Market Size and Spatial Growth—Evidence from Germany's Post-war Population Expulsions: A Comment

Antonio Ciccone and Jan Nimczik*

July 2024

Abstract

The scale effects that have become an integral part of growth theory imply that productivity should be increasing in population size. We use newly digitized data to estimate the relation between GDP per worker and refugee settlements in West Germany following the arrival of 8 million WWII refugees—more than 15% of the West German population in 1949. Our approach builds on the county-level analysis of the relation between GDP per capita growth and refugee settlements in Peters (2022). As we find that his estimates do not reflect the effect on GDP per capita, we also provide corrected per-capita estimates.

1 Introduction

When West Germany was founded in 1949, four years after the end of World War II, it had a population of around 50 million. Some 8 million, more than 15% of its population, were ethnic German refugees who had been expelled from former German territory and other regions in Eastern Europe since 1945. Not only did a large number of refugees arrive within a short period of time, the spatial distribution of refugee settlements within West Germany was uneven. Building on this uneven distribution, Braun and Kvasnicka (2014) find a negative effect of refugee settlements on the agricultural employment share in 1950 but that the effect becomes weaker and statistically insignificant by 1961. Schumann (2014) concludes that all effects of refugee settlements on observables disappear quickly after 1950, except for an effect on population growth that persists to 1970.[1] Due to the

*Ciccone: University of Mannheim, CEPR, CESifo; antonio.ciccone@uni-mannheim.de.
Nimczik: ESMT Berlin, IZA, Rockwool Foundation Berlin; jan.nimczik@esmt.org.
We thank the statistical offices of Baden-Württemberg, Bayern, Hessen, Niedersachsen, Nordrhein-Westfalen, Rheinland-Pfalz, and Schleswig-Holstein for their great help. We also thank Alexander Göppert for outstanding research assistance and Michael Peters for comments. Financial support by the German Research Foundation (DFG) through CRC TR 224 (project A04) is gratefully acknowledged.

[1] However, Ciccone and Nimczik (2024) find a positive long-run effect of refugee settlements on productivity, wages, income, and rents—key economic outcomes that were not available to Schumann.

lack of output data, research on the effect of WWII refugees on spatial productivity has been lagging. The first estimates have been published only recently in Peters (2022). He uses newly digitized GDP data for West Germany to estimate the effect of county-level refugee settlements on different measures of productivity and test for the scale effects that have become an integral part of growth theory (Jones, 2005; Akcigit, 2017; Peters, 2022).

Peters's main ordinary-least squares (OLS) and instrumental-variables estimates relate refugee settlements in 1950 to GDP per capita growth from 1935 to 1950 for the short-run effect and to GDP per capita growth from 1935 to 1961 for the longer-run effect. He finds that "income-per-capita growth between 1935 and 1950 is essentially unrelated to the inflow of refugees" and that "by contrast, [...], the relationship between refugee inflows and long-run income-per-capita growth is positive, suggesting a form of dynamic agglomeration" (Peters, 2022, p. 2369). He extends his OLS analysis beyond 1961 and finds that the growth effects of refugees persist into the 1970s. Peters argues that these findings are direct evidence for the scale effects at the core of (semi-)endogenous growth theories. He then employs the OLS estimates for the 1935-1950 and the 1935-1961 period—combined with OLS estimates of the relation between refugee settlements and income-per-capita growth from 1935 to 1980—to calibrate a semi-endogenous growth model with dynamic agglomeration effects of the West German economy.[2] The model is shown to account for the evolution of the growth effect of refugee settlements over time (Peters, 2022, Figure 4).

We start with a review of Peters's empirical analysis. We find that *per capita* refers to *per population* for per-capita income measures in 1935 and 1950, but not after 1957. As a result, Peters does not estimate the relation between refugee settlements and income-per-capita growth from 1935 to 1961 or later. Based on newly digitized data we then examine (i) the relation between refugee settlements and per-capita growth from 1935 to 1961 and later; (ii) the differences in the relation between refugee settlements and per-capita growth from 1935 to 1950 and from 1935 to 1961 when per-capita growth is consistently defined; and (iii) the relation between refugee settlements and GDP per worker in 1961 and 1970.

Estimating the effect of WWII refugees on productivity involves resolving two main challenges. The first challenge lies in the measurement of productivity. The two most direct measures are total factor productivity and GDP per worker (average labor productivity). Peters has county-level data on GDP starting in 1957. In 1950 and 1935, he uses a measure that he takes to be proportional to GDP.[3] As Peters does not have

[2]Because of concerns about the precision of his income measure for 1950, Peters also presents calibration results that do not rely on the OLS estimates for income-per-capita growth over the 1935-1950 period.

[3]A proportional measure would be as good as GDP itself as Peters's specifications employ log productivity.

any data on where people work for years before 1980, he estimates the relation between refugee settlements and GDP *per capita*. However, the outcome Peters refers to as *GDP per capita* is actually GDP per *Wirtschaftsbevölkerung* in the years from 1957 to 1974. *Wirtschaftsbevölkerung* is a statistic employed in West German regional statistics until the 1970s. It sums twice the net commuting inflow into a county to its population, with the intention of capturing the number of people living in households with some member working in the county. As *Wirtschaftsbevölkerung* is idiosyncratic to German statistics, the results for GDP *per capita* growth from 1935 to 1957-1974 in Peters cannot be compared with results for other countries. Nor can Peters's estimates of the relation between refugee settlements and GDP per capita growth for different time periods be compared with each other, as he uses the term *per capita* to refer to *per population* in 1935 and 1950, to *per Wirtschaftsbevölkerung* for the 1957-1974 period, and to *per worker* in 1980 and later years. Because of the inconsistent use of *per capita* in Peters, it is unclear whether there was an effect of refugees on per-capita growth and, if there was, how it evolved over the short, medium, and long run.

We therefore correct the per-capita estimates in Peters by re-estimating his specifications using a consistent definition of *per capita* as *per population*. This yield statistically insignificant effects of 1950 refugee settlements on GDP growth per capita from 1935 to 1961, to 1970, and to 1980. By contrast, using newly digitized data on employment by place of work, we obtain strong positive and statistically significant OLS estimates for the effect of 1950 refugee settlements on GDP *per worker* in 1961 and 1970, indicating scale effects on labor productivity following the arrival of WWII refugees in West Germany.

A second challenge when estimating the effect of WWII refugees on productivity is the endogeneity of refugee settlements. Refugees may select into places with rapidly growing labor demand, which could result in a spurious positive correlation between refugee shares and economic growth. On the other hand, refugees may be restricted to settle in economically unattractive places, which could result in a spurious negative correlation between refugees and economic growth. To address this endogeneity challenge, Braun and Kvasnicka (2014) use the spatial distance from the expulsion regions to instrument refugee settlements in West Germany. Schumann (2014) implements a spatial regression discontinuity approach based on the difference between the US occupation zone in post-WWII Germany that admitted refugees, and the French occupation zone that restricted immigration.

