Initial Conditions and the Big Push*

Dominick Bartelme† Zhimin Li‡ Daniel Velasquez§

August 2025

Abstract

This paper examines the long-run impact of temporary WWII spending on U.S. economic geography. Our empirical results show positive long-term effects on overall population and manufacturing activity only in counties with initially low manufacturing employment. Motivated by these findings, we develop a multisector economic geography model that features external economies of scale in manufacturing with non-constant elasticity. We calibrate the model and show that an S-shaped scale function in which sizable scale economies are limited to locations below 1500 manufacturing workers matches the data well. We use the calibrated model to reevaluate how temporary WWII spending shaped the evolution of U.S. postwar economic geography. Despite the presence of "big push" dynamics and multiple steady states, the aggregate effects of WWII spending on long-run outcomes are limited due to the restricted domain over which coordination failures operate.

JEL Classification: N40, N92, O14, O40, R11.

Keywords: WWII spending, economic geography, path dependence, external economies, big push.

*We thank seminar participants and conference discussants for their comments.

†Independent Researcher; email: dbartelme@gmail.com.

‡HSBC Business School, Peking University; email: zhimin.li@phbs.pku.edu.cn.

§The Robert Day School of Economics and Finance, Claremont McKenna College; email: dvelasquez@cmc.edu.

1 Introduction

During World War II, the United States government undertook the largest public spending program in U.S. history. Federal net outlays went from roughly 10% of GDP to over 40% in a few years, then subsequently retreated to less than 15% a few years later (Figure 1). Much of this increased spending consisted of supply contracts awarded to domestic firms for the production of materials and equipment for the war effort. What were the effects, if any, of this massive but temporary program on the long-run evolution of U.S. economic geography? This question has drawn interest in light of the broader literature on historical persistence, path dependence and place-based policies in economic geography, as well as the specific historical experience of the post-war United States with the decline of the Manufacturing Belt and the spread of industry to the Sun Belt.[1]

This paper revisits the long-run impact of this temporary WWII spending on U.S. economic geography empirically, theoretically, and quantitatively. We begin by linking a comprehensive dataset of war supply contracts during WWII with data on the long-run growth in population and manufacturing employment of U.S. counties from 1930 to 2000. Using a potential outcomes framework along with controls for initial conditions, we compare the long-run outcomes of counties that received large war supply contracts relative to their manufacturing employment to those that received small or no contracts. Consistent with the prior literature, we find that treated counties experienced higher long-run population growth and manufacturing employment growth relative to untreated counties. Our primary empirical contribution is to show that these positive treatment effects are driven almost entirely by counties with initially small manufacturing employment, which experience economically large and statistically significant long-run increases in population and manufacturing employment. Treated counties with initially large employment in manufacturing show strong short-run gains in popu-

[1]Econometric studies of WWII spending include Fishback and Cullen (2013); Fishback and Jaworski (2016); Jaworski (2017); Li and Koustas (2019); Garin and Rothbaum (2025); Jaworski and Yang (2025). A few examples from the vast historical literature include Wright (1986); Hooks and Bloomquist (1992).

lation and manufacturing employment, but these gains are almost completely reversed in subsequent decades.

We interpret our results as providing support for "big push" or "poverty trap" theories of regional specialization and development (Rosenstein-Rodan, 1943; Murphy, Shleifer, and Vishny, 1989; Krugman, 1991; Azariadis and Stachurski, 2005). In these theories, coordination failures can result in a region being stuck in an equilibrium that features a low level of economic activity relative to its potential. A large enough temporary shock induces a higher level of activity, which then becomes self-sustaining and permanent. However, as suggested by the term "poverty trap," most versions of the theory posit some scale of economic activity beyond which coordination failures diminish in importance and temporary shocks no longer have permanent effects. While these theories have motivated a sizable empirical literature on historical persistence and path dependence in a variety of contexts, the theoretical prediction of heterogeneity in the long-run response based on initial conditions has been largely neglected. It is precisely this prediction of the limited relevance of coordination failures beyond a certain scale of economic activity that we find support for in this paper.

Motivated by our findings, we develop a quantitative economic geography model with many regions that features the possibility of coordination failures leading to path dependence in long-run equilibria, but parameterize the scale of economic activity over which coordination failures can arise. We build on the recent framework of Allen and Donaldson (2020) by adding multiple sectors in a way that allow for long-run structural change, and more importantly the possibility of non-constant-elasticity local external economies of scale in the manufacturing sector that may be diminishing in the size of the local sector. We calibrate the model by choosing the scale function so that the response of the model to temporary demand shocks, stemming from temporary government purchases of manufactured goods, matches the heterogeneous responses that we found in the data. We find that the model matches the data quite well with a scale function that approximates an S-curve in manufacturing employment: initially flat up to 250 manufacturing workers in the county, increasing with elasticity of about 0.2 up to about 1500 manufacturing workers, then flat again thereafter. The calibrated model generates a large number of interior steady states and path dependence. In contrast, the constant elasticity version of the model, which is ubiquitous in the quantitative literature, has trouble matching the empirical moments precisely because it lacks a natural mechanism to limit the domain over which coordination failures apply.

We use the calibrated model to revisit how the WWII war supply effort shaped the evolution of U.S. postwar economic geography. We simulate the model with and without the WWII spending shocks, and compare the outcomes in the year 2000. We find that despite significant reallocation of population and manufacturing employment across smaller counties, the WWII spending had small long-run effects on U.S. aggregate manufacturing employment, welfare, and reallocation across larger regions. This is primarily due to the limited scope for reallocation to raise aggregate productivity in the calibrated model. History did matter, but it mattered little in this instance.

Our paper contributes several literatures. The first is the extensive empirical literature on historical persistence and path dependence in economic geography (e.g., Rauch (1993); Bleakley and Lin (2012); Davis and Weinstein (2002); Juhász (2018), surveyed in Lin and Rauch (2022)). Our main contribution is to focus on the heterogeneity of persistence with respect to initial conditions, a neglected topic in the empirical literature despite its theoretical importance. Our paper also relates to the literature on quantitative economic geography, recently surveyed in Redding and Rossi-Hansberg (2017). Our primary contribution to this literature is to develop a tractable quantitative model that embeds big push dynamics and parameterizes their scope. Finally, we contribute to the econometric literature on the impact of WWII on U.S. economic geography (Fishback and Cullen, 2013; Fishback and Jaworski, 2016; Jaworski, 2017; Rhode, Snyder Jr, and Strumpf, 2018; Li and Koustas, 2019; Garin and Rothbaum, 2025; Jaworski and Yang, 2025), which has generally found modest positive long-run effects of WWII spending on local economies. Our paper focuses explicitly on the temporary component of WWII spending (as opposed to wartime investment in long-lived structures as in Garin and Rothbaum (2025); Jaworski and Yang (2025)) and the heterogeneity of the impact with respect to initial conditions. Our findings support the view that the long-run aggregate impact of the *temporary* component of WWII spending was rather limited, despite some local effects.

The rest of the paper proceeds as follows. Section 2 briefly describes the data and the historical context. Section 3 conducts the empirical estimation. Section 4 describes the model and the structural estimation. Section 5 conducts counterfactual analysis, and Section 6 concludes.

2 Estimation and Identification

This section briefly describes the historical context of federal military supply contracts during WWII and our data sources, then discusses our approach to estimation and identification.

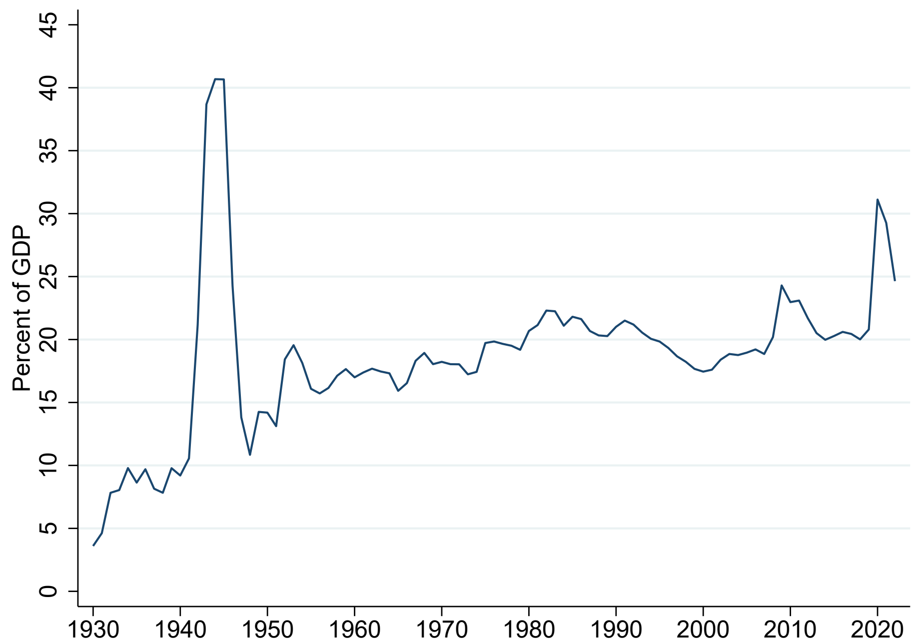
2.1 Historical Context and Data

During World War II, the federal government assumed an unprecedented level of control over the U.S. economy. At their peak, federal government expenditures accounted for over 40 percent of GDP, the highest proportion ever recorded (Figure 1). The federal government allocated $196 billion for military supply contracts (primarily for manufactured goods) and invested an additional $31 billion in new production facilities between June 1940 and June 1945.[2] This remarkable endeavor to turn the U.S. into the "Arsenal

[2] For reference, U.S. GDP was about $102 billion in 1940.

of Democracy" stands as the most significant economic intervention by the federal government in U.S. history.

