considered as a proxy for wealth. Lower income and wealth is related to expenditure patterns and it would therefore be an obvious explanation for the estimated difference in education expenditure.

Income and wealth is controlled for with several different variables. Household labor income measures the sum of labor income reported for the last month by household members. Household non-labor income measures the sum of pensions and allowances per month received by household members, while total household income is composed of the sum of the two aforementioned variables. The variable durables is the sum of the values of reported durable goods in the household, but not financial assets or real estate. Dummy variables for a reported value of zero for any of those variables are included in the regressions to make the estimation more flexible.

The results of the regressions are shown in Table 2.11. Each variable is included in four ways: the main variable, a dummy variable for a value of zero, and both interacted with displacement. Remember that the point estimates of the first row estimate the difference to the counterfactual of no reduction of income and/or durables. Also the unexplained part of the displacement effect is a lower bound as discussed above. However, the difference in education expenditure is surprisingly robust to the inclusion of income and durable goods variables. The difference decreases only slightly and rules out income and wealth differences as the main mechanism.

The dummy variable for displacement remains significant at the one-percent level in all specifications. The most flexible specification in column (5) shows a difference in education expenditure of 20.9% after controlling for labor and non-labor income and the stock of durable goods. This is still two-thirds of the total causal effect. Differences in income and the stock of durable goods do not seem to be the main mechanism at work.

Table 2.10: Descriptive Statistics: Income and Durable Goods

	All	Not Moved	Displaced	Difference
	(1)	(2)	(3)	(4)
Total HH Income	481	492	437	-54** (27.3)
Total HH Income (conditional on > 0)	527	539	473	-66** (28.2)
Share with Zero Total HH Income	0.086	0.088	0.075	-0.013 (0.016)
HH Labor Income	404	426	313	-113*** (27.4)
HH Labor Income (conditional on > 0)	521	527	490	-37 (33.9)
Share with Zero HH Labor Income	0.224	0.191	0.361	0.170*** (0.024)
HH Non-Labor Income	77	66	124	58*** (8.9)
HH Non-Labor Income (conditional on > 0)	188	174	229	55*** (15.7)
Share with Zero HH Non-Labor Income	0.590	0.621	0.458	-0.163*** (0.29)
Stock of Durable Goods	2,688	2,884	1,852	-1,032*** (244.7)
Stock of Durable Goods (conditional on > 0)	2,795	3,002	1,915	-1,087*** (251.8)
Share with Zero Stock of Durable Goods	0.038	0.040	0.033	-0.006 (0.011)
Observations	1,901	1,541	360	

Standard errors in parenthesis.Significance: $*$: 10% $**$: 5% $***$: 1%
Income is relates to monthly income. All income and wealth values are in *Convertible Mark*.

Table 2.11: Regression Output: Income and Durable Goods

	Log. of Education Expenditure				
	(1)	(2)	(3)	(4)	(5)
Displaced Family	-0.293***	-0.233***	-0.254***	-0.257***	-0.209***
	(0.0676)	(0.0701)	(0.0683)	(0.0669)	(0.0686)
Log. HH Total Income[1]	0.074*			0.051	
	(0.0401)			(0.0422)	
Displaced Family * Log. HH Total Income[1]	0.035			0.029	
	(0.0788)			(0.0799)	
Log. HH Labor Income[1]		0.042			0.009
		(0.0516)			(0.0555)
Displaced Family * Log. HH Labor Income[1]		0.076			0.072
		(0.1062)			(0.1128)
Log. HH Non-Labor Inc.[1]		0.026			0.023
		(0.0423)			(0.0422)
Displaced Family * Log. HH Non-Labor Inc.[1]		0.006			-0.001
		(0.0800)			(0.0799)
Log. Durable Goods[1]			0.065**	0.052*	0.060**
			(0.0266)	(0.0283)	(0.0289)
Displaced Family * Log. Durable Goods[1]			0.055	0.046	0.015
			(0.0559)	(0.0568)	(0.0589)
Zero Indicator[1]	Yes	Yes	Yes	Yes	Yes
Displaced Family * Zero Indicator[1]		X	X	X	X
F-test	7.51***	4.37***	5.92***	4.18***	2.99***
Controls	Yes	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes	Yes
No. Observations	1,901	1,901	1,901	1,901	1,901
R-squared	0.250	0.253	0.251	0.255	0.258

Standard errors clustered at the household level in parenthesis.
Significance: $*$: 10% $**$: 5% $***$: 1%
Control variables include number of children in HH enrolled, education of HHH, age of HHH and dummy variables for being the oldest child and females. "F-test" results from a test of joint significance of displacement and all its interactions.
[1] Measured as the difference to the mean of non-movers.

2.5.3 Employment Status of Parents

Another mechanism I want to explore is whether differences in the employment status of the parents cause the education expenditure difference. If both parents are working, parents may not have the time to help their children learn and therefore spend more on books and school materials to make up for less personal support (Houtenville and Smith Conway, 2008). The descriptive statistics in Table 2.12 show that displaced parents are less likely to be employed. This mechanism could therefore explain the difference in education expenditure.

I test this hypothesis by including indicator variables for employment status of both parents (or a single parent), the spouse of the household head, and if no parent is employed. Interaction terms of the employment indicators with displacement are included as well. Again the mean of the non-movers of each employment variable is subtracted from the indicator variable to interpret the displacement dummy as the aforementioned counterfactual. In the three following columns I add one dummy variable and the interaction term at a time. None of these coefficients is significant at a traditional level.

As with income, the coefficients show that only a small portion of the difference in education expenditure can be explained by the employment status of the parents, at most about one tenth in the last regression. The main mechanism through which displacement affects education expenditure is still undetected.

Table 2.12: Descriptive Statistics: Parent's Employment Status

	All	Not Moved	Displaced	Difference
	(1)	(2)	(3)	(4)
Both Parents Employed	0.29	0.30	0.23	-0.07***
				(0.026)
Spouse of HHH Employed	0.28	0.29	0.23	-0.06**
				(0.026)
No Parent Employed	0.34	0.30	0.48	0.17***
				(0.027)
Observations	1,901	1,541	360	

Standard errors in parenthesis.
Significance: $*$: 10% $**$: 5% $***$: 1%

Table 2.13: Regression Output: Parent's Employment Status

	Log. of Education Expenditure			
	(1)	(2)	(3)	(4)
Displaced Family	-0.281***	-0.290***	-0.276***	-0.274***
	(0.0687)	(0.0682)	(0.0696)	(0.0689)
Both Parents Employed[1]	0.097			0.056
	(0.0674)			(0.0723)
Displaced Family * Both Parents Employed[1]	0.165			0.203
	(0.1529)			(0.1689)
Spouse Employed[1]		0.095		
		(0.0676)		
Displaced Family * Spouse Employed[1]		0.062		
		(0.1476)		
No Parent Employed[1]			-0.147**	-0.129*
			(0.0705)	(0.0756)
Displaced Family * No Parent Employed[1]			0.017	0.099
			(0.1340)	(0.1470)
F-test	9.74***	9.34***	8.08***	6.14***
Controls	Yes	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes	Yes
No. Observations	1,901	1,901	1,901	1,901
R-squared	0.202	0.200	0.202	0.204

Standard errors clustered at the household level in parenthesis.
Significance: $*$: 10% $**$: 5% $***$: 1%
Control variables include number of children in HH enrolled, education of HHH, age of HHH and dummy variables for being the oldest child and females.
"F-test" results from a test of joint significance of displacement and all its interactions.
[1] Measured as the difference to the mean of non-movers.

2.5.4 Preferences, Uncertainty and Financial Constraints

In summary, neither income and durable goods levels nor the employment status of the parents are able to account for the majority of the difference in education expenditure. The natural question is then: How does the effect come about? Unfortunately, I am not able to fully answer this question. This section discusses some other, only partially testable, explanations and presents some crude evidence.

A possible explanation could be that displaced households are able to buy cheaper school materials and textbooks or share the supplies with other families. Displaced households in an area may build networks to help each other. However, it is hard to imagine that displaced households find a way to save on education expenditure that households who did not move during the war do not find, especially with their social network in place. The non-movers in Bosnia and Herzegovina are not exactly rich to pass up a possibility to save.