Peters (2022) uses two different approaches to identify the effect of WWII refugees on productivity growth. His main estimation method is OLS with controls for several factors that determine refugee settlements, such as housing availability and pre-WWII characteristics. To alleviate endogeneity concerns, Peters leverages that he has measures of county-level per-capita output growth from 1935 to 1950—shortly after the WWII refugees arrived—and from 1935 to 1961 and that he can proxy for the pre-WWII level of productivity using his measure of GDP per capita in 1935. His findings indicate that, conditional on the level of pre-WWII GDP per capita and some other controls, growth up to 1950 is unrelated to refugee settlements. A statistically significant positive relationship between refugee settlements and growth only emerges when growth after 1950 is included in the analysis. Using newly digitized output data, we find this is no longer the case when per-capita output growth from 1935 to 1961 is measured in a way that is conceptually consistent with Peters's measure for per-capita output growth from 1935 to 1950.

The second estimation method of Peters is an instrumental-variables (2SLS) approach building on Braun and Kvasnicka. He concludes that 2SLS estimates of the effect of refugee settlements on GDP per capita growth from 1935 to 1961 are statistically significant and similar to OLS estimates as long as certain control variables are included. We find that these 2SLS estimates turn statistically insignificant when GDP per capita growth is measured using 1961 GDP per capita rather than per *Wirtschaftsbevölkerung*. On the other hand, we find strong positive 2SLS effects on 1961 GDP per worker. However, to obtain significant 2SLS estimates using either GDP per *Wirtschaftsbevölkerung* or per worker, the instrument must be allowed to have a heterogeneous effect across West German states while the restriction must be imposed that control variables have a homogeneous effect across states. This restriction generally precludes a causal interpretation of the resulting 2SLS estimates, as Blandhol et al. (2022) and Abadie et al. (2024) have pointed out recently.

2 The Measurement of Output per Capita

What was the effect of WWII refugee settlements on GDP per capita at the county level, shortly after refugees arrived and in the longer run? The main results of Peters (2022) are in his *Table VI, Panel E: GDP per capita growth* (OLS) and his *Table VII, Panel E: GDP per capita growth* (2SLS). Peters finds statistically insignificant effects of refugee settlements on 1935-1950 GDP per capita growth, but that "by contrast, [...], the relationship between refugee inflows and long-run income-per-capita growth is positive" (Peters, 2022, p. 2369).

We want to clarify two aspects of German regional statistics that imply that these results do not capture the effect of refugees on GDP per capita.

Population versus *Wirtschaftsbevölkerung* The first aspect concerns the denominator in Peters's measure of GDP per capita in his longer-run estimations. While Peters obtains county-level GDP per capita in 1935 and 1950 by dividing his measures of 1935 and 1950 GDP by population, 1961 GDP per capita refers to GDP per *Wirtschaftsbevölkerung*. The same holds true for other years between 1957 and 1974. *Wirtschaftsbevölkerung* is a statistic used in official West German statistics until the mid-1970s. It is defined as

$$Wirtschaftsbevölkerung_c = Population_c + 2 \times NetCommutingInflow_c \quad (1)$$

where $NetCommutingInflow_c$ is the difference between commuting inflows into county c and commuting outflows from c.[4] The factor 2 applied to net commuting inflows approximates the population-to-worker ratio in West Germany in the 1960s. The intention of $Wirtschaftsbevölkerung_c$ was to capture the number of people living in households with some member working in county c (Statistische Landesämter, 1966). In any case, the *Wirtschaftsbevölkerung* of counties is generally not equal to their population and 1957-1974 GDP *per capita* in Peters is therefore not GDP per population.

Value added versus gross sales As the GDP data that Peters employs to calculate GDP per *Wirtschaftsbevölkerung* in 1961 is not available for 1935 and 1950, he uses a measure that he assumes is proportional to the GDP of counties. His measure comes from publications collecting tax data for 1935 and 1950 (Statistisches Reichsamt, 1938; Statistisches Bundesamt, 1955). Peters believes that these publications "report value-added taxes for each county" that he can take "as being proportional to local GDP" (Peters, 2022, p. 2366). A value-added based alternative measure to GDP would be extremely useful as GDP is obtained as the value added produced in a county. If, moreover, the measure was proportional to GDP, it would be as good as GDP itself as Peters's specifications employ log productivity. However, Statistisches Reichsamt (1938) and Statistisches Bundesamt (1955) do not refer to a value-added tax but a gross-sales tax (Bundesministerium der Finanzen, 2019) and the specific variable that Peters uses is the value of gross sales of firms in the county.[5] Hence, his measure for county-level GDP in 1935 and 1950 is the gross sales of firms in the county. While gross sales is highly positively correlated with GDP at the local level, and therefore a useful proxy, it is not proportional to value added because of differences in industry structure (as industries differ in value added relative to gross output) and vertical integration (which determine the extent to which inputs are produced within firms).[6] Peters's measure for GDP per capita in 1935 and 1950 is gross sales divided by population.

These two aspects of German statistics imply that in *Table VI, Panel E: GDP per*

[4] To the best of our knowledge, *Wirtschaftsbevölkerung* does not have any equivalent in other countries. The concept in US regional statistics that comes closest, but has a completely different objective, is *daytime population*, which is $Population_c + 1 \times NetCommutingInflow_c$. The objective of this population concept is to approximate the number of people who may be in an area during regular working hours.

[5] Only in 1968 was the gross-sales tax changed to a value-added tax (Bundesministerium der Finanzen, 2019).

[6] See also Braun and Kvasnicka (2014). As detailed below, there is data on gross sales at the county level for 1961. This allows us to calculate the ratio of GDP to gross sales in 1961 across counties. The average is 0.6 and the range of variation is from 0.16 to 1.8. The correlation coefficient between 1961 GDP and 1961 gross sales at the county level is 0.945.

capita growth and in *Table VII, Panel E: GDP per capita growth* of Peters, *GDP per capita growth* over the 1935-1950 period is gross sales growth per capita in the county. On the other hand, *GDP per capita growth* over the 1935-1961 period is the log difference between 1961 GDP per *Wirtschaftsbevölkerung* and 1935 gross sales per capita. It therefore is an open question whether there was a relation between refugee settlements in 1950 and GDP per capita growth over the 1935-1961 period if *per capita* was consistently defined as *per population*.

We first correct Peters's per-capita results for the 1935-1961 period and re-estimate the relation between refugee settlements and GDP per capita growth using GDP per capita instead of per *Wirtschaftsbevölkerung* in 1961. Then we use newly digitized data on gross sales at the county level in 1961 to compare the relation between refugee settlements and gross sales per capita growth from 1935 to 1961 with Peters's estimates of the relation between refugee settlements and gross sales per capita growth from 1935 to 1950.