Figure 1. Federal Net Outlays as Percent of GDP



Note: This figure shows the federal net outlays as percent of GDP over time. Source: Federal Reserve Bank of St. Louis, U.S. Office of Management and Budget; see the link here.

The allocation of war supply contracts during WWII was primarily determined by economic considerations rather than political factors. Rhode et al. (2018) find that the strongest predictors of the assignment of supply contracts at the state level were existing general industrial capacity prior to the war in 1939 and specialized industrial capacity for aircraft production. Their findings suggest that the government prioritized regions with established industrial infrastructure and specialized capabilities, reflecting a strategic approach to efficiently meet the demands of wartime production (Fishback and Cullen, 2013). The ability of the United States to rapidly procure and produce goods and services contributed significantly to the Allied victory in WWII (Harrison, 1998).

The data for WWII supply contracts are sourced from volumes originally issued by the Civilian Production Administration (formerly the War Production Board), which documented all defense contracts greater than \$50,000 issued between June 1940 and September 1945.[3] These military contracts were awarded for the production of combat supplies (e.g., planes, tanks), intermediate goods (e.g., propellers, gun fittings), and incidental materials (e.g., clothing, medical supplies).

We construct our measure of local government spending by aggregating the monetary value of these military contracts with private firms at the county level. The dataset comprises over 191,000 contracts with a total value of \$183 billion, a sum that substantially exceeded the 1940 U.S. GDP of \$103 billion. The distribution of contract values was highly skewed; while more than 90% of contracts were for less than \$1 million, a small number of high-value awards, particularly in aircraft production, accounted for a large share of total spending. For instance, major contracts included a \$669 million award to Boeing for B-17 bombers in Seattle, WA, and engine contracts with Pratt & Whitney (\$403 million in East Hartford, CT) and Wright Aeronautical (\$429 million in Lockland, OH). Furthermore, these procurement agreements were intentionally short-term, a design intended to mitigate the risk of renegotiation under wartime inflation. The median contract duration was seven months, with 90% of all contracts having a duration of 14 months or less. This short-term nature makes the contract data particularly suitable for measuring a temporary fiscal shock.

Data on county characteristics are the same as those in Kline and Moretti (2014) and come from a variety of sources including the Census of Population and Housing and Agricultural Census. We use the crosswalk provided by Eckert, Gvirtz, Liang, and Peters (2020) to obtain a sample of harmonized counties between 1900 and 2000. Additional details can be found in Online Appendix A.1.

[3]Planners divided war spending for 1940 to 1945 into two broad categories: contract spending and facilities spending. Facilities spending, which accounts for about 13% of total war spending, involved the building of government-financed military or industrial facilities of more than \$25,000 and included aircraft plants, shipyards, air fields, and cantonments (Fishback and Cullen, 2013). Recent papers that focus on the impact of facilities spending include Garin and Rothbaum (2025) and Jaworski and Yang (2025). We focus on contract spending in this paper because it accounts for the vast majority of war spending and, unlike facilities, which take a long time to depreciate, better reflects a temporary shock.

2.2 Empirical Model

We wish to estimate the impact of this massive but temporary increase in government spending on local economic outcomes in the long run, and in particular to understand how this impact varied with the initial size of the local manufacturing sector. Our motivation for exploring this potential heterogeneity comes from big push or poverty trap theories of regional specialization and development (Rosenstein-Rodan, 1943; Murphy et al., 1989; Krugman, 1991; Azariadis and Stachurski, 2005), in which coordination failures can result in a region being stuck in an equilibrium that features a low level of economic activity relative to its potential. Large temporary shocks to underdeveloped regions can induce a higher level of activity, which then becomes self-sustaining and permanent. This implication has been born out in previous research on WWII spending as well as in many (though not all) other contexts.[4]

We focus on a second important implication: temporary shocks can only have permanent effects on regions that are stuck in a low-activity equilibrium. For regions that are already highly developed, even a very large temporary shock will not affect its long run trajectory. This theoretical prediction of heterogeneity in the long-run response based on initial conditions has seen little or no empirical investigation, and is the focus of our analysis in this paper. Following a long tradition in the literature, we think of manufacturing as the sector generating increasing returns during this time period, and hence the size of the manufacturing sector as being the key initial condition for understanding heterogeneous responses to shocks.[5]

We draw on several additional implications of the theory to inform our empirical analysis. Since an underdeveloped manufacturing sector is a stable equilibrium, small shocks will have no long-run effect even in underdeveloped regions. On the other hand,

[4]For the U.S. experience in WWII, see, for example, Fishback and Jaworski (2016); Li and Koustas (2019). Examples of studies in other contexts include Bleakley and Lin (2012); Kline and Moretti (2014); Juhász (2018).

[5]See Bartelme, Costinot, Donaldson, and Rodriguez-Clare (2024) and Lashkaripour and Lugovskyy (2023) for some recent evidence on scale economies in manufacturing.

once a shock is large enough to move the region onto a trajectory leading to the higher equilibrium, further increases in the size of the shock will not alter the long-run outcome. These considerations suggest dichotomizing the war spending variable based on a threshold, dividing regions (counties in our analysis) into "treated" (large shocks) and "untreated" (small or no shocks) categories for the analysis. Similar reasoning suggests dichotomizing the initial conditions, dividing counties into high and low initial manufacturing employment (our measure of sector size). We divide observations into four quadrants based on these criteria, then compare long-run outcomes in the treated quadrants to their untreated counterparts.

To be clear, the conditional expectation of the treatment effect should vary continuously with both the initial conditions and the treatment size.[6] Our discretization strategy is dictated by the need to maximize the relevant information that can be obtained from limited data. Rather than estimate the entire two-dimensional surface of treatment effects, which is infeasible to do with any precision, we average the treatment effects over the most relevant domains by choosing appropriate thresholds that define the quadrants. The choice of thresholds is guided by both substantive and statistical considerations; we discuss our choices in more detail and conduct robustness checks in the following section, but we stress that there is no right or wrong way to choose these thresholds, only more or less informative ones. We also note that in no case are the treatment effects to be interpreted as structural parameters of some underlying model. We will use the estimated treatment effects as moments to be targeted when estimating our structural model in Section 4.

We follow Kline and Moretti (2014) and adopt the potential outcomes framework as

[6]Continuity is expected because the treatment size and initial manufacturing employment are not sufficient to identify which counties will be deflected from their counterfactual untreated long-run equilibrium trajectory by the treatment. For example, some counties with low initial manufacturing employment will already be on a trajectory towards a high equilibrium and will thus have a treatment effect of zero for even large treatment sizes, even though most counties with the exact same initial manufacturing employment will have positive treatment effects for most treatment sizes. Integrating over unobserved heterogeneity leads to continuity of the conditional expectation despite the discrete nature of the transition from a low to high equilibrium.

our empirical model, with the Average Treatment Effect on the Treated (ATT) as our primary parameter of interest.[7] If county i actually received treatment, the county i treatment effect is the difference between the actual county i outcome and the (unobserved) counterfactual outcome for county i had it not received treatment. The ATT is simply the average of these individual treatment effects taken over all counties that actually received treatment. Formally, let y_i^0 be the outcome of interest (e.g., the long-run change in manufacturing employment) in county i *if it had not received treatment.* We assume that

$$y_i^0 = \alpha_i^0 + \theta_0 X_i + \epsilon_i^0, \tag{1}$$

where X_i is a set of observable control variables and ϵ_i is an unobserved term. If county i *were to receive treatment,* the outcome y_i^1 would given by

$$y_i^1 = \alpha_i^1 + \theta_1 X_i + \epsilon_i^1. \tag{2}$$

Identification is achieved by assuming that both residuals are uncorrelated with the treatment conditional on the controls,

$$E[\epsilon_i^0 | X_i, D_i] = E[\epsilon_i^1 | X_i, D_i] = 0, \ \forall i. \tag{3}$$

The treatment effect for county i is given by $\alpha_i^1 - \alpha_i^0 + (\theta_1 - \theta_0)X_i$, and the ATT for each subgroup is obtained by averaging across treated counties in each subgroup. We allow for the possibility that the treatment effect is correlated with the observables, $\theta_0 \neq \theta_1$; if $\theta_0 = \theta_1$, the model collapses to a standard difference-in-differences regression. Equations (1) and (2) are estimated separately on the untreated and treated populations, re-

[7] See Imbens and Wooldridge (2009) for a review of the literature on estimation and inference in this model. The ATT can be consistently estimated under weaker assumptions than the average treatment effect (ATE). It also has an appealing interpretation in historical counterfactual analysis, especially when general equilibrium effects are important, since it avoids considering very different policies than those actually implemented.

spectively, and the estimates then combined to compute the ATT.[8]

The model (1) - (3) can be regarded as a local approximation to a structural model, such as the one we develop in Section 4. The controls X_i should include observables from the initial state that summarize all aspects of the local initial conditions useful in predicting the future evolution of local productivity and amenities. The controls should also account for WWII spending in nearby counties, which will affect outcomes in county i through spatial spillovers (e.g. those generated by trade and migration frictions). Given our focus on the temporary component of WWII spending, the controls should also include wartime government investment in long-lived capital assets such as new factories. Conditional on the appropriate set of controls, the residuals ϵ_i reflect the portion of future local productivity/amenity growth that is unforecastable based on the initial state. Our identification assumption is that the local WWII supply spending is uncorrelated with this unpredictable component of local productivity/amenity growth.