Voors et al (2012) present evidence from field experiments in Burundi showing that exposure to violence affects preferences. In detail they report more altruistic behavior towards neighbors, more risk taking, and a higher discount rate. For my purpose the higher discount rate is of special interest. If the same is true for displacement, then displaced parents could have a higher discount rate than parents who were not displaced and would invest less in projects that generate a payoff in the future - such as education. With the consumption data available, such a hypothesis can not be tested rigorously.

Nevertheless, some preliminary tests can be done and those results contradict the prediction of a higher discount rate. With the available annual total consumption and income data I construct a rough savings variable and the savings rate.[6]

Table 2.14 presents the results of some regressions with these variables. All these regressions show that conditional on income and the stock of durable goods, displaced households save more and consume less than their non-mover counterparts. However, an increased discount rate would, ceteris paribus, imply lower savings and higher consumption. Of course these variables contain a lot of measurement error and there are other econometric problems present, but the general pattern does not support the hypothesis that an increase in the discount rate of displaced persons is the reason for decreased spending on education.

[6]The savings rate is savings divided by total household income. Since households with low reported incomes produce large savings rates, I drop households with annual household income of less than 360 CM.

An interesting result in Table 2.14 is that displaced households consume about 11% less than comparable non-mover households. This is a lot less than the 29% I find for education expenditures. Two potential interpretation of this pattern come to mind. First, displaced households face a lot of uncertainty about the future and try to prepare themselves by cutting down spending on every non-vital position, which includes education expenditure. In a simple two period model, an agent with convex marginal utility reduces consumption in the first period if the risk of period 2 income increases. Kimball (1990) calls this phenomenon prudence and defines it as "the propensity to prepare and forearm oneself in the face of uncertainty". In 2001 the restitution of property to internally displaced households and the possibility to return to their homes from before the war was still an issue in Bosnia and Herzegovina. Many displaced households probably faced a highly uncertain future.

Table 2.14: Regression Output: Consumption and Savings

	Log. of Annual Consumption	Savings	Savings Rate
	(1)	(2)	(3)
Displaced Family	-0.113***	1,535***	0.412***
	(0.0220)	(320.0)	(0.1417)
Log. Ann. HH Income	0.107***	3,383***	3.584***
	(0.0119)	(346.0)	(0.2579)
Log. Durable Goods	0.156***	-1,479.6***	-0.742***
	(0.0081)	(128.1)	(0.1018)
Zero Income Indicator	Yes	Yes	Yes
Zero Durables Indicator	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Municipality FE	Yes	Yes	Yes
No. Observations	1,323	1,323	1,200
R-squared	0.572	0.346	0.615

Heteroskedasticity robust standard errors are reported in parenthesis.
Significance: $*$: 10% $**$: 5% $***$: 1%
Control variables include number of adults in HH, the number of children in HH, education of HHH, age of HHH. Outliers with an annual household income of less than 360 Convertible Mark were dropped in regression (4) because of unreasonably small saving shares. The estimate is only sensitive to the chosen income cutoff if a cutoff of 60 Convertible Mark or smaller is used. Other income cutoffs do not produce quantitatively different results.

The second potential interpretation of these results are financial constraints. For many households in the data savings are negative, which might have to do with underreporting income figures. Displaced households could face stricter financial constraints due to the loss of their social network for credit or the lack of property they could use as collateral. Unfortunately, financial assets are not included in the survey to infer about financial constraints.

Even if financial constraints are not binding now, expecting that they will be binding in the future would already make households cut back on expenditures today (Deaton, 1991).

However, I am not able to give a full explanation how the difference in education expenditure comes about. Increased risk and financial constraints are consistent with the pattern above, but changes in preferences and values of the parents can not be ruled out in general.

2.6 Conclusion

This paper contributes to the literature on the long-term consequences of conflict by identifying education as a mechanism through which displacement has a potential negative spill-over on the next generation. I find a robust statistical relationship that displaced parents spend a sizable amount less on the education of their children, that is between 20 and 30%, than parents who did not move during the war. In an environment like Bosnia and Herzegovina, where parents have to pay for textbooks and school materials etc., this difference in spending on education has the potential to negatively affect the quality of education children receive and hence their labor market outcomes. The estimated difference in the annual tuition payed for children in secondary school, indicates that children of displaced parents attend secondary schools of lower quality. The difference in education expenditure is robust to many specifications and a series of tests indicate that a selection bias is not the source of the result. Nearest-neighbor matching also finds quantitatively similar effects.

A number of channels of how displacement can affect education expenditure is tested. Differences in income and durable goods levels can explain at most one third of the baseline result. The employment of parents and support from outside the household can be ruled out as the main mechanisms. Further evidence is consistent with the hypothesis that the displaced households face more uncertainty about the future or more rigid financial constraints than non-movers. That would lead them to cut back spending on every non-vital position, including the education of their children.

More work needs to be done to fully understand how violent conflict influences peoples lives. Research has shown negative consequences of exposure to violence and displacement, but how exactly these changes in economic outcomes come about is not fully understood yet.

Chapter 3

On the Robustness of Inherited Trust and Growth

3.1 Introduction

The relationship between measures of social capital on the one hand, and economic performance on the other hand, has long been the subject of economic research. Empirical research on the topic using cross-country data has produced promising results. Knack & Keefer (1997) and Tabellini (2010), for instance, find trust and civic norms strongly correlated with income after controlling for other determinants of GDP. These results are suggestive, but they fall short of establishing a causal effect of social capital on income and growth: the question remains open whether it is the social capital or some underlying conditions that create and maintain the social capital, that are causing the results.

The recent paper in the AER, "Inherited Trust and Growth" by Yann Algan & Pierre Cahuc, hereafter AC, is a significant contribution to the literature (see the discussion by Guiso, Sapienza, Zingales, 2011) insofar as the authors claim to identify a causal effect of trust on growth. In their study, AC conclude that "[there is] a sizable causal impact of inherited trust on worldwide growth during the twentieth century" (p. 2060). Their point estimate implies that if the share of the population that generally trusts other people would increase by ten percentage points, GDP per capita would increase between USD 2,000 to 3,000.

When estimating the causal effect of trust on economic growth, the following problem arises: it is likely that trust towards others is influenced by income or that trust and income are driven by an unobserved variable. AC overcome this problem by using the inherited part of trust, which should be uncorrelated with contemporaneous income shocks and variation in an unobserved variable, when initial income is controlled for. To do this, it is necessary to estimate historic trust levels from descendants of immigrants to the US from different source countries and arrival periods.

In the present paper, I revisit their identification strategy and argue that the endogeneity problem mentioned above is not solved in the estimated model in AC. The analysis reproduces some of AC's regressions with some changes using the same data and some additional data from the same sources. The data and code were obtained from the AER website and the Maddison database (2007).

Intuitively, the following problem emerges with AC's regression model: the identification strategy rests on the assumption that the trust level of the previous generation, $t-1$, is uncorrelated with any income shock or omitted variable driving income growth up to period t after controlling for initial income. In practice however, AC control for initial income prior to $t-1$, say, $t-2$, in the main specifications. Imagine a large shock in an unobservable variable in $t-1$ that increases the trust level at that time, as well as GDP per capita. Since GDP per capita is persistent and trust is passed on to the next generation, a regression of income in t on trust from period $t-1$ (inherited trust) and initial income in $t-2$ would produce a positive coefficient on the inherited trust variable, even if there is no causal effect of trust on income.

AC seem to be aware of the problem and show regressions in their robustness checks that impose a minimum lag between inherited trust and the outcome variable of 50 and 75 years. This would reduce or eliminate the above discussed problem since initial income is controlled for with a 65 year lag. However, the highly significant results in AC do not hold if two problems are taken care of: including a constant in the regressions and measuring initial GDP for the first period in 1870 for all countries.

Since the main regression outputs in AC are little affected by these two corrections, I explore the robustness of these results. The assumptions of the model specification are varied over a range of values and the estimated effect of trust on income is recorded. As it turns out, there is only a narrow window of assumptions around the specifications in AC producing estimated effects of statistical significance. When the latest rounds of the General Social Survey are included to estimate inherited trust levels from more data no effect significantly different from zero can be found, even in the original specification.

The paper proceeds as follows: Section 2 discusses the unsolved endogeneity problem, while Section 3 corrects the programming error and the data inconsistency and replicates the robustness checks. In Section 4 variations on the main assumptions are tested and Section 5 concludes.