2.1 WWII Refugees and GDP per Capita Growth: OLS Results

Table 1: GDP per capita growth 1935 - 1961: OLS Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Main specifications								
Share of refugees in 1950	0.021	0.331	0.317	0.506*				
	(0.387)	(0.272)	(0.282)	(0.249)				
Observations	519	519	515	468				
Panel B. Robustness								
Share of refugees in 1950	0.331	0.334	0.305	0.321	0.338	-0.067		0.438
	(0.272)	(0.308)	(0.279)	(0.275)	(0.264)	(0.385)		(0.356)
Share of refugees in 1946							0.010	
							(0.328)	
Observations	519	519	471	519	471	334	392	392

Notes: The specifications in Panel A correspond to those in Peters's *Table VI, Panel E*, columns (5) to (8). The specifications in Panel B correspond to those in Peters's *Supplementary Material Table SM-2, Panel E*. See his table notes for details on the control variables and samples used. Standard errors clustered at the level of *Regierungsbezirke* in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In Table 1, we re-estimate the relation between refugee settlements in 1950 and 1935-1961 growth using the same estimation method (OLS) and controls as Peters. The sole change we make is that we measure 1961 GDP *per capita* using 1961 GDP per population instead of 1961 GDP per *Wirtschaftsbevölkerung* in Peters. Hence, the outcome variable in our Table 1 is the difference between log GDP per capita in 1961 and log gross sales per capita in 1935 where *per capita* refers to *per population*. Because gross sales is not proportional to GDP, the interpretation of this measure for per-capita growth from 1935 to 1961 is not completely straightforward. However, as in all specifications except for column (1) of Panel A, log gross sales per capita in 1935 is included as a control variable, we would obtain the same OLS estimates using the log level of GDP per capita in 1961 as an outcome variable.

The specifications in our Table 1, Panel A correspond to those in Peters's *Table VI, Panel E: GDP per capita growth* columns (5)-(8). The specifications in our Table 1, Panel B correspond to the specifications in Peters's *Supplementary Material Table SM-2, Panel E.* See his table notes for details on the controls used. In the specification with state fixed effects only (column (1) of Panel A) we obtain a very small and statistically insignificant point estimate of 0.02. By contrast, Peters's estimate using 1961 GDP per *Wirtschaftsbevölkerung* is large (1.159) and statistically significant. When we add controls, our estimates of the coefficient for 1950 refugee settlements remain statistically insignificant in all but one specification. By contrast, Peters reports statistically significant positive estimates of the coefficient for 1950 refugee settlements in all but one specification (the exception is a specification where he drops the state of Bavaria, which corresponds to column (6) in our Panel B, where his point estimate is positive but statistically insignificant).

The difference between our statistically insignificant estimates using 1961 GDP per capita and Peters's significant estimates using 1961 GDP per *Wirtschaftsbevölkerung* motivates us to examine the cross-county variation of *Wirtschaftsbevölkerung* relative to population more closely. Empirically, the range of cross-county variation of *Wirtschaftsbevölkerung* relative to population in 1961 is large, from 0.4 to 1.8 with a standard deviation of 0.2. To understand the determinants of this cross-county variation, we use (1) to obtain

$$\frac{Wirtschaftsbevölkerung_c}{Population_c} = 1 + 2\left(\frac{Workers_c}{Population_c} - \bar{\pi}\right) - 2(\pi_c - \bar{\pi}). \tag{2}$$

$Workers_c$ is the number of workers whose place of work is in the county or, equivalently, the number of jobs in the county; π_c is the share of the county population employed anywhere or, equivalently, the employment participation rate of the county population; and $\bar{\pi}$ is the worker-to-population ratio in West Germany. Recall that the factor 2 in the definition of *Wirtschaftsbevölkerung* in (1)—and hence in (2)—captures that the population-to-worker ratio in West Germany in the 1960s was around 2. Hence, the worker-to-population ratio $\bar{\pi}$ in (2) in the 1960s was approximately 1/2.

According to (2), cross-county differences between the *Wirtschaftsbevölkerung* and the population of a county reflect (i) differences in the number of workers (jobs) per capita and (ii) differences in the participation rate of the local population. If the *Wirtschafts-*

bevölkerung of counties differs from its population mostly because of cross-county differences in workers (jobs) per capita, then GDP per *Wirtschaftsbevölkerung* will be closely related to GDP per worker. This is seen most clearly when the participation rate in each county is the same and hence $\pi_c = \bar{\pi} = 1/2$. In this scenario, *Wirtschaftsbevölkerung* is proportional to the number of workers (jobs) and GDP per *Wirtschaftsbevölkerung* would therefore be proportional to GDP per worker. Hence, estimating the impact of refugees on log GDP per *Wirtschaftsbevölkerung* would be equivalent to estimating their effect of log labor productivity. On the other hand, if the *Wirtschaftsbevölkerung* of counties differs from its population mostly because of cross-county differences in participation rates, then GDP per capita is more closely related to GDP per worker. This is seen most clearly when each county has the same number of workers (jobs) per capita. In this scenario, GDP per capita would be proportional to GDP per worker. Hence, estimating the impact of refugees on log GDP per capita would be equivalent to estimating their effect of log labor productivity.

In sum, the spatial distribution of *Wirtschaftsbevölkerung* relative to population reflects spatial variation in jobs per capita and participation rates. Depending on the amount of spatial variation in jobs per capita compared to variation in participation rates, GDP per *Wirtschaftsbevölkerung* or GDP per capita may be more closely related to productivity.[7]

2.2 WWII Refugees and Sales per Capita Growth: OLS Results

In *Table VI, Panel E: GDP per capita growth*, Peters contrasts the relation between refugees and output per capita growth over the 1935-1950 period and over the 1935-1961 period. Because of the two aspects of German statistics discussed above, the per-capita growth rates he uses for the two periods differ conceptually in the implicit definition of *per capita* and in the measure of output.[8] It would be preferable to contrast the two time periods using conceptually consistent measures of per-capita growth. As there is no data on GDP at the county level before 1957, a conceptually consistent comparison can only be made by measuring growth over the 1935-1961 period in the same way Peters measures growth over the 1935-1950 period. As discussed above, his measure for 1935-

[7]The West German statistical offices were aware of these issues, see Statistische Landesämter (1966). In particular, they recognized that if the employment participation rates of the population were the same in all counties, GDP per *Wirtschaftsbevölkerung* would be proportional to average labor productivity at the county level.

[8]While Peters does not discuss the conceptual difference between the two measures of per-capita growth due to the different definitions of *per capita*, he reviews the implications of his different output measures in his Supplementary Material S.M.-2.6 (as mentioned above, he takes his measure for GDP per capita in 1950 to be relatively less reliable and therefore also calibrates his model without the OLS estimates for the 1935-1950 period in his robustness section). There he compares OLS estimation results using what he takes to be value-added taxes per capita in 1970—but are actually taxable gross sales per capita in 1970 (Statistisches Bundesamt, 1972)—and GDP per capita in 1970. In his comparison, 1970 GDP per capita refers to GDP per *Wirtschaftsbevölkerung*.

1950 growth is growth in gross sales per capita (per population). We have obtained and digitized county-level data on gross sales in 1961 and 1962 for all West German states from their statistical offices.[9] This data allows us to obtain the growth in gross sales per capita over the 1935-1961 period and the 1935-1962 period.