This description raises the question of where our identifying variation is coming from, given that treatment is non-randomly assigned and that we control for many of the observables in the initial state. We think of our identifying variation as coming from three primary sources. One is the composition of local industry; two places with the same level of manufacturing employment (which we control for in the regression) may have very different wartime demand depending on how useful the products they specialize in are for the war effort (see Rhode et al. (2018) for some evidence). Since the composition of wartime demand is very different from the composition of peacetime demand, this variation is plausibly exogenous with respect to the future (long-run) evolution of local productivity and amenities. The second source of variation is political, social or informational connections between wartime decisionmakers and particular geographic locations or firms, which again is plausibly exogenous with respect to long

[8]This is known as the Oaxaca-Blinder estimator (Oaxaca, 1973; Blinder, 1973). See also Kline (2011) for some robustness properties of the estimator and Kline and Moretti (2014) for an application in a similar context to ours.

run changes in fundamentals. The third source is the winner-take-all nature of most contracts, which implies that there might be very little difference between a county that received a large contract and one that did not. This last source could be particularly influential for counties with an initially small manufacturing base, for which the allocation of a single large contract could have a material effect on output and employment.

Admittedly, it is difficult to be completely confident that there is no confounding selection into treatment whenever the treatment is not explicitly randomly assigned. We address this concern, as best we are able, with two data exercises and a theoretical point. First, we use a placebo test to show that treatment status does not predict outcomes in the pre-treatment period. Second, we use data on industry composition to show that our results are robust to controlling for this source of variation. Finally, we note that the *difference* in the ATT between locations with initially low and high manufacturing employment, a key estimand of interest, is identified under the weaker assumption that selection bias equally affects both groups.

3 Empirical Results

Treatment and Initial Conditions We split the set of counties into those with an initially high manufacturing base—where log manufacturing employment in 1930 is above the 50th percentile nationally—and those with an initially low manufacturing base. We assign treatment status to each county based on the WWII war supply contract value per capita. Counties with a high level of war supply contracts were more densely populated and had higher manufacturing employment in the prewar period (see Table A.1 for summary statistics). This is not surprising because war supply contracts were allocated based primarily on existing production capacity (Fishback and Cullen, 2013).

We choose the definition of treatment so as to balance the dual objectives of ensuring an adequate number of control units for both groups and having the treatment group assignment lead to similar distributions of contract value within the treatment and control groups, across initial conditions. Our baseline is to assign counties to the treatment group if their log contract value per 1930 manufacturing employee exceeds 8, a round number very close to the mean. We conduct extensive robustness checks on both cutoffs (initial conditions and treatment) and discuss them further below.

Pretrend Test To adjust for pre-existing differences across counties before the WWII, we follow previous literature and control for a rich set of county-level variables related to prewar economic, social, demographic, and geographical characteristics in 1930 (Kline and Moretti, 2014; Li and Koustas, 2019). The covariates include log population, log employment, log manufacturing employment, agricultural employment share, manufacturing employment share, log manufacturing wages, dummies for wages in manufacturing or trade being missing, farm values, urban population share, the share of white population, the share of the population age 5 or above that are illiterate, the share of whites who are foreign-born, the share of households with a radio, maximum elevation, elevation range, and area.[9]

Moreover, since contract assignment is strongly correlated with initial manufacturing conditions (Rhode et al., 2018), we control for relevant variables in the manufacturing sector including log total manufacturing employment, manufacturing employment share, and log manufacturing wages. We control for an indicator variable for whether a region is classified as having initially high manufacturing employment, as well as an interaction of this indicator with the log prewar population, log prewar manufacturing employment, and the prewar manufacturing share. We also control for trends in population and manufacturing employment between 1900 and 1920. To account for potential spillover effects from neighboring counties, we control for the war supply contract

[9]To capture potential temporal correlation of shocks, we control for socioeconomic characteristics in 1930 and 1900 (except log population and log manufacturing employment in 1900 to prevent multicollinearity, as we also control for trends in these variables).

TABLE 1. Pretrend Tests of Outcomes in 1900-1930

	(1) Pop 1900-1930	(2) ManufEmp 1900-1930	(3) Pop 1900-1930	(4) ManufEmp 1900-1930	(5) Pop 1900-1930	(6) ManufEmp 1900-1930
Treat	−0.001	0.004	−0.004	−0.008	0.000	0.008
	(0.004)	(0.007)	(0.004)	(0.015)	(0.005)	(0.008)
Location	All	All	LowM	LowM	HighM	HighM
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows pretrend tests for the "treatment" effect on the growth rate between 1900 and 1930 (defined as the log difference between 1930 and 1900 levels divided by three) of the outcome variables. Columns (1)–(2) shows pretrend tests for the whole sample, columns (3)–(6) break them down by the initial levels of manufacturing employment: whether a county has manufacturing employment in 1930 below or above the 50th percentile. Treated counties correspond to those that received WWII war supply contract value per capita exceeds 8 in logarithmic terms. Block-bootstrapped standard errors at the state level are in parenthesis.

value per capita of neighboring counties. Importantly, we also control for WWII facilities spending to purge the any lasting direct effects of long-lived structures on productivity. To better balance the treated and control groups, we drop counties whose pre-WWII characteristics yield extreme predicated probability of treatment higher than 99% or lower than 1%.

We first assess whether the set of control variables perform well to adjust for observable differences across counties. To do so, we conduct a placebo test by estimating the "effects" of defense spending on 1900-1930 changes in population and manufacturing employment, for all counties together as well as separately for initially low and high manufacturing counties. This test attempts to detect if, conditional on the control variables, the outcome variables exhibit differential trends before WWII across treated and control counties. Because that time period predates the war, a finding of a significant relationship would signal a potential failure of our identification assumption. As shown in Table 1, no statistically or economically significant effects are discernable across all outcomes and subgroups.

Main Results Panel (a) of Table 2 shows the results of estimating equations (1) and (2) and computing the ATT across all observations, regardless of initial conditions. Columns (1)–(2) of Table 2 show the long-run effects of WWII spending on the growth of population and manufacturing employment between 1930 and 2000, with results expressed in terms of decadal growth rates. The impacts on the growth in population and manufacturing employment between 1930 and 2000, with the latter date being over 5 decades after the war ended and wartime spending subsided, are 1% and 1.6% per decade, respectively. These modest positive long-run average effects are consistent with the previous literature on the geographic impact of WWII spending (Fishback and Cullen, 2013; Fishback and Jaworski, 2016; Li and Koustas, 2019). For the immediate postwar period—up until 1960—we see much larger differential growth rates of 2.3% and 6% for population and manufacturing employment in treated counties (columns (3)-(4). However, these sizable impacts die out or are partially reversed over the next several decades (columns (5)-(6)). This dynamic pattern of growth and (partial) reversal reflects the temporary nature of the WWII demand shock and the gradual adjustment of treating counties to the loss of wartime demand for their products.

Panel (b) of Table 2 reports the ATT separately for counties with high (*HighM*) and low (*LowM*) initial manufacturing bases. For regions with initially low manufacturing employment, wartime defense spending increased the long-term growth rate of population and manufacturing employment levels by 2.9% and 3.8%, respectively, between 1930 and 2000. For regions with initially high manufacturing employment, these effects are much smaller both economically and statistically: the long-run effects are 0.5% and 1.0%, respectively, with both statistically indistinguishable from zero. Columns (3)-(6) report the dynamic patterns, which show that both treated groups experienced similar large positive effects on both population and manufacturing between 1930 and 1960. However, these positive impacts almost completely reverse over the next 40 years for the initially high manufacturing group, leaving the long-term impact for this group close to

TABLE 2. Long-Term Impact on Growth Rate of Outcomes

Panel (a): Baseline

	(1) Pop 1930-2000	(2) ManufEmp 1930-2000	(3) Pop 1930-1960	(4) ManufEmp 1930-1960	(5) Pop 1960-2000	(6) ManufEmp 1960-2000
Treat	0.010	0.016	0.023	0.060	0.002	−0.017
	(0.006)	(0.007)	(0.008)	(0.012)	(0.006)	(0.007)
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Panel (b): Initial Condition

	(1) Pop 1930-2000	(2) ManufEmp 1930-2000	(3) Pop 1930-1960	(4) ManufEmp 1930-1960	(5) Pop 1960-2000	(6) ManufEmp 1960-2000
Treat*LowM	0.029	0.038	0.029	0.087	0.028	−0.003
	(0.008)	(0.010)	(0.009)	(0.019)	(0.009)	(0.010)
Treat*HighM	0.005	0.010	0.021	0.053	−0.004	−0.021
	(0.007)	(0.008)	(0.009)	(0.013)	(0.007)	(0.009)
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows the impact of wartime spending on the growth rate of outcomes. Columns (1)–(2) show the treatment effect on the growth rate of outcome variables between 1930 and 2000 (defined as the log difference between 2000 and 1930 levels divided by seven). Columns (3)–(6) show the treatment effect on the growth rate of outcome variables by period: 1930-1960 and 1960-2000. Panel (a) shows the treatment effects for the whole sample, and panel (b) breaks them down by the initial levels of manufacturing employment: whether a county has manufacturing employment in 1930 below or above the 50th percentile. Treated counties correspond to those that received WWII war supply contract value per capita exceeds 8 in logarithmic terms. Block-bootstrapped standard errors at the state level are in parenthesis.

zero. In contrast, the initially small treated locations experienced little or no reversal from 1960-2000, leaving their population and manufacturing employment significantly higher after 60 years, relative to the controls.