3.2 Endogeneity Problem

The primary challenge when estimating a causal relationship between trust and income is that either trust and income are correlated with an omitted variable or that income influences trust. In what follows I argue that this problem is not solved in AC's main specifications and even if there is no endogeneity problem, the causal effect of interest is not identified.

This section is based on a number of assumptions that I spell out here to facilitate the analysis.

A1. Trust is inherited from the previous generation.

A2. There are temporary shocks that influence income and trust at the same time.

A3. The inherited part of trust is independent of current temporary shocks.

B1. The inherited part of trust is known or can be precisely estimated.

B2. Income at time t is a sufficient statistic for the history up to time t.

Assumption A1 - A3 are taken from AC, while assumptions B1 and B2 are additionally imposed to abstract from other econometric problems. B1 is an assumption I make to avoid dealing with the estimation of historic trust levels. I take this variable as given to focus on the estimation of the causal effect of trust on income. For a discussion on the estimation of historic trust levels see Müller, Torgler, Uslaner (2012). The assumption B2 is implicitly made by AC by controlling for lagged income in most of their regressions.[1] The assumptions above plus linearity can be formulated in the following equations:

$$
\begin{aligned}
Y &= \alpha_0 + \alpha_1 S + \alpha_2 L1.Y + \varepsilon && (3.1)\\
S &= \gamma_0 + \gamma_1 L1.S + \gamma_2 L1.Y + \nu && (3.2)\\
Cov\left(\varepsilon, S\right) &\neq 0\\
\varepsilon &\perp L1.S, L1.Y
\end{aligned}
$$

where Y is income per capita, S stands for trust towards others, ε is an income shock, and $L1$ indicates a lagged variable by one period.

For now I will focus on a cross-sectional case without time or country fixed effects. This is to make the argument and the discussion of the intuition simpler. Appendix B.1 examines the case with panel data and country fixed effects.

3.2.1 Identification

The novel strategy by AC to identify the parameter of interest α_1 is to use the inherited part of trust $\gamma_1 L1.S$ as exogenous variation in S. Let the regression equation be

$$Y = a_0 + a_1\left(\gamma_1 L1.S\right) + a_2 L1.Y + e. \tag{3.3}$$

When the structural equation (3.1) is used and S is substituted for by (3.2), it is easy to derive that

$$Y = \alpha_0 + \alpha_1\gamma_0 + \alpha_1\left(\gamma_1 L1.S\right) + \left(\alpha_1\gamma_2 + \alpha_2\right) L1.Y + \alpha_1\nu + \varepsilon. \tag{3.4}$$

Since $\gamma_1 L1.S$ is uncorrelated with the error term $(\alpha_1\nu + \varepsilon)$, the estimator $\hat{a}_1$ in regression (3.3) is an unbiased estimator of α_1.

[1] AC include a variable X into their model and describe it as "This vector might include the past economic development of the economy with the lagged values for income per capita or education." In most of their regressions AC control for lagged income, which leads to my strong assumption B2 that all history is controlled for if past income is included in the regression.

3.2.2 Regressions in AC

However, in their regressions AC use a higher lag in initial income than they do for inherited trust. Let me illustrate the problem with this strategy by examining this regression model where initial income two periods before is included as a control variable.

$$Y = b_0 + b_1 (\gamma_1 L1.S) + b_2 L2.Y + e \tag{3.5}$$

First, let me explain the intuition of the problem with an example. Assume there are two identical groups of countries, A and B, in $t-2$. Let there be no causal effect of trust on income, ie. $\alpha_1 = 0$. While the countries in group B face no shock of any sort in both periods, $t-1$ and t, there is a war in the countries of group A in period $t-1$. This war constitutes a negative shock to national income and also to the trust level of the countries in A, ie. $Cov(\varepsilon, S) > 0$. There is no shock to any country in period t, but because national income follows a persistent process, income in period t is lower in the countries of group A.

A regression of income in t on inherited trust from period $t-1$ and lagged income from $t-2$, as in AC, would produce a negative point estimate of inherited trust, even though there is no causal effect of trust on income.

For a derivation of the resulting bias use equations (3.1) and (3.2), which describe the data generating process. These equations can be used to substitute in for lagged income $L1.Y$ and trust S.

$$\begin{aligned} Y &= \alpha_0 + \alpha_1 (\gamma_0 + \gamma_1 L1.S + \gamma_2 (\alpha_0 + \alpha_1 L1.S + \alpha_2 L2.Y + L1.\varepsilon) + \nu) \quad (3.6) \\ &\quad + \alpha_2 (\alpha_0 + \alpha_1 L1.S + \alpha_2 L2.Y + L1.\varepsilon) + \varepsilon \\ &= \alpha_0 (1 + \alpha_1\gamma_2 + \alpha_2) + \alpha_1\gamma_0 \quad (3.7) \\ &\quad + \alpha_1 \left(1 + \alpha_1 \frac{\gamma_2}{\gamma_1} + \frac{\alpha_2}{\gamma_1}\right)(\gamma_1 L1.S) \\ &\quad + \alpha_2 (\alpha_1\gamma_2 + \alpha_2) L2.Y \\ &\quad + (\alpha_1\gamma_2 + \alpha_2) L1.\varepsilon \\ &\quad + \alpha_1\nu + \varepsilon \end{aligned}$$

The error term in regression (3.5) is $e = (\alpha_1\gamma_2 + \alpha_2) L1.\varepsilon + \alpha_1\nu + \varepsilon$. This regression faces two problems.

Even if there is no endogeneity problem (ie. $Cov(\varepsilon, S) = 0$) and, thus,

$$E\left[\hat{b}_1\right] = \alpha_1 \left(1 + \alpha_1 \frac{\gamma_2}{\gamma_1} + \frac{\alpha_2}{\gamma_1}\right), \tag{3.8}$$

the parameter of interest α_1 is not identified because γ_1 and γ_2 are unknown. This is an omitted variable problem because $L1.Y$ is not included in the regression equation and $L1.S$ has a direct effect on $L1.Y$ as has $L1.Y$ on S.

Second, if there is endogeneity (ie. $Cov(\varepsilon, S) \neq 0$), then the estimator $\hat{b}_1$ is potentially more biased. It is difficult to express the bias in terms of the structural equations (3.1) and (3.2), but an omitted variable bias argument can determine the sign of the additional bias. In fact,

$$E\left[\hat{b}_1\right] = \alpha_1 \left(1 + \alpha_1 \frac{\gamma_2}{\gamma_1} + \frac{\alpha_2}{\gamma_1}\right) + c_1 \left(\alpha_1 \gamma_2 + \alpha_2\right), \tag{3.9}$$

where c_1 comes from the hypothetical auxiliary regression $L1.\varepsilon = c_0 + c_1 (\gamma_1 L1.S) + c_2 L2.Y + e$ according to the omitted variable bias formula (Greene, 2008, p. 133). Note that c_1 has the same sign as $Cov(\varepsilon, S)$. The term $(\alpha_1 \gamma_2 + \alpha_2)$ is the effect of $L1.\varepsilon$ on Y, where at least α_2 can be plausibly assumed to be positive and sizable, since positive income shocks in the past should increase income today. The effect of a positive income shock in the past on todays trust level is $\alpha_1 \gamma_2$, which should not be negative.

The result suggests that the estimator $\hat{b}_1$ from (3.5) is upward biased except in the special case when $\alpha_2 = 0$ and $\gamma_2 = 0$. This special case implies that there is no persistence in income and past income has no effect on todays trust level.

3.2.3 Country Fixed Effects

The inclusion of country fixed effects does not solve this problem. Those fixed effects account for time-invariant factors of a country, which might be correlated with the trust variable. A temporary shock between the time of measurement of initial income and inherited trust would still bias $\hat{b}_1$. Appendix B.1 derives the estimator for the fixed effect estimation.

3.2.4 Discussion

Almost every regression in AC results in a sizable positive coefficient for initial income, which suggests that income shocks are persistent and $\alpha_2 > 0$. The results in AC should therefore be interpreted with caution since the endogeneity problem is not solved and the estimates might be severely upward biased. The estimated effect would not be the causal effect of trust on GDP per capita, even if there is no endogeneity between trust and income shocks.