Table 2: Gross sales per capita growth 1935 - 1961: OLS Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Main specifications								
Share of refugees in 1950	-0.895***	0.248	0.271	0.538				
	(0.287)	(0.344)	(0.352)	(0.362)				
Observations	519	519	515	468				
p-value for $H_0 : \beta^{ref}_{1950} = \beta^{ref}_{1961}$	0.123	0.161	0.185	0.016				
Panel B. Robustness								
Share of refugees in 1950	0.248	0.260	0.119	0.242	0.141	-0.0188		0.372
	(0.344)	(0.458)	(0.327)	(0.345)	(0.352)	(0.535)		(0.450)
Share of refugees in 1946							0.259	
							(0.441)	
Observations	519	519	471	519	471	334	392	392
p-value for $H_0 : \beta^{ref}_{1950} = \beta^{ref}_{1961}$	0.082	0.323	0.105	0.077	0.084	0.098	0.334	0.117

The estimation method and model specifications in Panel A correspond to those in Peters's *Table VI, Panel E*, columns (5)-(8). The specifications in Panel B correspond to those in Peters *Supplementary Material Table SM2, Panel E*. See his table notes for details on the control variables and samples used. Standard errors clustered at the level of *Regierungsbezirke* in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Our results for the 1935-1961 period are in Table 2 (results for the 1935-1962 period are similar). The specifications in our Panel A correspond to those in *Table VI, Panel E: GDP per capita growth* columns (5)-(8) in Peters, while the specifications in our Panel B correspond to those in Peters's *Supplementary Material Table SM-2, Panel E*. The sole change we make is that we use 1961 gross sales per capita instead of 1961 GDP per *Wirtschaftsbevölkerung*. We find that the relation between 1950 refugee settlements and 1935-1961 gross sales growth per capita is negative and statistically significant in the specification with state fixed effects only (column (1) of Panel A) and becomes positive in the specifications with additional controls. Hence, in contrast to Peters's results for 1935-1961 in *Table VI, Panel E: GDP per capita growth*, our estimates do not point in the same direction in the specification with state fixed effects only and with additional controls. Moreover, the coefficient for refugee settlements is never statistically significantly different from 0 with additional controls. The conclusion from Table 2 is therefore that 1950 refugee settlements did not have a statistically significant, positive effect on gross sales growth per capita 1935-1961. As our measure for 1935-1961 growth per capita is conceptually consistent with the measure for 1935-1950 growth per capita in *Table VI, Panel E: GDP per capita growth* in Peters, we can compare the effect of refugee settlements on growth

[9] This has only been possible thanks to the kind support of the librarians of the statistical offices of the German federal states. The exact data sources are listed in our appendix.

over these two periods of time. For each specification, our Table 2 reports *p*-values from a simple F-test of equality of the coefficients for 1950 refugee shares for the 1935-1950 period (reported in Peters's *Table VI, Panel E*), columns (1)-(4) and for the 1935-1961 period (reported in our Table 2). Overall, the evidence for statistically different relations between refugee settlements and sales per capita growth between the two periods is weak.

2.3 WWII Refugees and GDP per Capita in 1970 and 1980: OLS Results

In *Table X: Additional moments for quantitative analysis*, Peters reports estimates of the relation between 1950 refugee settlements and GDP per capita growth from 1935 to 1980. His estimate is 0.201 with a standard error of 0.198. While Peters's measures of GDP per capita for earlier years are (proxies for) GDP per population or per *Wirtschaftsbevölkerung*, GDP per capita in 1980 and later years is GDP per worker (by county of work). As a result, his estimates for 1935-1980 cannot be compared to earlier periods.

To see how the relation between refugees and GDP per capita growth evolved over time when a consistent definition of *per capita* is employed, we re-estimate the specification in Peters's *Table X* using newly digitized data on county-level population in 1980 (Statistisches Bundesamt, 1981). The sole change we make is that we measure 1935-1980 GDP per capita growth using 1980 GDP per capita instead of 1980 GDP per worker. In this case, we obtain a smaller point estimate for 1950 refugee settlements of 0.075 with a standard error of 0.294. We also obtain a statistically insignificant point estimate for 1950 refugee settlements on GDP per capita growth from 1935 to 1970, with a point estimate of 0.206 and a standard error of 0.256. Hence, the relation between 1950 refugee settlements and GDP per capita growth from 1935 to 1961, to 1970, and to 1980 is statistically insignificant.[10]

2.4 WWII Refugees and GDP per Worker: OLS Results

The West German statistical offices recognized the drawbacks of GDP per capita and GDP per *Wirtschaftsbevölkerung* for the analysis of labor productivity at the county level. In a special effort in the 1960s, they therefore calculated county-level GDP per worker for 1961—the first year with such data in the history of German official statistics to the best of our knowledge—based on consistent methodologies for GDP by county of production and the number of workers by county of work (Statistische Landesämter,

[10]Following Peters, regressions for growth from 1935 to 1970 and from 1935 to 1980 are based on county-level data using the county boundaries of 1975 (there were extensive territorial reforms around 1970 that reduced the number of counties in West Germany). The magnitude of the point estimates for the 1935-1970 period and the 1935-1980 period can therefore not be directly compared to our estimates of GDP per capita growth in Table 1.

1966). We digitized the data to estimate the relation between 1950 refugee settlements and GDP per worker in 1961.

Table 3: GDP per Worker 1961: OLS Eestimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Main specifications								
Share of Refugees in 1950	-0.651***	0.583***	0.696***	0.808***				
	(0.239)	(0.184)	(0.175)	(0.172)				
Observations	519	519	515	468				
Panel B. Robustness								
Share of refugees in 1950	0.583***	0.675***	0.638***	0.571***	0.665***	0.573*		0.687***
	(0.184)	(0.193)	(0.193)	(0.193)	(0.204)	(0.281)		(0.233)
Share of refugees in 1946							0.0672	
							(0.252)	
Observations	519	519	471	519	471	334	392	392

Notes: The estimation method and model specifications in Panel A correspond to those in Peters's *Table VI, Panel E*, columns (5)-(8). The specifications in Panel B correspond to those in Peters *Supplementary Material Table SM2, Panel E*. See his table notes for details on the control variables and samples used. Standard errors clustered at the level of *Regierungsbezirke* in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Our results are in Table 3. As before, we use the same estimation method and specifications as in Peters's *Table VI* and *Table SM-2*. However, the outcome variable is the log level of GDP per worker in 1961, not a growth rate over the 1935-1961 period.[11] The control variables in our Panel A correspond to those in *Table VI, Panel E: GDP per capita growth* columns (5)-(8) in Peters, while the specifications in our Panel B correspond to those in *Supplementary Material Table SM-2, Panel E* in Peters. In the specification in column (1) with state fixed effects only, the point estimate is negative and statistically significant. When we add controls, point estimates turn positive and statistically significant. Our point estimates with controls indicate that an increase in the 1950 population share of WWII refugees of 10 percentage points is associated with a 6-8% higher level of labor productivity in 1961. Hence, the results in our Table 3 indicate strong scale effects on labor productivity following the arrival of WWII refugees in West Germany.

[11] We find this outcome variable easier to interpret than the 1935-1961 growth rate obtained as the difference between log GDP per worker in 1961 and log gross sales per capita in 1935. In any case, except for column (1) of Panel A, log gross sales per capita in 1935 is always among the controls and estimates are therefore unchanged when the outcome variable is the growth of 1961 GDP per worker relative to 1935 gross sales per capita.