Our results show that the strong positive long-run effects of WWII spending on population and manufacturing employment are found only for regions with initially small manufacturing sectors, as predicted by big push/poverty trap models. The dynamic pattern of the effects also supports the big push theory; it is quite reassuring that the short-term effects are similar for both groups, since we would expect a large temporary expansion in employment due to the demand shock regardless of any endogenous productivity effect. However, as the demand shock recedes only the locations that have

TABLE 3. Long-Term Impact on Growth Rate of Outcomes by Treatment Size

	(1) Pop 1930-2000	(2) ManufEmp 1930-2000	(3) Pop 1930-1960	(4) ManufEmp 1930-1960	(5) Pop 1960-2000	(6) ManufEmp 1960-2000
Treat Big*LowM	0.050	0.052	0.042	0.117	0.056	−0.011
	(0.017)	(0.016)	(0.019)	(0.022)	(0.015)	(0.017)
Treat Small*LowM	0.014	0.029	0.021	0.067	0.009	0.001
	(0.008)	(0.011)	(0.011)	(0.022)	(0.010)	(0.013)
Treat Big*HighM	0.014	0.022	0.038	0.091	−0.001	−0.025
	(0.009)	(0.013)	(0.013)	(0.022)	(0.008)	(0.011)
Treat Small*HighM	−0.000	0.001	0.009	0.027	−0.007	−0.017
	(0.007)	(0.008)	(0.008)	(0.011)	(0.007)	(0.009)
Controls	Yes	Yes	Yes	Yes	Yes	Yes

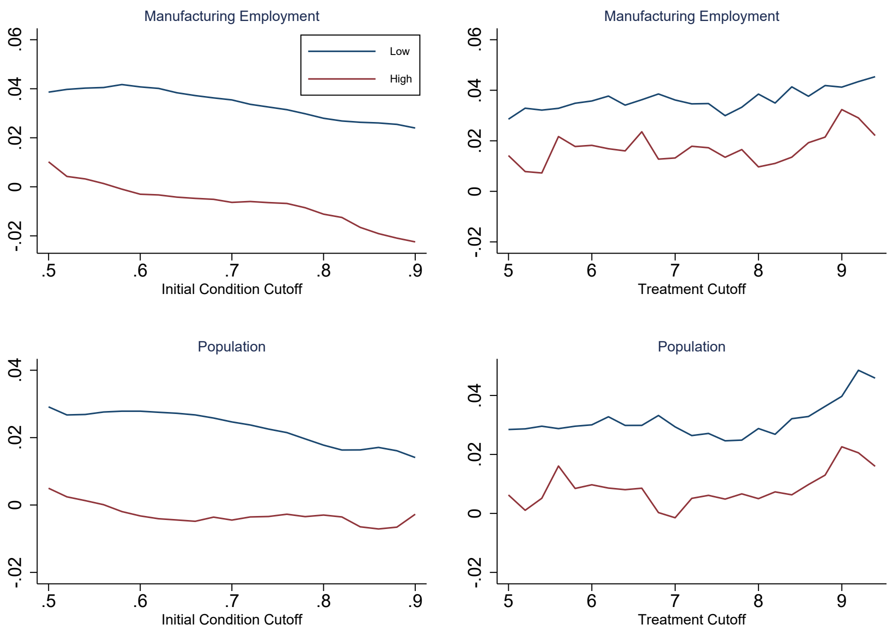
Notes: This table shows the impact of wartime spending on the growth rate of outcomes by treatment size. Columns (1)–(2) show the treatment effect on the growth rate of outcome variables between 1930 and 2000 (defined as the log difference between 2000 and 1930 levels divided by seven). Columns (3)–(6) show the treatment effect on the growth rate of outcome variables by period: 1930-1960 and 1960-2000. Low and high initial manufacturing base is defined by being above or below the 50th percentile of manufacturing employment in 1930. Treated counties correspond to those that received WWII war supply contract value per capita exceeds 8 in logarithmic terms. The boundary between "small" and "large" treatment is defined by a log per capita supply contract value of 9.36, which is the median value in the treated group. Block-bootstrapped standard errors at the state level are in parenthesis.

been pushed to a higher productivity state retain significant employment gains, while they largely fade away for locations that were not.

Additional Results and Robustness Checks Table 3 breaks the treatment size into two categories, with the log contract value per 1930 manufacturing employment above or below median, and re-estimates baseline regression equations (1) and (2). The results are as expected; larger treatments lead to larger short term (reflecting the larger direct effect) and long term (reflecting the higher probability of being pushed onto a higher productivity trajectory) gains for all groups. The long-run effects continue to be much larger for the initially small manufacturing group.

Figure 2 shows how varying the cutoffs for the initial high/low split as well as the treatment size affects the results. Classifying more locations as "small" and less as "large" results in a monotonic decline in both coefficients, with the effects for large locations turning negative as the category comes to include only the very largest locations. The

Figure 2. Coefficients with Varying Cutoffs of Treatment and Initial Conditions



Notes: This figure shows the coefficients as the cutoffs of treatment and initial condition definitions vary. The left panel correspond to estimates as the cutoff of classifying initially high- and low-manufacturing locations ranges from 50th to 90th percentiles. The right panel correspond to estimates as the treatment cutoff varies from 4 to 10 in log contract value per 1930 manufacturing employee.

spread between the estimates is fairly robust. These results suggest that it is indeed the smaller percentiles of the initial conditions that are driving the positive long-run effect. The results are also not too sensitive to the definition of treatment, as is also shown in Figure 2.

As mentioned above, our identifying variation comes from 3 primary sources: the composition of local industry, locally specific political and informational connections, and true randomness from the lumpy and winner-take-all nature of the contracts. The only observable element is the composition of local industry, and we now explore how using (or not using) this information affects our estimates. We divide manfuacturing sectors into those highly related to the war effort (e.g. auto, ship and aircraft manufacturing) and all others. Table A.2 shows that the share of prewar manufacturing employment in these sectors is highly correlated with both treatment status and contract value. It is plausible that this industry composition is also related to the future evolution of productivity and amenities, despite our other controls, which would violate our exclusion restriction. Therefore, as a robustness check Table A.3 reports the results of re-running our basic specification while controlling for industry composition. Reassuringly, the results are very similar to those reported in Table 2.

Finally, we discuss the concern that our results are driven by possible serial correlation of military spending. While military spending dropped precipitously at the end of the war, it did not drop all the way to the prewar level and it could be that locations with high wartime spending disproportionately benefited from this continued spending. We conduct two tests to address the concern. First, we show that long-term effects, especially for locations with initially low manufacturing, are present in across different categories of contract spending including those meant for civilian use. Specifically, we break contract spending data into three coarse categories as follows: "food, textile, chemicals, miscellaneous," "metal," and "machinery, electronics, transportation." As shown in Table A.5, many of the key coefficients are statistically significant for all categories of contract spending. Second, we show that the correlation between contract spending and defense spending in later, post-war periods (1966 or 2000) at state level is quite low and insignificant, as shown in Table A.6.[10]

4 Quantitative Model and Structural Estimation

4.1 Quantitative Model

We build on the recent framework of Allen and Donaldson (2020), which endows the basic quantitative economic geography model with a dynamic structure that can accommodate multiple steady states and path dependence. We add a) multiple sectors (agri-

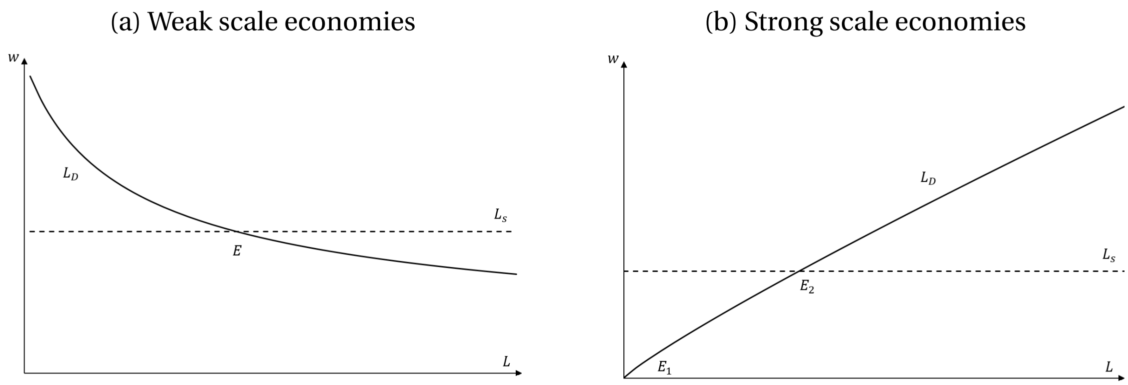
[10] We analyze this correlation at the state level because reliable military spending data are only available at the state level in post-war years from 1960s (Nakamura and Steinsson, 2014).

culture, manufacturing, services), b) intermediate goods, c) a government sector, and d) local external economies of scale in the manufacturing sector that may affect productivity with non-constant elasticity. The first three modifications help match the specific experiment in our setting or add quantitative realism, and our modeling of these is standard. The fourth, external economies of scale with non-constant elasticity, is motivated by our empirical findings, as we explain below.