The paper by AC focuses on the causal effect of trust on long-term growth. It is therefore not practicable to change the initial income variable to 25 years before the outcome variable, so that the time lags of inherited trust and initial GDP per capita is the same. Also because trust is changing slowly over time. Another way to deal with the problem is already done in the robustness checks in AC. The authors include regressions where the minimum lag between immigration and the measurement of the outcome variable is 50 years and 75 years to potentially address the point raised in this section. In that case the lag of inherited trust and initial income is similar or equal and the bias is reduced or eliminated in the latter case. These regressions show highly significant estimates of the same magnitude than the main results, but these results do not hold if a programming error and a data problem are corrected, as is demonstrated in the next section.

3.3 Programming Error and Data Problem

There are two issues with the regressions in AC, which both change the results of the robustness checks significantly. The first is an inconsistency in the time of measurement of GDP per capita. Income per capita is not measured in the years stated in the paper, but for some observations at different points in time. For instance, the variable "initial income in 1870" is measured for some countries as the difference to Sweden in 1820, while for others it is the difference to Sweden in the period 1900-1909. Table 3.1 lists the periods of measurement of GDP per capita in AC. It is not only initial GDP per capita in 1870 that is measured inconsistently, but also other variables (eg. GDP per capita in 1935-38) have its flaws.

The second point concerns the regressions with country fixed effects. In those specifications

Table 3.1: Maddison Data on GDP per Capita

Variable "GDP per Capita..."		
...in 1910	1900:	Africa
	1900-1913:	Canada, Czech Republic, Denmark, Great Britain, France, Germany, Ireland, Italy, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland
...in 1935	1940:	Africa*
	1920-1938:	Austria, Belgium, Canada, Czech Republic, Denmark, Great Britain, Finland, France, Germany, Hungary, India, Ireland, Italy, Mexico, Netherlands, Norway, Poland, Portugal, Russia, Spain, Sweden, Switzerland, Yugoslavia
...in 1950	1948-1955:	Africa, Canada, Czech Republic, Denmark, Great Britain, France, Germany, Ireland, Italy, Mexico, Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland
...in 2000	2000-2003:	Africa, Austria, Belgium, Canada, Czech Republic, Denmark, Great Britain, Finland, France, Germany, Hungary, India, Ireland, Italy, Mexico, Netherlands, Norway, Poland, Portugal, Russia, Spain, Sweden, Switzerland, Yugoslavia
Variable "Initial GDP per Capita..."		
...in 1870	1820**:	Africa, Austria, Denmark, Great Britain, Finland, France, Germany, Italy, Mexico, Netherlands, Norway, India, Portugal
	1820:	Ireland
	1870-1875:	Belgium, Spain
	1870-1880:	Canada, Hungary, Poland, Switzerland
	1900-1909:	Russia, Yugoslavia
	Unknown:	Czech Republic
	Accordingly:	Sweden
...in 1935	1940:	Africa*
	1920-1938:	Austria, Belgium, Canada, Czech Republic, Denmark, Great Britain, Finland, France, Germany, Hungary, India, Ireland, Italy, Mexico, Netherlands, Norway, Poland, Portugal, Russia, Spain, Sweden, Switzerland, Yugoslavia

*Value used because of data availability
**From print edition

the constant is omitted. The Stata command used is:

```
reg gdppc inh_trust lag_gdppc i.cty, nocons
```

where "gdppc" is GDP per capita, "inh_trust" is the measure of inherited trust, "lag_gdppc" is initial GDP per capita, and "cty" is the country variable. All the variables are measured as the difference to Sweden, so a time fixed effect is not necessary. In this regression there are N-1 country dummies included, but no constant. In a fixed effects regression there should be either N-1 dummy variables and a constant or N fixed effects and no constant. Intuitively this is the same as using the observations of Africa (the country with the lowest country code) as the difference to Sweden, while all other observations are used as the deviation to the country mean. Figure 3.1 illustrates the point using regression 2 from Table 7 in AC as an example.

3.3.1 Replication

Table 3.2: Replication of Table 7 (3)

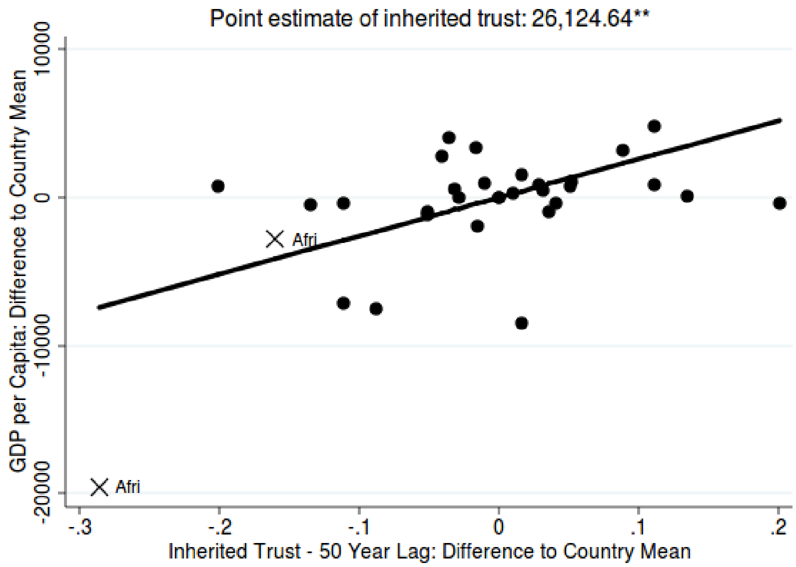
	GDP per Capita in 1935 and 2000			
	(1)	(2)	(3)	(4)
	From AC	New	New	New
Inherited Trust in 1935/2000	14,903.50**	3,291.79	9,620.18	-755.53
Minimum 50 years lag	(6,905.16)	(7,681.39)	(9,408.99)	(8,357.42)
Initial GDP per	4.44***	3.96***	3.45**	2.95**
Capita in 1870/1935	(1.04)	(0.86)	(1.58)	(1.22)
Political Institutions	-71.63	15.61	-223.12	-182.45
in 1930/2000	(123.88)	(117.11)	(220.02)	(206.62)
Country Fixed Effects	X	X	X	X
Constant Included			X	X
Homogenous Measurement of GDP per Capita		X		X
No. Observations	32	32	32	32
R-squared	0.823	0.842	0.812	0.834

Significance: $*$: 10% $**$: 5% $***$: 1%
Standard errors in parenthesis.
All variables are measured as the difference to Sweden.
Column (1) replicates the result from AC for reference with my data.

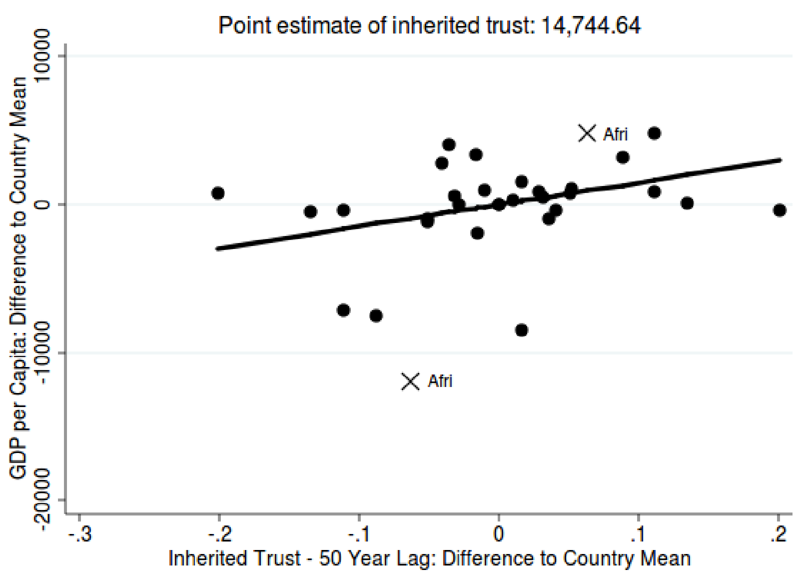
Here I replicate the results of AC by sequentially eliminating the two aforementioned problems. For this task, I use the estimated measure of inherited trust from AC. The new data of GDP per capita is measured as AC describe it in their paper, which is the difference to Sweden of the average of GDP per captia in 1935-38 (2000-03) and initial income is measured in 1870 (1935-1938), with the exception of Africa, where GDP per capita is only available for 1940, which is used instead of 1935-38.

By solving these two problems, the main results of the paper (Table 5 and 6 in AC) do not change much (not shown). But in almost all the robustness checks (Tables 7, 10, and 15 in AC) the point estimates are lower and not significantly different from zero anymore. The focus is on the regressions with country fixed effects and a control variable for initial income, since in this specification both issues come into play.