2.5 WWII Refugees and GDP per Worker in 1970 and 1980: OLS Results

To see whether the relation between refugee settlements and GDP per worker in 1961 persisted to 1970, we also digitized data on workers by county of work for 1970 (Statistische Landesämter, 1973). We find a relation between 1950 refugee settlements and GDP per worker in 1970 that is statistically significant with point estimates between 0.51 and 0.78.[12] As we explained in Section 2.3, Peters's estimate of the relation between 1950 refugee settlements and GDP per capita growth from 1935 to 1980 in his *Table X* is based on GDP per worker in 1980. Hence, his estimate of 0.201 with a standard error of 0.198 indicates a drop in the effect of refugee settlements on GDP per worker during the 1970s.[13]

3 Instrumental Variables Results

With the goal to provide complementary evidence that the OLS results for 1935-1961 per-capita growth reported in his *Table VI, Panel E: GDP per capita growth* reflect the causal effect of WWII refugees, Peters (2022) also reports instrumental-variables (2SLS) estimates. His 2SLS approach builds on Braun and Kvasnicka (2014) in using the (population-weighted) distance to the expulsion regions as an instrument for the 1950 share of refugees in the local population within West Germany. However, distance to the expulsion regions turns out to be a weak instrument for refugee settlements at Peters's more granular geographic level. Rather than using distance to the expulsion regions as a single instrument, Peters therefore employs seven instruments obtained by multiplying distance to expulsion regions with dummy variables for the seven federal states in his dataset. This approach allows distance to the expulsion regions to have different effects on county-level refugee settlements within different states.

Peters's 2SLS estimates of the effect of refugee settlements on GDP per capita growth from 1935 to 1961 using 1961 GDP per *Wirtschaftsbevölkerung* are in his *Table VII, Panel E: GDP per capita growth*. He finds that 2SLS estimates are similar to OLS estimates as long as certain controls are included. Our Table 4, Panel A, replicates Peters's 2SLS

[12] The estimate depends on the level of geographic aggregation (there were extensive territorial reforms around 1970 that reduced the number of counties in West Germany) and the source used for 1970 GDP (the statistical offices published GDP figures in Statistische Landesämter, 1973, and revised figures in Statistische Landesämter, 1978). The point estimate of 0.51 (standard error 0.17) is for the specification in column (2) of Panel A of our Table 3 using the 1973 GDP source and the level of geographic aggregation that we employed for the 1935-1961 period. The point estimate of 0.78 (standard error 0.38) corresponds to the specification in Peters's *Table X* using the 1978 GDP source and the level of geographic aggregation employed by Peters in *Table X* for the 1935-1980 period.

[13] While Peters's outcome variable is the difference between log GDP per worker in 1980 and log gross sales per capita in 1935, his estimate would be unchanged if the outcome used was log GDP per worker in 1980 as log gross sales per capita in 1935 is among his control variables.

estimates. In the specification with state fixed effects only in column (1), the 2SLS point estimate of the effect of 1950 refugee settlements on growth is negative. The 2SLS point estimate turns positive with the controls in column (2). When the additional controls in columns (3) and (4) are included, 2SLS estimates turn statistically significant.

Table 4: Instrumental Variables Estimates. The Role of Bavaria

	(1)	(2)	(3)	(4)
Panel A. Specification in Peters (2022)				
Share of refugees in 1950	-0.671	0.208	0.517*	0.471**
	(0.615)	(0.370)	(0.271)	(0.238)
(Log) distance to inner-German border		0.074	0.162***	0.102**
		(0.0665)	(0.0447)	(0.0520)
Observations	519	519	515	468
First-stage F-stat	82.778	23.307	24.381	22.087
Overid Restrictions Test p-value	0.024	0.000	0.000	0.000
Panel B. Interacting distance to inner-German border with Bavaria				
Share of Refugees in 1950	-0.671	-0.0814	-0.189	-0.575
	(0.615)	(0.519)	(0.419)	(0.493)
(Log) distance to inner-German border × Not Bavaria		0.008	0.002	-0.134
		(0.128)	(0.106)	(0.132)
(Log) distance to inner-German border × Bavaria		0.099	0.223***	0.186**
		(0.098)	(0.065)	(0.084)
Observations	519	519	515	468
First-stage F-stat	82.778	18.970	18.831	16.901
Overid Restrictions Test p-value	0.024	0.000	0.000	0.005

Notes: The estimation method and model specifications in Panel A replicate Peters's *Table VI, Panel E*, columns (5)-(8). See his table notes for details on the control variables and samples used. The specifications in Panel B include one additional control variable, an interaction of the (log) distance to the inner-German border with the dummy variable that equals one for counties in Bavaria. Standard errors clustered at the level of *Regierungsbezirke* in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Due to the measurement issues discussed in Section 2, Peters's 2SLS estimates for income-per-capita growth from 1935 to 1961 do not capture the 2SLS effect of refugee settlements on per-capita growth. We report the corrected 2SLS estimates where per-capita GDP growth is calculated using 1961 GDP per capita instead of GDP per *Wirtschaftsbevölkerung* in our Appendix Table A2, Panel A. In this case, 2SLS estimates are statistically insignificant in all specifications. Panel B of Appendix Table A2 reports 2SLS estimates when the outcome variable is the growth of gross sales per capita from 1935 to 1961 (and results are therefore comparable to Peters's 2SLS estimates for gross sales per capita growth from 1935 to 1950). Again, 2SLS estimates are statistically insignificant in all specifications. Finally, in Panel C of Appendix Table A2, we use the log level of

GDP per worker in 1961 as the outcome variable. 2SLS estimates of the effect of 1950 refugee settlements on productivity in 1961 are positive and statistically significant in all specifications except with state fixed effects only.

We now abstract from the measurement issues in Peters and discuss whether his 2SLS estimates of the effect of 1950 refugee settlements on income-per-capita growth from 1935 to 1961 can be interpreted as causal effects. Our discussion also applies to all other 2SLS estimates that use the identification strategy of Peters (including the 2SLS estimates of the effect of refugee settlements on 1961 GDP per worker in our Appendix Table A2).

The 2SLS estimator of Peters is based on seven instruments for 1950 refugee settlements. As a result, his 2SLS estimator is overidentified and it becomes possible to test the implied overidentifying restrictions using the Sargan test. As can be seen from the p-values reported in our Table 4, Panel A, the overidentifying restrictions are rejected in every specification of Peters. This also holds true for every specification in our Appendix Table A2.

The Sargan test of overidentifying restrictions can be interpreted as a test whether the 2SLS estimates obtained using one instrument at a time identify the same causal relation (e.g., Angrist and Pischke, 2009). Hence, the p-values in Table 4 indicate that the different instruments in Peters do not identify the same causal relation. His multiple-instrument 2SLS estimates could still have the causal interpretation of weighted averages of local treatment effects (LATE). However, based on the recent contributions of Blandhol et al. (2022) and Abadie et al. (2024), we will argue that Peters's empirical specification for the control variables precludes a causal interpretation of his 2SLS estimates.