Our structural estimation will show that external economies of scale that are diminishing in sector size beyond a threshold value are critical for matching the empirically observed difference in treatment effects across initial conditions. Some intuition for why this is the case can be gleaned from considering the case of a single competitive sector in partial equilibrium, with labor l the only factor of production, facing a flat labor supply curve and constant elasticity sector-level product demand with elasticity $-\sigma < -1$ (see Figure 3). The elasticity of the sector-level inverse labor demand curve is $\frac{(\sigma-1)\gamma(l)-1}{\sigma}$, where $\gamma(l)$ is the elasticity of productivity with respect to sectoral employment which we assume is driven by external economies. When $\gamma(l)$ is constant, then (ignoring the case of equality) either $\gamma(l) > \frac{1}{\sigma-1}$ and the inverse labor demand curve is always downward sloping (panel a) of Figure 3) or $\gamma(l) < \frac{1}{\sigma-1}$ and the inverse labor demand curve is always upward sloping (panel b) of Figure 3).

In Panel a) there is a unique equilibrium at point E that is interior and stable and hence there is no poverty trap. In Panel b) there are two equilibria, a stable one at zero employment (E_1) and an unstable interior equilibrium at E_2. The equilibrium at E_1 could be described as a poverty trap, since a temporary shock large enough to carry employment past the point E_2 would result in a permanently higher level of employment and productivity. However, this model predicts that locations with with *lower* initial employment should be *less* responsive to positive demand shocks in the long run than large places, the opposite of our empirical findings. More generally, the problem with this model of a poverty trap is that it is too stylized for quantitative work, since the only

Figure 3. Labor market equilibria, CES scale function

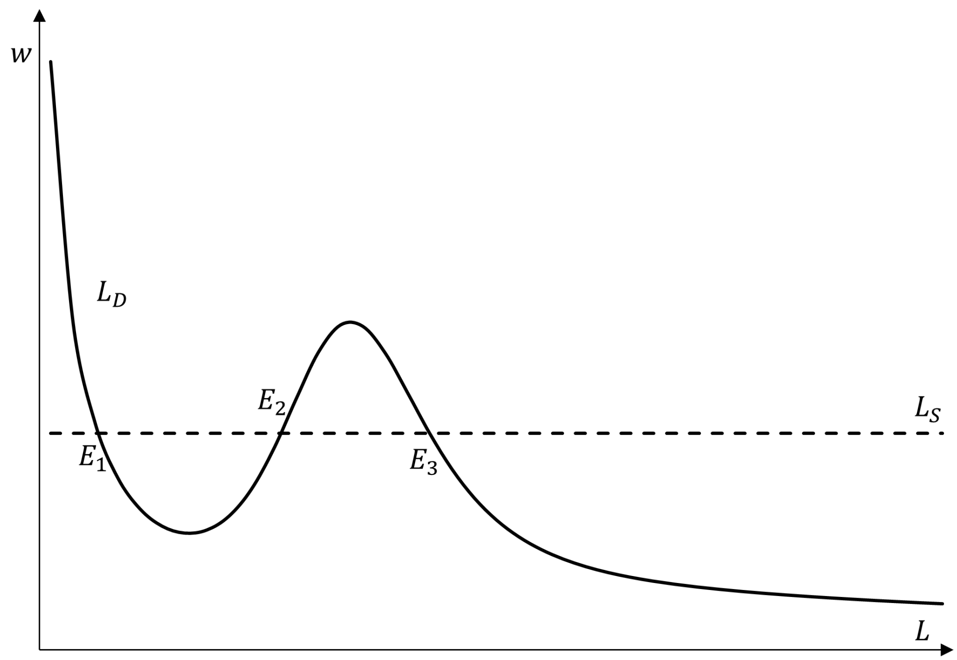


two long-run outcomes are an employment level of zero or infinity, as it lacks partial equilibrium mechanisms to limit the growth of the sector.

In contrast, consider an S-shaped scale function that is initially flat in the sector size, increases rapidly over some range, then flattens again as the sector grows large (Figure 4). Given downward sloping product demand, this type of scale function can generate a sectoral inverse labor demand curve that crosses the labor supply curve three times in the interior, implying a stable "low" equilibrium E_1 and a stable "high" equilibrium E_3, with an unstable equilibruim between them at E_2. A large enough temporary shock can move the sector from the low to the high equilibrium, which then becomes self-sustaining; however, an economy that is initially in the high equilibrium always returns to the same steady state in response to a positive shock (or a small enough negative one). This framework thus naturally implies differential responses to temporary shocks based on initial conditions.

This analysis is too simplistic in that it ignores general equilibrium considerations. As a matter of fact, the constant-elasticity model of Allen and Donaldson (2020) is capable of generating multiple stable interior steady states in general equilibrium, in both the original and our version. However, we find that the partial equilibrium *tendency* towards extreme sectoral allocations still manifests itself in general equilibrium, at least in

Figure 4. Labor market equilibria, S-shaped scale function



our calibration. Thus the constant elasticity specification tends to imply that either a) no location is in a poverty trap, or b) *every* location is in a poverty trap, unless the initial equilibrium is one that is already extremely (and counterfactually) highly concentrated (a "black hole" in the terminology of Fujita, Krugman, and Venables (2001)). This feature is in conflict with the empirical evidence from Section 3. For this reason, our estimation procedure strongly prefers scale economies that are sufficiently diminishing with large sector sizes. We discuss this issue further in Section 4.3.

Setup

Consider $i \in \{1, ..., N\}$ locations and discrete time $t \in \{0, 1, ...\}$. Individuals live for two periods. In the first period ("childhood"), they are born in their parents' location and consume the same as their parents. At the beginning of the second period ("adulthood"), they realize their own preferences and choose their residence as adults. In this location, they supply a unit of labor inelastically, consume, and have ζ_t children apiece.

Production

There are two tradable sectors, "agriculture" and "manufacturing," and a non-tradable sector, "services." Each location produces a differentiated variety of agricultural and manufacturing goods. The production in agriculture and services uses labor with constant returns to scale,

$$Q_{A,i,t} = T_{A,i,t} l_{A,i,t}, \tag{4}$$

$$Q_{S,i,t} = T_{S,i,t} l_{S,i,t}, \tag{5}$$

where T_A and T_S are exogenous productivities, and l_A and l_S are labor inputs. Manufacturing uses labor and intermediate inputs from the manufacturing sector. Producers perceive constant returns, but external economies of scale exist in total manufacturing labor. Manufacturing output is given by

$$Q_{M,i,t} = T_{M,i,t} f\left(l_{M,i,t-1}\right) l_{M,i,t}^{\alpha} q_{M,i,t}^{1-\alpha}, \tag{6}$$

where T_M is exogenous productivity, $f(l)$ is the endogenous productivity that depends on lagged sector scale, l_M is labor input, and q_M is a bundle of intermediate inputs.

Government

Government purchases manufacturing output from each location and throws it away. Aggregate government consumption is

$$G_t = \sum_i g_{i,t}, \tag{7}$$

where $g_{i,t} = \kappa_{i,t} G_t$ are government purchases from location i, and $\kappa_{i,t}$ is the purchase share. To fund these purchases, the government imposes a per-person lump-sum tax,

τ_t, satisfying $\tau_t L_t = G_t$, where L_t is total population at time t.

Demand

Preferences are given by a nested CES utility function, with an upper tier across sectors and a lower tier across locations within sectors,

$$C_{i,t} = \sum_{s=\{A,S,M\}} \rho_s^{\frac{1}{\eta}} C_{s,i,t}^{\frac{\eta-1}{\eta}}, \tag{8}$$

$$C_{s,i,t} = \sum_{n=N} C_{s,i,n,t}^{\frac{\sigma-1}{\sigma}}, \tag{9}$$

where $C_{s,i,n,t}$ is the consumption in location i of the sector s good produced in location n at time t. The lower tier elasticity of substitution σ is assumed to be the same for each sector, as well as for intermediate goods demand in the manufacturing sector. Let

$$\omega_{s,i,t} = \rho_s \left(\frac{P_{s,i,t}}{P_{i,t}}\right)^{1-\eta}, \forall s \in \{A, S, M\}, \tag{10}$$

where $\omega_{s,i,t}$ is the share of total expenditure $E_{i,t} = w_{i,t}L_{i,t} - \tau_t L_{i,t}$ on goods from sector s, $P_{s,i,t}$ is the sector s price index, and $P_{i,t}$ is the overall price index. Then expenditure on each sector is given by

$$E_{s,i,t} = \omega_{s,i,t}\left(w_{i,t}L_{i,t} - \tau_t L_{i,t}\right), \forall s \in \{A, S, M\} \tag{11}$$

where $L_{i,t}$ is the total population in location i and time t.

Goods market clearing and trade

There are three sources of demand for each good: final demand by consumers, intermediate demand, and final demand by the government. Manufactured goods trade incurs iceberg costs $\tau_{i,n,t}$. Under perfect competition in input markets, total manufacturing wages are proportional to total sales,

$$w_{i,t}l_{M,i,t} = \alpha \left[c_{M,i,t}^{1-\sigma} \left(\sum_{n \in N} \tau_{i,n,t}^{1-\sigma} \frac{E_{M,n,t} + \frac{1-\alpha}{\alpha} w_{n,t} l_{M,n,t}}{P_{M,n,t}^{1-\sigma}} \right) + \kappa_{i,t} G_t \right], \tag{12}$$

where $c_{M,i,t} = \frac{w_{i,t}^{\alpha} P_{M,i,t}^{1-\alpha}}{f(l_{M,i,t}, l_{M,i,t-1}) T_{M,i,t}}$ is proportional to the unit cost in manufacturing and $P_{M,j,t} = (\sum_i (\tau_{i,j} c_{M,i,t})^{1-\sigma})^{\frac{1}{1-\sigma}}$ is the intermediate goods price index. Agricultural goods are freely traded, and their market clearing conditions are

$$w_{i,t} l_{A,i,t} = \left(\frac{w_{i,t}}{T_{A,i,t}} \right)^{1-\sigma} P_{A,t}^{\sigma-1} \sum_{n \in N} E_{A,n,t}, \forall i, t, \tag{13}$$

where $P_{A,t} = (\sum_i (c_{A,i,t})^{1-\sigma})^{\frac{1}{1-\sigma}}$ is the price index for agricultural goods, with unit costs $c_{A,i,t} = \frac{w_{i,t}}{T_{A,i,t}}$. Finally, market clearing in the services sector yields

$$w_{i,t} l_{S,i,t} = E_{S,i,t}, \tag{14}$$

with $P_{S,i} = c_{S,i} = \frac{w_{i,t}}{T_{S,i,t}}$.