Figure 3.1: FE Estimation in Table 7 (2)



Panel A: Constant Surpressed



Panel B: Constant Included

Table 3.3: Replication of Table 10 (2)

	GDP per Capita in 1910 and 2000			
	(1)	(2)	(3)	(4)
	From AC	New	New	New
Inherited Trust in 1910/2000	17,286.24**	800.59	18,411.92**	161.91
Minimum 25 years lag	(6,346.56)	(7,225.72)	(8,135.95)	(7,591.42)
Initial GDP per	4.86***	4.15***	5.16**	3.89***
Capita in 1870/1935	(1.22)	(0.87)	(1.81)	(1.07)
Country Fixed Effects	X	X	X	X
Constant Included			X	X
Homogenous Measurement of of GDP per Capita		X		X
No. Observations	30	30	30	30
R-squared	0.772	0.820	0.756	0.809

Significance: $*$: 10% $**$: 5% $***$: 1%
Standard errors in parenthesis.
All variables are measured as the difference to Sweden.
Column (1) replicates the result from AC for reference with my data.

Table 3.4: Replication of Table 15 (3)

	GDP per Capita in 1950 and 2000			
	(1)	(2)	(3)	(4)
	From AC	New	New	New
Inherited Trust in 1950/2000	24,195.65***	17,068.33*	15,285.88*	10,712.33
Minimum 75 years lag	(6,824.92)	(8,787.95)	(8,593.23)	(8,449.07)
Initial GDP per	3.64***	2.88***	1.86	0.15
Capita in 1870/1935	(0.79)	(0.78)	(1.35)	(1.46)
Political Institutions	-25.74	42.97	-224.37	-358.00
in 1950/2000	(91.15)	(107.75)	(152.39)	(211.70)
Country Fixed Effects	X	X	X	X
Constant Included			X	X
Homogenous Measurement of GDP per Capita		X		X
No. Observations	34	34	34	34
R-squared	0.893	0.865	0.881	0.865

Significance: $*$: 10% $**$: 5% $***$: 1%
Standard errors in parenthesis.
All variables are measured as the difference to Sweden.
Column (1) replicates the result from AC for reference with my data.

Tables 3.2, 3.3, and 3.4 show the results of the replications. After the two issues are resolved, none of the estimates remain significantly different from zero at the 10-percent level. For the robustness checks with a minimum lag of 50 years between the measurement of inherited trust and the outcome variable, and when the initial period is changed to 1910, the point estimate are almost zero. But also the point estimate of the model with a 75 years lag decreases by 55 percent. These results suggest that the relationship between inherited trust on GDP per capita might not be especially robust to changes in the specifications.

3.4 Robustness to Varying Assumptions

Considering the failure of the robustness checks in the original paper, I now test to what extent the main result of AC is robust to changes in the various assumptions. In particular, I look at three assumptions: the minimum lag of the measurement of GDP per capita and the point in time that ancestors immigrated to the US of whom inherited trust is estimated, the length of a generation, and the point in time the outcome variable is measured.

In a first set of simulations, I use the estimates of inherited trust from AC. However, as Müller, Torgler, and Uslaner (2012) have pointed out, the estimates of inherited trust in AC are based on few observations for some countries. For this reason in a second set of simulations, I include the latest waves of the GSS from 1972 to 2012 to increase the number of observations in the estimation of inherited trust.

The baseline assumptions in AC are a minimum lag of 25 years between immigration to the US and the measurement of national income, the length of a generation is 25 years, and two periods are 1935 and 2000. AC change two of the three assumptions in their paper to convince the reader of the robustness of their results (see previous section in this note that those do not hold). The simulations in this section are done in the following way: I change one parameter at a time and leave the others at the original values, and estimate the whole model for a range of values. For the minimum lag the range of values is from 0 to 50 years, for the length of a generation from 15 to 40 years and for the measurement of income from 1910 to 1960, with the second period being the minimum of 65 years later or the year 2000. The macro regression is income on the measure of inherited trust and initial income with country fixed effects, where all variables are measured as the difference to Sweden.

The results are presented in the form of a graph. Figure 3.2 uses the same data set (GSS

1972-2002) and procedure to estimate inherited trust as AC. The estimated effect is only significantly different from zero in small windows around the baseline specification in AC (marked with a vertical line). With varying assumptions on the minimum lag or the time of measurement of income, the estimated effect are often close to zero and not significantly different from zero.

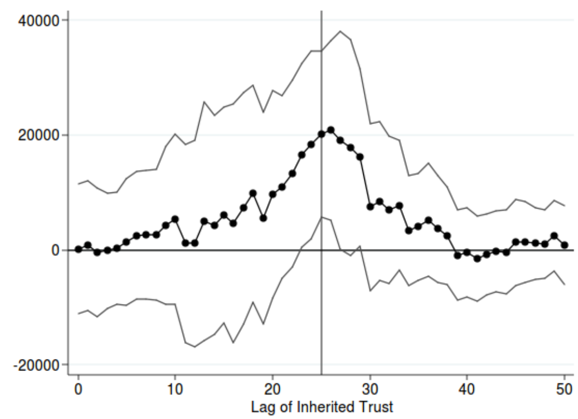
When additional waves from the GSS are included in the estimation of inherited trust, the results are never significantly different from zero, as Figure 3.3 shows. Remember that the only changes made are additional data from the GSS.

3.5 Conclusion

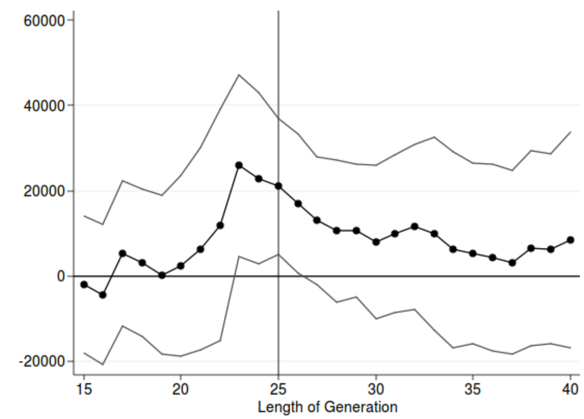
In this note on "Inherited Trust and Growth" (Algan & Cahuc, AER, 2010), I argue that the main results of the paper should be considered with care. First, I argue that the large and significant effect of inherited trust on GDP per capita during the twentieth century can not be interpreted as a causal relationship. The estimation faces an unsolved endogeneity problem. In the regressions AC introduce problems by using inconsistent data and do not include a constant in their regression models. The correction of these problems invalidates most of the robustness checks in the paper.

In a final section, I check the robustness to small changes in the assumptions made by AC. The main results hold true only for a narrow range of assumptions and no statistically significant result in the growth regression can be found when newly available rounds of the General Social Survey are included to estimate inherited trust.

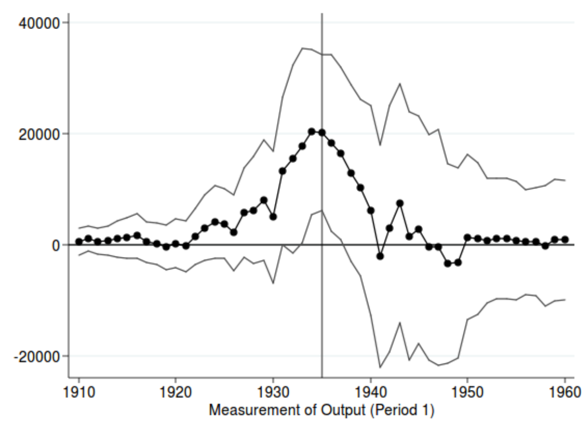
Although the paper by AC is considered an important contribution to the growing literature on the relationship between trust and economic outcomes, this paper encourages the reader to consider those results with care.