The reason why the multiple-instrument 2SLS estimates in Peters do not have a causal interpretation can be illustrated by examining the role played by the state of Bavaria. Peters finds that there is no first-stage effect of distance from the expulsion regions on refugee settlements across counties within Bavaria.[14] The non-existent first-stage effect for Bavaria is not necessarily a threat to causal identification. Under certain assumptions, multiple-instrument 2SLS estimates can be interpreted as positive-weighted averages of individual-instrument 2SLS estimates, with the weight on each individual 2SLS estimate increasing in the strength of the first-stage effect of the respective instrument (Angrist and Pischke, 2009, Sec. 4.5.2.). Hence, the only consequence of the non-existent first-stage effect within Bavaria may be that the individual-instrument 2SLS estimate corresponding to Bavaria enters Peters's multiple-instrument 2SLS estimate with a near-zero weight. However, the interpretation of multiple-instrument 2SLS estimates as positive-weighted averages of individual-instrument 2SLS estimates requires a sufficiently flexible specification for the effect of the control variables (Angrist and Pischke, 2009). What constitutes

[14] See Peters's *Supplementary Material Table SM-4*. See also the first-stage estimates in column (1) of our Table 5, which—in order to be consistent with the results in Peters's main tables—employs a slightly different set of controls than his *Supplementary Material Table SM-4*.

a sufficiently flexible specification depends on the empirical framework (Blandhol et al., 2022). Peters's setting is a special case of the framework in Abadie et al. (2024).

Abadie et al. examine the properties of the multiple-instrument 2SLS estimator in a framework where observations belong to different groups and the strength of a single instrument can vary across groups. The single instrument is employed to obtain multiple instruments (as many instruments as there are groups) by multiplying it with dummy variables for each group. For the multiple-instrument 2SLS estimator to have a causal interpretation in Abadie et al.'s framework, the specification must allow for the control variables to have different effects within the different groups. This is not the case for Peters's specifications in his *Table VII, Panel E: GDP per capita growth*. Quite in contrast, in these specifications the effect of all the control variables is restricted to be the same within all the states. Intuitively, this precludes a causal interpretation of his 2SLS estimates because the effect of (endogenous) refugee settlements—and hence the effect of the control variables—on growth is not identified in states where the instrument (distance from expulsion regions) does not have a first-stage effect on refugee settlements. However, multiple-instrument 2SLS estimates will nevertheless reflect the (non-identified) relationship between growth and the control variables in these states when the restriction is imposed that control variables have the same effect on growth in all states.

The consequences of restricting controls to have the same effect within different states can be illustrated by examining the role of one of Peters's key control variables, distance from the inner-German border (the border between West and East Germany). Peters argues that to identify the effect of refugee settlements on growth, it is necessary to control for distance from the inner-German border (Peters, 2022, p. 2370)[15] and his significant 2SLS estimates replicated in our Table 4 turn insignificant without this control. Distance from the inner-German border is therefore always included among the controls. As the slope estimate for distance from the inner-German border is restricted to be the same in all states, Bavaria could—even though the effect of distance from the inner-German border on growth in the state is not identified—have a substantial impact on Peters's 2SLS estimates.

We therefore re-estimate Peters's 2SLS regressions but allow the slope estimate for distance from the inner-German border to be different in Bavaria than in the rest of Germany. Our 2SLS results are in Panel B of Table 4. Allowing for a different slope estimate in Bavaria does not have a strong impact on the first-stage F-statistic of the set of seven instruments, which is not much lower than in the specifications of Peters that we replicate in Panel A. However, the positive and statistically significant 2SLS estimates of the effect of refugee settlements on growth in Panel A become negative and

[15] As counties closer to the inner-German border could have seen lower growth because of political uncertainty or the loss in market access and distance from the inner-German border is correlated with distance from the expulsion regions.

statistically insignificant. To understand why, it is useful to inspect the slope estimates for distance from the inner-German border reported in Table 4. The common (restricted) slope estimate in Panel A is positive and statistically significant in the specifications where 2SLS estimates are statistically significant. However, when we permit different slope estimates in Bavaria and in the rest of West Germany in Panel B, the slope estimate is only statistically significant (positive) in Bavaria. Allowing for this difference between Bavaria and the rest of West Germany turns the positive and statistically significant 2SLS estimates of the effect of refugee settlements on growth in Panel A negative and statistically insignificant.[16]

Table 5: Decomposition of 2SLS Estimation in Peters (2022)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Multiple Instruments IV			Individual IV			
	First Stage	Reduced Form	2SLS	First Stage	Reduced Form	2SLS	Weights
All Instruments			0.517*				
			(0.271)				
First-stage F-stat			24.38				
Overidentifying Restrictions Test			34.58***				
Baden-Württemberg × Expulsion Distance	-0.64***	0.26		-0.54***	-0.07	0.13	0.549
	(0.07)	(0.26)		(0.05)	(0.19)	(0.33)	
Bavaria × Expulsion Distance	-0.02	0.70**		0.17	0.61***	3.54	-0.023
	(0.09)	(0.27)		(0.10)	(0.20)	(2.33)	
Hesse × Expulsion Distance	-0.20	1.09***		0.03	0.74***	21.11	-0.005
	(0.13)	(0.32)		(0.09)	(0.21)	(53.95)	
North-Rhine Westphalia × Expulsion Distance	-0.19**	0.05		0.02	-0.34	-19.29	-0.005
	(0.08)	(0.34)		(0.08)	(0.27)	(93.77)	
Lower Saxony × Expulsion Distance	-0.45***	0.14		-0.34***	-0.32	0.95*	0.467
	(0.12)	(0.31)		(0.09)	(0.20)	(0.51)	
Rhineland-Palatine × Expulsion Distance	-0.07	-0.95**		0.10	-1.16***	-11.84	-0.002
	(0.13)	(0.47)		(0.11)	(0.42)	(15.50)	
Schleswig-Holstein × Expulsion Distance	-0.32***	-0.17		-0.14**	-0.46***	3.28***	0.020
	(0.09)	(0.21)		(0.06)	(0.08)	(1.25)	

Notes: This table decomposes the 2SLS estimate in Peters (2022), column (7) of *Table VII: The Effects of Refugee Inflows on the Local Economy: IV Estimates.* Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In Table 5, we broaden our analysis of Peters's 2SLS estimates by considering the role of all seven states. We start by looking at the first-stage effects of distance from the expulsion regions on 1950 refugee settlements within each states. Moreover, we also examine the reduced-form effects of distance from the expulsion regions on growth within each state. We focus on Peters's multiple-instrument 2SLS estimate in column (7) of his *Table VII, Panel E: GDP per capita growth.* The corresponding first-stage OLS estimates in column (1) of our Table 5 indicate that greater distance from the expulsion regions is associated with fewer refugees within each state. The (negative) first-stage

[16] We allow for a different slope estimate in the state of Bavaria only to illustrate the role of a state where the instrument does not have a first-stage effect. Allowing for different slope estimates in every state also turns 2SLS estimates for refugee settlements statistically insignificant as shown by Peters in his robustness analysis.

effect is statistically significant in four states and statistically insignificant in three states, including Bavaria.