Migration

Adults in i at $t-1$ have ζ_t children. At time t, children make their migration choice to maximize their adult welfare. Utility depends on location-specific amenity $H_{i,t}$, migration frictions $\xi_{i,n,t}$, and idiosyncratic preference $\epsilon_n(\ell)$,

$$U_{i,n,t}(\ell) = H_{n,t} \xi_{i,n,t} \frac{w_{n,t} - \tau_t}{P_{i,t}} \epsilon_n(\ell) \equiv U_{i,n,t} \epsilon_n(\ell). \tag{15}$$

We assume that $\epsilon_n(\ell)$ is distributed Fréchet with shape parameter θ, and thus the probability of choosing n is $\pi_{i,n,t} = P(U_{i,n,t} \geq \max_{j \neq n} U_{i,j,t})$. Note that $P(U_{i,n,t}(\ell) \leq u) =$

$\exp[-U_{i,n,t}^{\theta}u^{-\theta}]$. The migration flow from i to n at time t is

$$L_{i,n,t} = \zeta_t \frac{U_{i,n,t}^{\theta}}{\sum_{j \in N} U_{i,j,t}^{\theta}} L_{i,t-1}. \tag{16}$$

Dynamic equilibrium

Definition 1. *Given initial conditions* $\{L_{n,0,}, l_{M,i,0}, l_{A,i,0}\}$*, an equilibrium is a sequence of wages and labor allocations such that:*

1. *Total sales are equal to payments to labor in each sector as given by equations* (12) *and* (13).

2. *Government budget balance is satisfied.*

3. *Migration flows are given by equation* (16).

4. *Local and global labor markets clear:*

$$\begin{aligned} L_{i,t} &= l_{M,i,t} + l_{A,i,t} + l_{S,i,t}, \forall i, t, \\ \sum_{n \in N} L_{n,t} &= L_t = \zeta_t L_{t-1}, \forall t. \end{aligned}$$

The dynamics of the model are driven by costly migration and by the lagged effect of scale on productivity. Because the scale effects operate with a lag and there is no forward looking behavior, the within-period equilibrium conditions at time t are the same as those of a static model without scale effects, and hence we expect that there is always a unique equilibrium path from every initial allocation.[11] However, depending on the parameter values there may be multiple long-run steady states depending on the initial allocations (Allen and Donaldson, 2020).

[11] Rigorous theoretical conditions for unique equilibria have not been established for this type of static multisector economic geography model without scale economies. However, both theoretical priors and practical experience suggest that it is reasonable to expect uniqueness.

4.2 Estimation using indirect inference

We calibrate the majority of the parameters of the model to standard values in the literature (Table 4) and use indirect inference to estimate parameters of the external economies of scale function $f(l)$. As described earlier in the section, we conjecture that matching the empirically observed difference in treatment effects across locations according to the initial size of the manufacturing sector may require $f(l)$ to diminish in sector size. We parameterize the scale function as cee

$$f(l) = \begin{cases} 250^{\gamma} & l_{Mit-1} \leq 250, \\ l_{Mit-1}^{\gamma} & 250 < l_{Mit-1} < l_{*}, \\ l_{*}^{\gamma} & l_{Mit-1} \geq l_{*}, \end{cases} \tag{17}$$

where γ is the slope parameter and l_* is the location parameter. This function approximates an S-curve in manufacturing employment: initially flat, increasing over some intermediate range, and flat again after local manufacturing scale exceeds l_*.[12] This is a simple two-parameter family that nests the standard constant elasticity case when the location parameter goes to infinity yet is flexible enough (as we will show) to match the empirical moments. It would certainly be ideal to estimate a more flexible functional form, but our current specification is adequate to our purpose and we do not think the data will support robust inference on additional parameters.

We estimate the slope and location parameters using classical minimum distance, sometimes referred to as "calibration with standard errors." We minimize the distance between the empirical ATTs for long-term manufacturing employment, for initially low and high manufacturing locations (column 2 of Table (2)) and the corresponding mo-

[12]The initially flat segment plays a dual role. First, it ensures that the marginal revenue product of labor in manufacturing tends to infinity as local manufacturing employment tends to zero in each county, just as in a model with no scale economies, ensuring positive manufacturing employment in every location. Second, it avoids placing weight on the large number of counties with very small reported manufacturing employment when calibrating the model. We suspect these entries are severely contaminated with measurement error. Our results are robust to varying the cutoff for the initial flat segment.

ments generated by the model,

$$Loss = \left(\beta^{model} - \beta^{data}\right)\left(\beta^{model} - \beta^{data}\right)'. \tag{18}$$

We generate the model moments using the following procedure. First, we guess the parameters of the scale function. Next, we invert the model to obtain the exogenous local sectoral TFPs and amenities that rationalize the observed equilibrium in 1940. Then we shock the model with the empirically observed temporary local war spending shocks $\{\kappa\}_i$ for each location simultaneously, and solve forward for the steady state holding all exogenous TFPs and amenities constant at their 1940 levels, and with population growth set to zero. This establishes the "treatment" scenario for each location. Then, for each treated location i, we return to the observed equilibrium in 1940 and set $\kappa_i = 0$, but retain all treatments in other locations at their observed values and solve the model forward again to obtain the control observation for i, i.e. the outcome for county i in the absence of treatment in county i, all else equal. Finally, we compute the treatment effect for i as the difference between county i's manufacturing employment in the two scenarios; averaging across treated counties in each group (initially high or low manufacturing employment) gives the model-implied ATT for each group (β^{model}).

We systematically search the parameter space by repeating this procedure for different pair of parameters (scale and location) using a combination of global and local methods. Once we arrive at the pair that minimizes equation (18), we compute standard errors for the slope and location parameters (collected in the vector ε) using the delta method (Newey and McFadden, 1994) ,

$$Z = (G'G)^{-1}G'VG(G'G)^{-1}, \tag{19}$$

where Z is the variance-covariance matrix of $\hat{\varepsilon}$, V is the variance-covariance matrix of the empirical moments (computed via block bootstrap), and $G = \partial\beta^{model}/\partial\varepsilon'$, which we

TABLE 4. Summary of calibration

Parameter	Notation	Value	Source
Migration elasticity	θ	3.3	Monte, Redding, and Rossi-Hansberg (2018)
Trade elasticity	σ	8	Eaton and Kortum (2002)
Manufacturing input share	α	0.5	Valentinyi and Herrendorf (2008)
Elasticity of substitution across sectors	η	0.33	Comin, Lashkari, and Mestieri (2021)
Sectoral share parameter	ρ_s	See notes	Add up to 1
Population growth	ζ_t	See notes	Population data
Slope parameter in scale function	γ	0.19	Indirect inference
Location parameter in scale function	l_*	1464	Indirect inference
Calibrated parameter in scale function	.	250	See notes

Notes: Sectoral share parameters are $\rho_{A,t} = [0.142, 0.094, 0.064, 0.035, 0.029, 0.023, 0.029, 0.024]$ and $\rho_{S,t} = [0.049, 0.040, 0.077, 0.129, 0.184, 0.250, 0.379, 0.572]$. Population growth is $\zeta_t = [1.208, 1.067, 1.156, 1.128, 1.209, 1.234, 1.224, 1.290]$. The first location parameter of the scale function is calibrated based on the average manufacturing share multiplied by the 10th percentile of population in 1940. Note that $\rho_{A,t}$, $\rho_{S,t}$ and ζ_t only vary by time for the counterfactual analysis in Section 5, but not for the structural estimation.

compute numerically.

Since our estimation procedure is based on comparing steady states with identical exogenous parameters that differ only by their history, by construction any parameter combination that does not exhibit long-run path dependence in response to the empirical shocks will have $\beta^{model} = 0$, which is quite far from matching the empirical moments. Thus estimation is a matter of finding a parameter vector that generates the correct empirically observed pattern of path dependence, one that primarily affects counties with initially low manufacturing employment. Speaking loosely, the size of the scale elasticity γ is almost exclusively responsible for the existence or non-existence of path dependence in the model, while the location parameter l_* primarily controls which counties exhibit path dependence in response to shocks. Both parameters influence the size of the long-run ATT for those counties that exhibit path dependence, which also depends on the empirical distribution of the shocks.