(a) Variation in the Time Lag

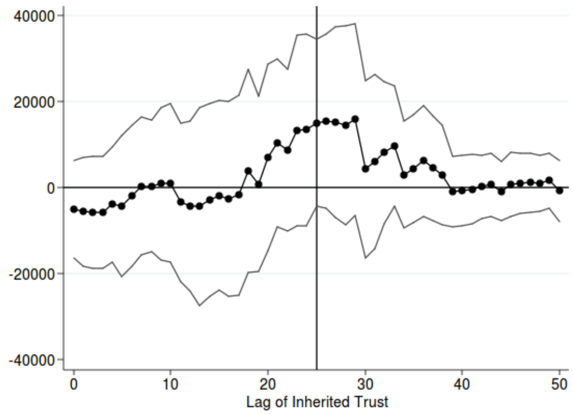


(b) Variation in the Generation Length

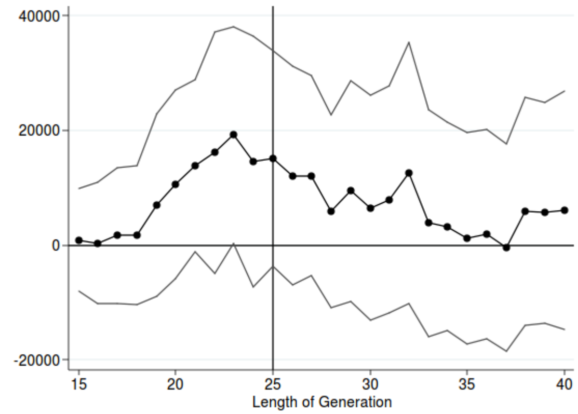


(c) Variation in the Time of Measurement of GDP per Capita

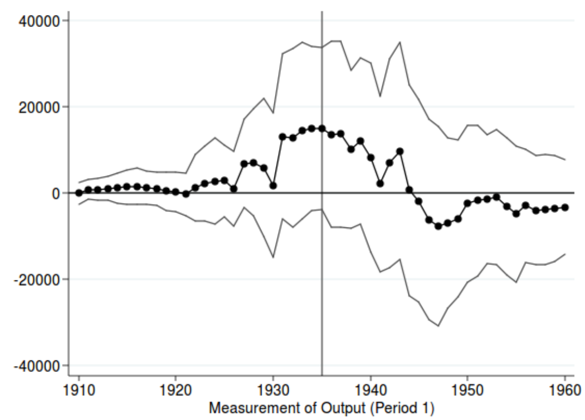
Figure 3.2: Robustness Simulation on Levels of GDP per Capita with Estimation of Inherited Trust as in AC (Point Estimates and 95% Confidence Intervals)



(a) Variation in the Time Lag



(b) Variation in the Generation Length



(c) Variation in the Time of Measurement of GDP per Capita

Figure 3.3: Robustness Simulation on the Level of GDP per Capita with Estimation of Inherited Trust from GSS 1972-2012 (Point Estimates and 95% Confidence Intervals)

Bibliography

Abdel Rahim Asha, Dany Jaimovich & Aleksi Ylnen. (2013). "Forced displacement and behavioral change: An empirical study of returnee households in the Nuba Mountains." *HiCN Working Paper 157.*

Acemoglu, Daron, Tarek Hassan & James Robinson. (2011). "Social Structure and Development: A Legacy of the Holocaust in Russia." *Quarterly Journal of Economics*, 126(2), 895-946.

Akresh, Richard & Damien de Walque. (2008). "Armed Conflict and Schooling: Evidence from the 1994 Rwandan Genocide." *HiCN Working Paper 47.*

Algan, Yann & Pierre Cahuc. (2010). "Inherited Trust and Growth." *American Economic Review,* 100(5), 2060-2092.

Angrist, Joshua & Jörn-Steffen Pischke. (2009). *"Mostly Harmless Econometrics."* Princeton, New Jersey: Princeton University Press.

Bauer, Helfried, Alfred Paleczny & Agnes Schulmeister. (1977.) *Aufgaben der Gemeinden.* (Wien: Institut fr Stadtforschung).

Bauer, Thomas, Sebastian Braun & Michael Kvasnicka. (2013). "The Economic Integration of Forced Migrants: Evidence from Post-War Germany." *The Economic Journal,* Vol. 123(571), 998-1024.

Beevor, Antony. (2012.) *The Second World War.* (New York: Little, Brown and Co).

Bellou, Andriana & Emanuela Cardia. (2013). "Occupations after WWII: The Legacy of Rosie the Riveter." *IZA Discussion Paper*, No. 7615.

Bellows, John & Edward Miguel. (2009). "War and Local Collective Action in Sierra

Leone." *Journal of Public Economics,* Vol.93(11-12), 1144-1157.

Bolognese-Leuchtenmller et al. (1985). *Frauen der ersten Stunde 1945-1955.* (Wien: Europaverlag).

Bukey, Evan Burr (2000). *Hitler's Austria: Popular Sentiment in the Nazi Era, 1938-1945.* (Chapel Hill, N.C.: University of North Carolina Press.).

Chamarbagwala, Rubiana & Hilcías Morán. (2009). "The Human Capital Consequences of Civil War: Evidence from Guatemala." *HiCN Working Paper 59.*

Chaney, Eric & Hornbeck Richard. (2013). "Economic Growth in the Malthusian Era: Evidence from the 1609 Spanish Expulsion of the Moriscos." *mimeo.*

Das, Jishnu, Stefan Dercon, James Habyarimana, Pramila Krishnan, Karthik Muralidharan & Venkatesh Sundararaman. (2013). "School Inputs, Household Substitution, and Test Scores." *American Economic Journal: Applied Economics,* Vol. 5(2), 29-57.

Davis, Donald & David Weinstein. (2002). "Bones, Bombs, and Break Points: The Geography of Economic Activity." *American Economic Review*, 92(5), 1269-1289.

Davis, Donald & David Weinstein. (2008). "A Search for Multiple Equilibria in Urban Industrial Structure." *Journal of Regional Science*, 48(1), 29-65.

Deaton, Angus. (1991). "Savings and Liquidity Constraints." *Econometrica,* Vol. 59(5), 1221-1248.

Dell, Melissa. (2010). "The Persistent Effects of Peru's Mining Mita." *Econometrica*, 78(6), 1863-1903.

EU. (2008). "Institutional and Capacity Building of Bosnia and Herzegovina Education System." Available at: www.erisee.org

Fiala, Nathan. (2009). "The Consequences of Forced Displacement in Northern Uganda." *HiCN Working Paper 65.*

Greene, William. (2008). *"Econometric Analysis."* Sixth Edition. Prentice Hall.

Grosfeld, Irena, Alexander Rodnyansky & Ekaterina Zhuravskaya. (2013). "Persistent Antimarket Culture: A Legacy of the Pale of Settlement after the Holocaust." *American Economic Journal: Economic Policy*, 5(3), 189-226.

Guiso, Luigi, Paola Sapienza & Luigi Zingales. (2011). "Civic Capital as the Missing

Link." In *Handbook of Social Economics,* Vol. 1A, ed. Jess Benhabib, Alberto Bisin and Matthew Jackson, 417-480. North-Holland.

Hagspiel, Hermann. (1995). *Die Ostmark: Österreich im Grossdeutschen Reich 1938 bis 1945.* (Wien: Braumüller).

Hanushek, Eric. (2003). "The Failure of Input-Based Schooling Policies." *The Economic Journal,* Vol. 113(485), 64-98.

Hornbeck, Richard & Suresh Naidu. (2014). "When the Levee Breaks: Black Migration and Economic Development in the American South." *American Economic Review*, 104(3), 963-990.

Houtenville, Andrew & Karen Smith Conway. (2008). "Parental Effort, School Resources, and Student Achievement." *Journal of Human Resources,* Vol. 43(2), 437-453.

IDMC. (2009). *"A profile of the internal displacement situation: Croatia."* Available at: www.internal-displacement.org

Justino, Patricia. (2011). "Violent Conflict and Human Capital Accumulation." *HiCN Working Paper 99.*

Kennedy, Peter. (1981). "Estimation with correctly interpreted dummy variables in semilogarithmic equations." *American Economic Review,* Vol. 71(4), 801.

Kimball, Miles. (1990). "Precautionary Savings in the Small and in the Large." *Econometrica,* Vol. 58(1), 53-73.

Knack, Stephen & Philip Keefer. (1997). "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation." *Quarterly Journal of Economics,* 112(4), 1251-1288.

Kondylis, Florence. (2010). "Conflict Displacement and Labour Market Outcomes in Post-War Bosnia and Herzegovina." *Journal of Development Economics,* Vol. 93, 235-248.

Kongsamut, Piyabha, Sergio Rebelo & Danyang Xie. (2001). "Beyond Balanced Growth." *Review of Economic Studies*, 48(4), 869-882.