The reduced-form OLS estimates in column (2) of our Table 5 indicate growth effects of distance from expulsion regions that go in different directions in different states. Greater distance from the expulsion regions is associated with lower growth within Rhineland-Palatine and within Schleswig-Holstein. The (negative) reduced-form effect is statistically significant for Rhineland-Palatine but not for Schleswig-Holstein. The sign of the reduced-form effect in these two states is consistent with Peters's hypothesis that greater distance from expulsion regions led to fewer refugees and therefore (because of the positive effect of refugees on GDP) lower GDP growth. However, greater distance from the expulsion regions is associated with higher growth within five of seven states. The (positive) reduced-form effect is statistically significant in two states, including Bavaria. The sign of the reduced-form effect in these five states is inconsistent with Peters's hypothesis on how distance from the expulsion regions affected growth.

In Table 5, columns (3)-(7), we examine the role of the different instruments in Peters for his multiple-instrument 2SLS estimate. Column (3) replicates the multiple-instrument 2SLS estimate in column (7) of his *Table VII, Panel E: GDP per capita growth*. In columns (4)-(6) of our Table 5 we consider seven separate (just-identified) 2SLS models, each based on exactly one of Peters's seven instruments. The effect of the control variables is restricted to be the same in the seven models.[17] Columns (4) and (5) contain the first-stage and reduced-form OLS estimates for the seven (just-identified) 2SLS models. Column (6) reports the just-identified 2SLS estimates. Finally, in column (7) we employ the results in Windmeijer (2019) to obtain the weights that Peters's multiple-instrument 2SLS estimate puts on each of the just-identified 2SLS estimates in column (6). Four of the seven weights are negative. Hence, Peters's multiple-instrument 2SLS estimate cannot be interpreted as a positive-weighted average of the 2SLS effects obtained one instrument at a time. Interestingly, there is a one-to-one correspondence between the negative weights in column (7) of our Table 5 and the statistically insignificant first-stage estimates in column (4).

In Appendix Table A3, we show—in line with the argument in Blandhol et al.—that the negative weights in our Table 5 are a consequence of restricting the slope estimates for the control variables to be the same in all states. In the Appendix Table we employ the same specification as in Table 5 but allow the effect of control variables to be state specific. In this case, the seven just-identified 2SLS estimates in column (6) enter the multiple-instrument 2SLS estimate with the positive weights reported in column (7).

[17] To do so, we first residualize 1935-1961 growth, 1950 refugee settlements, and distance from the expulsion regions multiplied by dummies for the seven states (the seven instruments) by regressing them on state fixed effects and the control variables assuming that controls have the same effects in all states. Then we use the residuals to estimate seven (just-identified) 2SLS models, each based on exactly one of the seven instruments.

However, the multiple-instrument 2SLS estimate in column (3) becomes negative and statistically insignificant.

4 Conclusion

Between 1945 and 1949, some 8 million WWII refugees arrived on the territory of what became West Germany. Not only was the population inflow very large, it was also unevenly distributed within West Germany. It is natural to wonder whether something can be learned from this population inflow regarding the effect of population size on local income per capita and productivity. While Peters (2022) aims at answering this question, his estimation results for GDP *per capita* between 1957 and 1974—the time period for which he finds statistically significant scale effects—refer to GDP per *Wirtschaftsbevölkerung*. *Wirtschaftsbevölkerung* is a statistic used in West German statistics until the mid-1970s. It is defined as the sum of the population of a county and twice the net commuting inflow into the county. The intention of *Wirtschaftsbevölkerung* was to capture the number of people in households with some member working in the county (Statistische Landesämter, 1966).

As *Wirtschaftsbevölkerung* is idiosyncratic to German statistics until the mid-1970s, the results for different time periods in Peters cannot be compared with each other nor can they be compared with results for other countries. We therefore re-estimate his specifications using GDP per capita and newly digitized data on gross sales per capita. The West German statistical offices recognized the drawbacks of GDP per *Wirtschaftsbevölkerung* and GDP per capita for the analysis of labor productivity and therefore also published data on the number of workers by county of work starting in 1961. We digitized this data to examine the effect of the refugee settlements on labor productivity in 1961 and 1970.

In addition to clarifying the measurement of county-level income per capita and the consequences for estimation, we also discuss Peters's instrumental-variables estimation strategy. In particular, based on the recent contributions of Blandhol et al. (2022) and Abadie et al. (2024), we examine whether his 2SLS estimates can be interpreted as causal.

References

Abadie, A., J. Gu, and S. Shen (2024): "Instrumental Variable Estimation with First-Stage Heterogeneity," *Journal of Econometrics*, 240.

Akcigit, U. (2017): "Economic Growth: The Past, the Present, and the Future," *Journal of Political Economy*, 125, 1736–1747.

Angrist, J. and J.-S. Pischke (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, Princeton.

Blandhol, C., J. Bonney, M. Mogstad, and A. Torgovitsky (2022): "When is TSLS Actually LATE?" mimeo.

Braun, S. and M. Kvasnicka (2014): "Immigration and Structural Change: Evidence from Post-War Germany," *Journal of International Economics*, 93, 253–269.

Bundesministerium der Finanzen (2019): "100 Jahre Umsatzsteuer in Deutschland," *Monatsbericht des BMF, Dezember*.

Ciccone, A. and J. Nimczik (2024): "The Long-Run Effects of Immigration: Evidence Across a Barrier to Refugee Settlement," mimeo.

Jones, C. (2005): "Growth and Ideas," *Handbook of Economic Growth*, 1063–1111.

Peters, M. (2022): "Market Size and Spatial Growth–Evidence from Germany's Post-War Population Expulsions," *Econometrica*, 90, 2357–2396.

Schumann, A. (2014): "Persistence of Population Shocks: Evidence from the Occupation of West Germany after World War II," *American Economic Journal: Applied Economics*, 6, 189–205.

Statistische Landesämter (1973): *Das Bruttoinlandsprodukt der kreisfreien Städte und Landkreise 1961, 1968 und 1970. Heft 4*, Stuttgart: Statistische Landesämter.

——— (1978): *Bruttoinlandsprodukt und Bruttowertschöpfung der kreisfreien Städte und Landkreise in der Bundesrepublic Deutschland: Revidierte Ergebnisse 1970, 1972 und 1974. Heft 8*, Stuttgart: Statistische Landesämter.

Statistische Landesämter (1966): "Das Bruttoinlansprodukt der kreisfreien Städte und Landkreise der Bundesrepublik Deutschland: 1957, 1961, 1964," in *Sozialproduktsberechnung der Länder*, ed. by Gemeinschaftsveröffentlicungen der statistischen Landesämter, Wiesbaden, vol. 2.

Statistisches Bundesamt (1955): *Statistik der Bundesrepublik Deutschland. Volume 122. Die Umsätze der Umsatzsteuerpflichtigen und deren Besteuerung*, Stuttgart: Kohlhammer.

——— (1972): *Fachserie L. Finanzen und Steuern. Reihe 7. Umsatzsteuer (Ergebnisse der Umsatzsteuerstatistik) 1970*, Stuttgart: Kohlhammer.

——— (1981): *Bevölkerung der Gemeinden mit Schlüsselnummern der Gemeinden und*

Verwaltungsbezirke:1980, Wiesbaden: Statistisches Bundesamt.