Calibration and data

To invert the model we require data on $\{w_{i,t}, l_{A,i,t}, l_{M,i,t}, l_{S,i,t}, G_t, \kappa_{i,t}\}_{t\in\mathcal{T}}$ and knowledge of some parameters.[13] We calibrate the trade and migration elasticities to $\sigma = 8$ and $\theta = 3.3$, and the labor share to $\alpha = 0.5$. We set $\eta = 0.33$. Furthermore, we set the sectoral shares ρ_s such that aggregate labor supply adds up to aggregate labor demand in each sector. While these do not vary over time for our structural estimation described above in Section 4.2, they will for some counterfactual scenarios described below in Section 5, in which we calibrate their changes over time from the postwar data. Similarly, while our structural estimation sets population growth to zero, for other counterfactuals we set population growth as $\zeta_t = \frac{\sum_i L_{it}}{\sum_i L_{it-1}}$ using the observed data over the postwar time period. Finally, we set trade costs as

$$\tau_{i,n,t} = \begin{cases} \tau_F distance_{in}^{\frac{1}{\sigma-1}} & i \neq n, \\ \tau_F & i = n, \end{cases} \tag{20}$$

where $\tau_F = \frac{0.3}{\sigma-1}$, and migration costs as

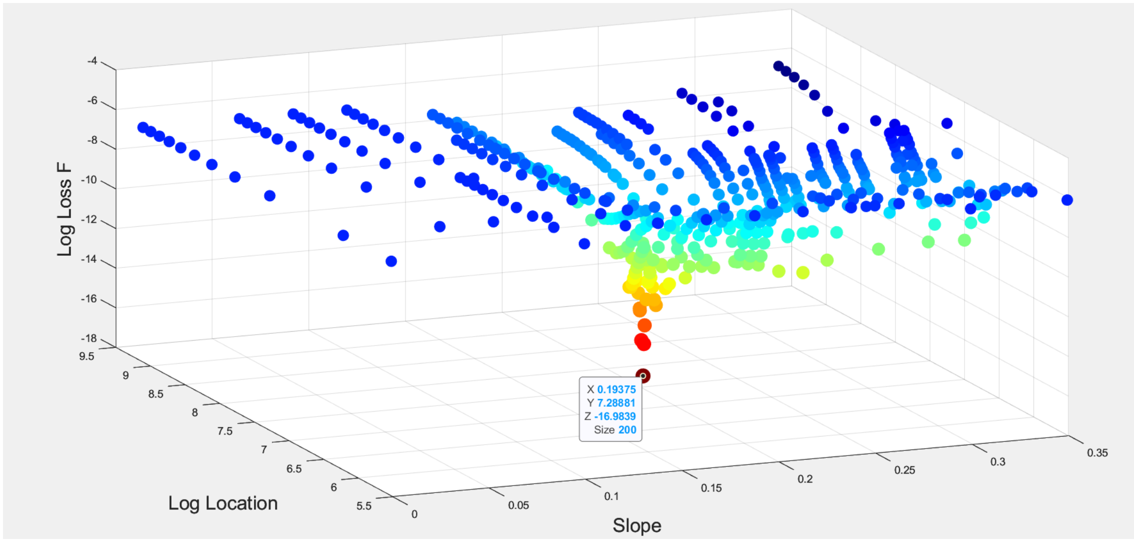
$$\xi_{i,n,t} = \begin{cases} distance_{in}^{-\frac{1}{\theta}} & i \neq n, \\ 1 & i = n \end{cases} \tag{21}$$

where the distance between i and n is defined in miles. Table 4 displays our calibrated parameters. Table A.7 displays the variables we require for the model inversion along with their data counterparts and the sources.

The data used for simulation come from the same sources as those described in Section 2.1. Since we need wage information to invert the structural model, we impute missing wages for the simulation sample using a dynamic panel regression specifica-

[13] See Appendix A.2 for details of the model inversion.

Figure 5. Unrestricted Loss Function



Notes: This graph plots the loss (equation (18)) as a function of the slope and location parameters.

tion. The analysis is based on a harmonized county-level data set of war supply contract values, manufacturing employment, and wages for the decadal years from 1900 to 2000. Additional details are documented in Online Appendix A.1.

4.3 Results

Figure 5 plots the loss as a function of the slope and location parameters. We achieve a loss that is numerically indistinguishable from zero at the minimum, matching the empirical estimates exactly up to the fourth decimal place. We can see that both local and global identification is fairly strong in the depicted region of the parameter space, which we verify holds outside the depicted region as well.

Table 5 shows our parameter estimates and standard errors. We estimate a slope parameter $\gamma = 0.19$ and a location parameter $l_* = 1464$. The location parameter is around the 66^{th} percentile of the county manufacturing employment distribution in 1940. Our estimates therefore imply that manufacturing productivity grows rapidly, about 140% in total, as a county moves from the 10^{th} percentile to the 66^{th} percentile of the manufac-

TABLE 5. Indirect Inference

Parameter	Estimate
Slope: γ	0.19
	(0.02)
Location: l_*	1464
	(993)

Notes: See text for details.

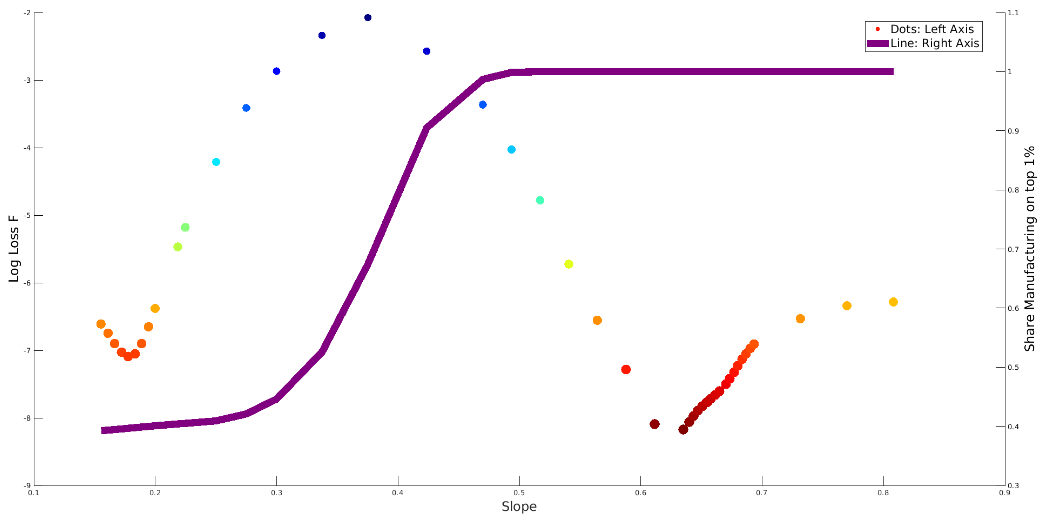
turing employment distribution, with no growth in productivity for further increases in the manufacturing labor force.

It is difficult to make a clean comparison of our estimate of γ to those in the literature. First, we are not aware of any estimates of a local scale function at this level of geographic and sectoral disaggregation. Second, estimation of similar functions, for example scale elasticities in manufacturing at the country level (Bartelme et al., 2024; Lashkaripour and Lugovskyy, 2023), assume a constant elasticity function which is not comparable to ours. Third, as is common in the quantitative literature, our estimation is sensitive to other details of the model calibration such as the value of the trade elasticity and our suppression of other potentially relevant agglomeration or congestion forces. Rather than focus on the specific estimated scale function, which is after all only an approximation that may be highly context-specific, we find it more illuminating to focus on how the estimated scale function fits with the rest of the model to match the observed empirical treatment effects.[14]

Due to its prominence in the literature, it is interesting to evaluate the ability of the constant elasticity model to match the empirical moments. Figure 6 plots the log loss function (left axis) against the scale elasticity, with the location parameter restricted to be infinite (the constant elasticity case). This restricted log loss function has a local minimum at around $\gamma \approx 0.19$, which is our unrestricted estimate, and a global minimum at $\gamma \approx 0.6$. The fit, however, is very poor compared to what we achieve with a

[14]With that said, our value of γ fits comfortably in the range of manufacturing scale elasticities at the country level estimated in Bartelme et al. (2024) and Lashkaripour and Lugovskyy (2023).

Figure 6. Restricted Loss Function



Notes: The dots plot the loss in equation (18) as a function of the slope parameter, assuming an infinite location parameter. The solid line plots the share of manufacturing in the top 1% of counties at the steady state.

finite location parameter; indeed, the lowest loss in Figure 6 is about the same as the highest loss displayed in Figure 5. Moreover, the nature of the long-run equilibrium as γ increases beyond 0.4 or so in the constant elasticity model is "black hole"-like. Figure 6 also plots the share of manufacturing accounted for by the top 1% of counties in the steady state, which moves from being close to the actual 1940 value of about 40% to about 100% when $\gamma > 0.4$. In fact, for larger values of γ our simulations show that 3 counties (Monroe County, MS; Gibson County, TN; and Dinwiddie County, VA) account for essentially all of the nations manufacturing employment in the steady state.[15] We conclude that the constant elasticity model is unable to come close to matching the estimated heterogeneity in the long-run impact of WWII spending and, at least in this setting, the domain over which scale economies that are strong enough to generate path dependence is limited.

[15]In contrast, at our optimal parameter estimates the top 1% share in the steady state is around 40%, as in the data.

5 Counterfactual Analysis

We use our estimated model to re-examine the impact of the temporary component of WWII spending on the long-run evolution of U.S. economic geography. There is a sizable literature in history and economics on the topic, with some arguing that the WWII spending catalyzed long-run manufacturing growth in the South (e.g., Wright (1986); Hooks and Bloomquist (1992)) or the Pacific West (e.g., Rhode (2003)). Others have pointed out that the manner in which spending was allocated, primarily towards counties with a large established manufacturing base, could have had the opposite effect of reinforcing existing patterns of specialization (Fishback and Cullen, 2013; Rhode et al., 2018; Fishback and Jaworski, 2016). We do not directly address this historical debate, which considers the effects of both temporary and permanent (or very long-lived) changes in government spending associated with WWII, but focus solely on understanding the long-run effects that might be attributable to big push dynamics. Specifically, we compare the actual treated world given by the data from the years 1930-2000 to a counterfactual world in which the WWII spending did not occur, but all other exogenous variables evolved as they did in the data. We then simulate the model (without the WWII spending) forward to the year 2000 to generate the counterfactual equilibrium and compare the actual and counterfactual equilibria in the year 2000.