León, Gianmarco. (2012). "Civil conflict and Human Capital Accumulation: The Long Term Consequences of Political Violence in Perú". *Journal of Human Resources,* Vol. 47(4), 991-1022.

Müller, Daniel, Benno Torgler & Eric Uslaner. (2012). "Inherited Trust and Growth - Comment." CREMA Working Paper No. 2012-04.

Maddison, Angus. (2007). "Historical Statistics." Groningen Growth and Development Centre. www.ggdc.net/MADDISON/Historical_Statistics/horizontal-file_03-2007.xls (accessed March 10, 2012.)

Merrouche, Ouarda. (2006). "The Human Capital Cost of Landmine Contamination in Cambodia." *HiCN Working Paper 25.*

MHHR (Ministry for Human Rights and Refugees). (2003). *"Comparative Indicators on Refugees, Displaced Persons and Returnees, Property Law Implementation, and Reconstruction in BiH from 1991 to 30 June 2003"*

Miguel, Edward & Gérard Roland. (2011). "The long-run impact of bombing Vietnam." *Journal of Development Economics*, 96(1), 1-15.

Mooney, Erin & Colleen French. (2005). "Barriers and Bridges: Access to Education for Internally Displaced Children." The Brookings Institution-University of Bern.

Morawek, Elisabeth & Wolfgang Neugebauer. (1989). *Österreicher und der Zweite Weltkrieg.* (Wien: Österreichischer Bundesverlag).

Nunn, Nathan & Wantchekon Leonard. (2011). "The Slave Trade and the Origins of Mistrust in Africa." *American Economic Review*, 101(7), 3221-3252.

Nunn, Nathan. (2008). "The Long-term Effects of Africa's Slave Trades." *Quarterly Journal of Economics*, 123(1), 139-176.

Nunn, Nathan. (2014). "Historical Development." in *Handbook of Economic Growth*, Vol. 2A, ed. Philippe Aghion and Steven Durlauf (Elsevier), 347-402.

OECD, Mission to Bosnia and Herzegovina. (2006). *"Action Plan on School Enrolment and Completion in Bosnia and Herzegovina."*

Pivovarova, Margarita & Eik Swee. (2012). "Quantifying The Microeconomic Effects of War: How Much Can Panel Data Help?" *HiCN Working Paper 116.*

Rahim, Abdel, Dany Jaimovich & Ylönen Aleksi (2013). "Forced displacement and behavioral change: An empirical study of returnee households in the Nuba Mountains." *HiCN Working Paper 157.*

Rodriguez, Catherine & Fabio Sanchez. (2009). "Armed Conflict Exposure, Human Capital Investments and Child Labor: Evidence from Colombia." *HiCN Working Paper 68.*

Sarvimäki, Matti, Roope Uusitalo & Markus Jäntti. (2009). "Long-Term Effects of Forced Migration." *IZA Discussion Paper No. 4003.*

Shemyakina, Olga. (2011). "Labour Market, Education and Armed Conflict in Tajikistan." *HiCN Working Paper 106.*

Silber, Laura & Allan Little. (1996). *"The Death of Yugoslavia."* London, England: Penguin Books.

State Agency for Statistics of BiH, Statistical Institute of the Federation of BiH, Republika Srpska Statistical Institute & The World Bank. *Living Standard Measurement Survey - Bosnia and Herzegovina.* Available at: tinyurl.com/bdetznv

Swee, Eik. (2009). "On War and Schooling Attainment: The Case of Bosnia and Herzegovina." *HiCN Working Paper 57.*

Swee, Eik. (2010). "Post-conflict Partition, Ethnic Homogenization, and the Provision of Public Schooling." *Working Paper,* University of Toronto.

Tabellini, Guido. (2010). "Culture and Institutions: Economic Development in the Regions of Europe." *Journal of the European Economic Association,* 8(4), 677-716.

Thalmann, Friedrich. (1954). "Die Wirtschaft in Österreich." in *Geschichte der Republik Österreich*, ed. Heinrich Benedikt (Wien: Verlag für Geschichte und Politik).

UNESCO. (1996). "Review of the education system in the Federation of Bosnia and Herzegovina." Available at: www.unesco.org

UNESCO. (1997). "Review of the education system in the Republika Srpska - Bosnia and Herzegovina." Available at: www.unesco.org

UNHCR. (2009). *"Global Report 2009."* Available at: www.unhcr.org/gr09

van Garderen, Kees Jan & Chandra Shah. (2002). "Exact interpretation of dummy variables in semilogarithmic equations." *The Econometrics Journal,* Vol. 5(1), 149-159.

Vocelka, Karl. (2010). *Österreichische Geschichte.* (München: C.H. Beck).

Voors, Maarten, Eleonora Nillesen, Philip Verwimp, Erwin Bulte, Robert Lensink & Daan van Soest. (2012). "Violent Conflict and Behavior: A Field Experiment in Burundi." *American Economic Review,* Vol. 102(2), 941-964.

Appendix A

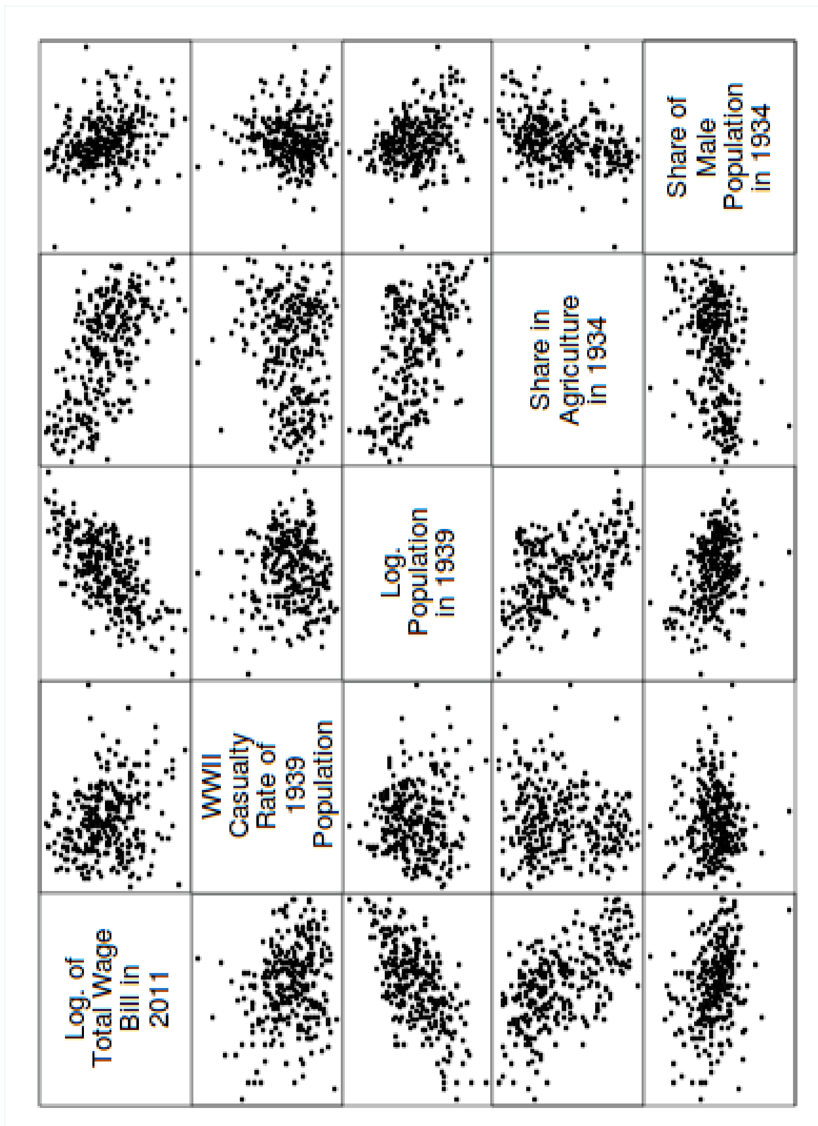
Appendix to Chapter 1

A.1 Data Sources

The data for the economic variables come from various publications of *Statistik Austria*, the Austrian statistical agency (earlier names of the agency: *K. K. Statistischen Zentralkommission*, *Bundesamt für Statistik*, and *Österreichisches Statistisches Zentralamt*): "Gemeindelexikon der im Reichsrate vertretenen Königreiche und Länder: Bearbeitet auf Grund der Ergebnisse der Volkszählung vom 31. Dezember 1900", "Die Ergebnisse der österreichischen Volkszählung vom 22. März 1934", "Die Ergebnisse der österreichischen Volkszählung vom 1. Juni 1951", "Die Ergebnisse der österreichischen Volkszählung vom 21. März 1961", "Die Ergebnisse der österreichischen Volkszählung vom 12. Mai 1971", "Arbeitsstättenzählung 1973", "Volkszählung 1981", "Arbeitsstättenzählung 1981", "Volkszählung 1991", "Arbeitsstättenzählung 1991", "Ein Blick auf die Gemeinde" (`www.statistik.at/blickgem`).