Statistisches Reichsamt (1938): *Statistik des Deutschen Reichs. Volume 1. Heft 511. Umsatzsteuerstatistik 1935*, Berlin: Verlag für Sozialpolitik, Wirtschaft und Statistik Paul Schmidt.

Windmeijer, F. (2019): "Two-Stage Least Squares as Minimum Distance," *Econometrics Journal*, 22, 1–9.

A Additional Tables

Table A1: Data Sources for Gross Sales at the County Level in 1961 and 1962

State	**Data Source**
Schleswig-Holstein	Kamp (1963): "Die Umsätze und ihre Besteuerung im Jahre 1961" in *Statistische Monatshefte Schleswig-Holstein* Vol. 15. Statistisches Landesamt Schleswig-Holstein (1964): "Die Umsätze der steuerpflichtigen Unternehmen in Schleswig-Holstein im Jahre 1962" in *Statistische Berichte* Vol. L/II/3-j.62.
Lower Saxony	Niedersächsisches Landesverwaltungsamt (1962): "Die Umsatzsteuerstatistik 1961 in Niedersachsen" in *Statistik von Niedersachsen* Vol. 18. Niedersächsisches Landesverwaltungsamt (1964): "Die Umsatzsteuer 1962 in Niedersachsen—Ergebnisse der Umsatzsteuerstatistik" in *Statistik von Niedersachsen* Vol. 41
North-Rhine Westphalia	Statistisches Landesamt Nordrhein-Westfalen (1963): "Umsätze und Umsatzsteuer in Nordrhein-Westfalen 1961" in *Beiträge zur Statistik des Landes Nordrhein-Westfalen* Vol. 159. Statistisches Landesamt Nordrhein-Westfalen (1964): "Umsätze und Umsatzsteuer in Nordrhein-Westfalen 1962" in *Beiträge zur Statistik des Landes Nordrhein-Westfalen* Vol. 180.
Hesse	Statistisches Bundesamt Wiesbaden (1963): "Umsatzsteuer (Ergebnisse der Umsatzsteuerstatistik 1961)" in *Finanzen und Steuern* Vol. 7. Hessisches Statistisches Landesamt (1965): "Die Umsätze und ihre Besteuerung 1962" in *Beiträge zur Statistik Hessens* Vol.8.
Rhineland-Palatinate	Statistisches Landesamt Rheinland-Pfalz (1962): "Der Umsatz und seine Besteuerung in Rheinland-Pfalz im Jahre 1961" in *Statistik von Rheinland-Pfalz* Vol. 105. Statistisches Landesamt Rheinland-Pfalz (1964): "Der Umsatz und seine Besteuerung in Rheinland-Pfalz im Jahre 1962" in *Statistik von Rheinland-Pfalz* Vol. 137.
Baden-Württemberg	Statistisches Bundesamt Wiesbaden (1964): "Umsatzsteuer (Ergebnisse der Umsatzsteuerstatistik 1961)" in *Finanzen und Steuern* Vol. 7. Statistisches Landesamt Baden Württemberg (1969):"Umsätze und ihre Besteuerung in Baden-Würtemberg 1962" in *Statistik von Baden-Württemberg*. Vol. 132
Bavaria	Bayerisches Statistisches Landesamt (1962): "Umsätze und Umsatzsteuer in Bayern—Ergebnisse der Umsatzsteuerstatistik für das Jahr 1961" in *Statistische Berichte* Vol. L/II/3 j.61. Bayerisches Statistisches Landesamt (1962): "Umsätze und Umsatzsteuer in Bayern—Ergebnisse der Umsatzsteuerstatistik für das Jahr 1962" in *Beiträge zur Statistik Bayerns* Vol. 256.

Table A2: Instrumental Variables Estimates

	(1)	(2)	(3)	(4)
Panel A. GDP per capita growth 1935-1961				
Share of Refugees in 1950	-0.585	0.232	0.391	0.267
	(0.603)	(0.363)	(0.367)	(0.278)
Observations	519	519	515	468
First-stage F-stat	82.778	23.307	24.381	22.087
Overid Restrictions Test p-value	0.001	0.021	0.006	0.070
Panel B. Gross sales per capita growth 1935-1961				
Share of Refugees in 1950	-0.331	0.493	0.631	0.492
	(0.457)	(0.591)	(0.593)	(0.502)
Observations	519	519	515	468
First-stage F-stat	82.778	23.307	24.381	22.087
Overid Restrictions Test p-value	0.021	0.001	0.000	0.017
Panel C. GDP per worker 1961				
Share of Refugees in 1950	-0.134	0.499*	0.760***	0.704***
	(0.448)	(0.280)	(0.275)	(0.249)
Observations	519	519	515	468
First-stage F-stat	82.778	23.307	24.381	22.087
Overid Restrictions Test p-value	0.000	0.000	0.000	0.000

Notes: The estimation method and controls correspond to *Table VII, Panel E: GDP per capita growth* columns (5)-(8) in Peters (2022). Panel A differs from Peters's *Table VII* in that we use 1961 GDP per capita instead of GDP per *Wirtschaftsbevölkerung.* Panel B differs from Peters's *Table VII* in that we use 1961 gross sales per capita instead of GDP per *Wirtschaftsbevölkerung.* Panel C differs from Peters's *Table VII* in that the outcome variable is 1961 log GDP per worker. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: 2SLS Estimation in Peters (2022) with heterogeneous effects of controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Multiple Instruments IV			Individual IV			
	First Stage	Reduced Form	2SLS	First Stage	Reduced Form	2SLS	Weights
All Instruments			-0.072				
			(0.836)				
First-stage F-stat			¿9,000,000				
Overidentifying Restrictions Test			32.86***				
Baden-Württemberg × Expulsion Distance	-0.61***	-1.78***		-0.61***	-1.78***	2.93***	0.151
	(0.14)	(0.47)		(0.14)	(0.46)	(0.98)	
Bavaria × Expulsion Distance	-0.01	0.68**		-0.013	0.68**	-50.40	0.005
	(0.09)	(0.27)		(0.09)	(0.27)	(351.3)	
Hesse × Expulsion Distance	0.39	2.03***		0.39	2.03***	5.15	0.024
	(0.38)	(0.17)		(0.38)	(0.16)	(5.09)	
North-Rhine Westphalia × Expulsion Distance	-0.44	-8.38		-0.44	-8.38	19.27	0.004
	(0.58)	(5.31)		(0.58)	(5.28)	(20.90)	
Lower Saxony × Expulsion Distance	-1.26***	0.81		-1.26***	0.81	-0.65	0.808
	(0.36)	(1.13)		(0.35)	(1.13)	(0.98)	
Rhineland-Palatine × Expulsion Distance	-0.15*	0.34		-0.15**	0.34	-2.24	0.004
	(0.08)	(1.39)		(0.08)	(1.38)	(8.05)	
Schleswig-Holstein × Expulsion Distance	0.20***	2.62***		0.20***	2.62***	13.33	0.004
	(0.00)	(0.00)		(0.00)	(0.00)	(0.00)	

Notes: The specification corresponds to Table 5 but allows controls to have heterogeneous effects across states. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.