A.2 Partial Correlations

Figure A.1: Partial Correlations of the Main Variables



A.3 Descriptive Statistics of Robustness Variables

Table A.1: Descriptive Statistics

	Obs.	Mean	Std. Dev.	Min.	Max.
	(1)	(2)	(3)	(4)	(5)
Share of Market Status 1945	300	0.167	0.373	0	1
Share of Males in 1934	300	0.501	0.025	0.395	0.593
Vote Share of NSDAP 1930	300	0.012	0.025	0	0.247
Vote Share of Social Dem. 1930	300	0.197	0.189	0	0.710
Vote Share of Christian Dem. 1930	300	0.636	0.241	0	1
Share of Jewish Population 1934	300	0.001	0.007	0	0.117
Elevation of Municipality	300	621.4	346.7	117	1628

A.4 Gravity Model of the Commuting Pattern

Statistik Austria also publishes commuting streams between two municipalities for each census. For the 2011 census, I prepared the data to estimate a gravity model, which provides more detail on the determinants of commuting streams as the previous regressions in Table 1.7. An observation is a pair of municipalities if the share of dead soldiers is available for both locations and there are at least 20 people commuting in one direction, as the number of commuters is left-censored. This requirement creates 437 pairs out of the 300 municipalities available. Since there are several destination municipalities for some municipalities of origin and vice versa, I am able to include municipality of destination/origin fixed effects. The basic estimated equation is

$$\log C_{o,d} \quad = \quad \beta_o S_o^{WW2} + \beta_d S_d^{WW2} + M_{o,d}\gamma + X_o\delta_o + X_d\delta_d + \varepsilon_{o,d}, \tag{A.1}$$

where $C_{o,d}$ is the number of people commuting from municipality o (origin) to municipality d (destination), S^{WW2} is the share of dead soldiers in WWII in each municipality, $M_{o,d}$ contains the log. of the distance between the two locations and an indicator whether the municipalities are in different districts, X is the set of control variables of each municipality, and $\varepsilon_{o,d}$ is an error term. A district of origin fixed effect is included in column 1, while there are municipality of origin or destination dummies in columns 2-5. Of course multi-collinear control variables are dropped when municipality fixed effects are included. Columns 4 and 5 use a Tobit model to acknowledge the left-censoring of the dependent variable.

Table A.2: Commuting Pattern in 2011 - Gravity Model

	Log. Number of Commuters				
	(1)	(2)	(3)	(4)	(5)
WWII Casualty Rate in Destination Municipality	-10.28***	-13.58***		-14.43***	
	(2.537)	(3.244)		(2.604)	
WWII Casualty Rate in Origin Municipality	1.80		0.09		0.17
	(1.882)		(1.838)		(1.629)
Control Variables	Yes	Yes	Yes	Yes	Yes
District Fixed Effects	Yes				
Municipality FE		Yes	Yes	Yes	Yes
Tobit Model				Yes	Yes
No. Observations	437	437	437	437	437
R-squared	0.58	0.72	0.75		

Significance: $*$: 10% $**$: 5% $***$: 1%
Standard errors in parenthesis.
Control variables include the log. population in 1939, an indicator variable for market status in 1945, the share of population in agriculture in 1934 and the share of the male population in 1934.

The results in Table A.2 present a clear picture: municipalities with many casualties during WWII attract fewer commuters, while there is no significant effect on the out-commuting pattern. This finding confirms the results from Table 1.7.

A.5 Share of Foreigners

During and after WWII Austria experienced a large influx of foreigners. These includes displaced people who settled in or transited through Austria, or guest workers looking for employment in the booming Austrian economy of the 1950s and 60s. Foreigners could have selected into municipalities where the workforce has reduced through war casualties. If foreigners were less likely to establish new firms in the service sector, the large output reduction found in this paper could be explained by the influx of foreigners to replace war casualties.

The population censuses of 1951, 1971, 2001, and 2011 record the number of non-Austrian citizens of each municipality. In Table A.3 the share of foreigners of the total population in each year is regressed on the war casualties and the usual control variables to test this hypothesis. None of the estimated effects is significantly different from zero and the point estimates are much smaller than the standard errors. Based on these results, I reject the hypothesis that selection of foreign settlement into high-casualty municipalities can explain the total effect on output.

Table A.3: Share of Foreigners

	Share of Foreigners in			
	1915	1971	2001	2011
	(1)	(2)	(3)	(4)
WWII Casualty Rate of 1939 Population	-0.00	-0.06	-0.11	-0.14
	(0.117)	(0.123)	(0.233)	(0.194)
Control Variables	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
No. Observations	300	300	300	300
R-squared	0.35	0.42	0.37	0.38

Significance: $*$: 10% $**$: 5% $***$: 1%
Standard errors in parenthesis.
Control variables include the log. population in 1939, an indicator variable for market status in 1945, the share of population in agriculture in 1934 and the share of the male population in 1934.

Appendix B

Appendix to Chapter 3

B.1 Endogeneity Problem with Panel Data

Let the data generating process be as in AC:

$$Y_{ct} = \alpha_0 + \alpha_1 S_{ct} + \alpha_2 Y_{ct-1} + F_c + F_t + \varepsilon_{ct} \tag{B.1}$$

$$S_{ct} = \gamma_0 + \gamma_1 S_{ct-1} + \gamma_2 Y_{ct-1} + \Phi_c + \Phi_t + \nu_{ct} \tag{B.2}$$

where the country fixed effect F_c is possibly correlated with the trust variable S_{ct} or lagged income Y_{ct-1}.

The estimated regression equation in AC is:

$$Y_{ct} = a_0 + a_1 \left(\gamma_1 S_{ct-1}\right) + a_2 Y_{ct-2} + a_c + e_{ct} \tag{B.3}$$

AC use the difference to Sweden for every variable which accounts for the time fixed effect F_t.

I show that the problem of biased estimates discussed in section 2 with a cross section argument is still present in the case of estimation with country fixed effects. Instead of country dummies, I use the first differencing method to eliminate the country fixed effects to make for easier algebra.

Note that in AC the dependent variable is measured in 1935 and 2000, inherited trust comes from 25 years before that, while the lagged dependent variable is measured in 1870 and 1935. Consistent with this structure, the model observes the dependent variable in t and $t-2$, while inherited trust is from periods $t-1$ and $t-3$ and lagged income from periods $t-2$ and $t-4$. The differences need therefore be taken over two periods. Using the structural equations and substitute for S_{ct} and Y_{ct}, we get:

$$\begin{aligned} Y_{ct} - Y_{ct-2} &= \alpha_1 \left(1 + \alpha_1 \frac{\gamma_2}{\gamma_1} + \frac{\alpha_2}{\gamma_1}\right)(\gamma_1 S_{ct-1} - \gamma_1 S_{ct-3}) \\ &\quad + \alpha_2 (\alpha_1 \gamma_2 + \alpha_2)(Y_{ct-2} - Y_{ct-4}) \\ &\quad + (\alpha_1 \gamma_2 + \alpha_2)(\varepsilon_{ct-1} - \varepsilon_{ct-3}) \\ &\quad + (\varepsilon_{ct} - \varepsilon_{ct-2}) + \alpha_1 (\nu_{ct} - \nu_{ct-2}) \end{aligned} \tag{B.4}$$

As in AC, I ignore the problems of dynamic panel models to focus on the endogeneity issue that arises from the correlation between S and ε.

It is easy to see that the structural equation (B.4) predicts the same relationship between income, inherited trust and lagged income as in the cross-sectional model in section 2. Also the correlation between the lagged error term $(\varepsilon_{ct-1} - \varepsilon_{ct-3})$ and inherited trust is still an issue. Therefore the regression model (B.3) estimated with panel data and country fixed effects does not make headway in identifying the parameter of interest.