Table 6 highlights the key finding: while the temporary WWII spending had a sizable long-term impact on some individual counties, the effects were not consequential at the aggregate level. Columns 1–2 show the weighted and unweighted average manufacturing shares across counties in the two scenarios, which are quite similar. At the individual county level, the average absolute difference in manufacturing employment (column 3) is a bit over 1%, indicating that at least some counties experienced long-term effects on manufacturing employment. Column 4 shows no major shifts in the concentration of manufacturing in rural areas. Finally, columns 5-6 depicts no substantial effect on real

TABLE 6. Aggregate Differences in Manufacturing Growth in 2000

	(1) Agg. Manuf. Share	(2) Avg. Manuf. Share (County)	(3) Avg. Absolute Value of % Difference	(4) Rural Share Manuf. Emp.	(5) % Change in Avg. Real Income	(6) % Change in Avg. Ex-ante Welfare
Data	0.1227	0.1380	.	0.1107	.	.
Counterfactual	0.1227	0.1376	0.0124	0.1103	0.00045	0.00020

Notes: This table depicts aggregate statistics. Real income is wages divided by the price index. Ex-ante welfare is given by $\Phi_{i,t} = \left(\sum_{j\in N} U_{i,j,t}^{\theta}\right)^{1/\theta}$. The last two columns show the average change in real income and ex-ante welfare per person.

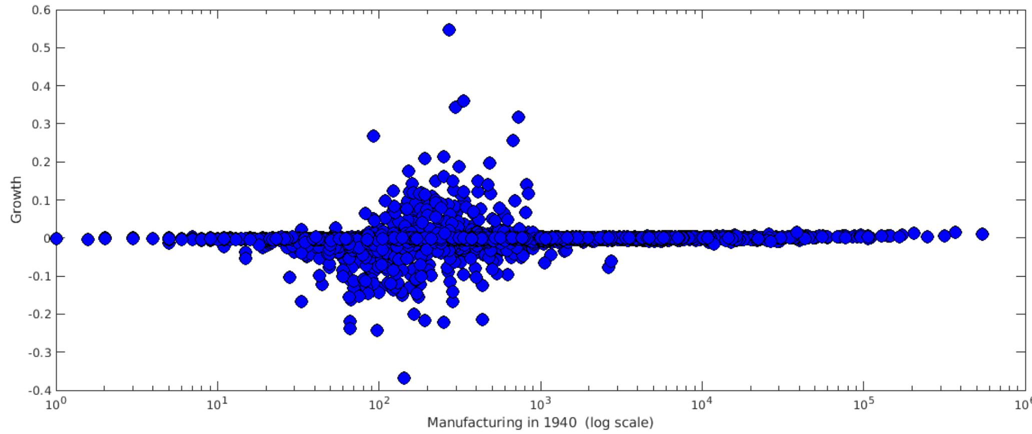
income per person and ex-ante welfare per person.[16]

We shed further light on this finding by examining the changes at the individual county level. Figure 7 plots the log difference in manufacturing employment between actual and counterfactual for each county, against the initial manufacturing employment in 1940. Unsurprisingly, the differences are primarily found in counties with under 1500 manufacturing employees. Many differences are negative, implying that the long run growth of manufacturing in some smaller counties came at the expense of other small counties where manufacturing shrank. The lack of an aggregate impact implies that positive and negative effects roughly balanced out. While efficiency considerations would tend to dictate an overall reallocation of manufacturing activity to smaller counties and a reduction in the overall manufacturing employment share (due to the EoS across sectors η being smaller than 1), the WWII spending shock did not catalyze much of a shift in this direction. This is of course consistent with the historical evidence that regional reallocation of manufacturing activity was not an objective of the wartime contract awards.

The map in Figure 8 highlights regional shifts in manufacturing employment. Overall, the Northeastern and Midwestern manufacturing belt saw modestly increased manufacturing employment, consistent with the view that the WWII spending actually (modestly) reinforced their regional comparative advantage in manufacturing. Most counties

[16]Ex-ante welfare is defined as $\Phi_{i,t} = \left(\sum_{j\in N} U_{i,j,t}^{\theta}\right)^{1/\theta}$.

Figure 7. Long-term Impact of WWII Manufacturing Contracts on Manufacturing Employment Across U.S. Counties



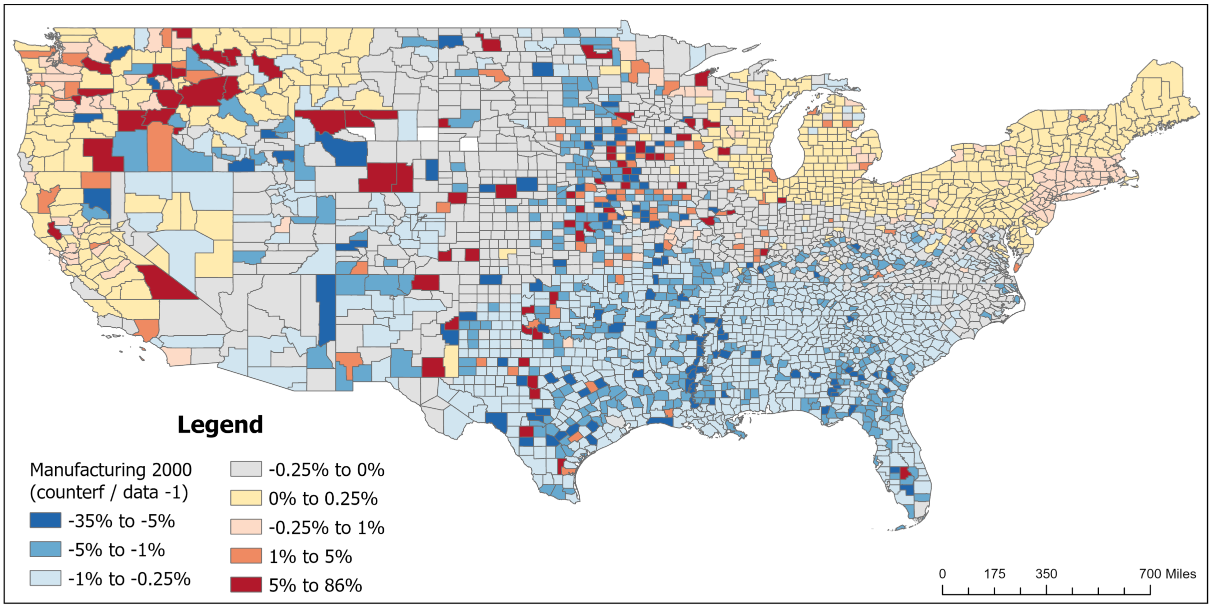
Notes: This scatter compares the outcome in the treated world versus the untreated world in 2000 across counties.

in the South experienced declines in manufacturing employment relative to the counterfactual world. Regionally, the largest growth in manufacturing employment came in the Pacific Northwest, consistent with Rhode (2003), but even there the regional increase in manufacturing employment was small (0.5%). In all regions we see a rich pattern of impacts across counties, with some winners and some losers, even as regional outcomes are correlated on the whole. This reflects the balance between agglomeration economies at the county level, stemming from production externalities embodied in the scale function, and regional agglomeration economies due to the presence of trade costs and input linkages.

In sum, the counterfactual analysis reveals that while temporary World War II contracts had localized long-term impacts, prompting some counties to escape poverty traps, these effects were not significant at the aggregate level. This lack of aggregate impact seems to primarily reflect the limited scope for reallocation in the calibrated model, due to the restricted domain over which coordination failures operate.[17]

[17] In principle, the lack of aggregate impact could also reflect the untargeted nature of the intervention, raising the possibility that some *other* temporary policy might have been more consequential for reallo-

Figure 8. Long-term Impact of WWII Manufacturing Contracts on Manufacturing Employment Across U.S. Counties



Notes: This map compares the outcome in the treated world versus the untreated world in 2000 across counties.

6 Conclusion

This paper makes two primary contributions by analyzing the long-term effects of WWII spending, arguably the largest fiscal shock in U.S. history, on economic geography. Empirically, we show that sustained increases in population and manufacturing employment occur primarily in counties with initially low manufacturing activity. Methodologically, we develop a quantitative multiregion economic geography model that features non-constant scale elasticity and path dependence in long-run equilibria.

The results of our re-evaluation of the long-run geographic effects of a very large temporary shock provide something for both enthusiasts and skeptics of path dependence in economic geography. On the one hand, there is clear evidence of very long-

cation and welfare. While we cannot rule this possibility out completely due to the very large number of feasible policies, we did a number of counterfactual experiments with different policies that redistributed the WWII contracts more heavily towards smaller locations that might experience a big push. We were not able to generate large long-term aggregate reallocations or welfare effects from even substantial redistributions of contracts, suggesting that more targeted temporary policies would also have had limited long-run effects.

run effects of a temporary intervention, in a way that is consistent with the predictions of big push theory. On the other hand, history seems to have mattered but little for the overall geographic allocation of economic activity. The key implication going forward is the importance of having a framework that allows for flexibility in the scale of activity over which coordination failures operate.

References

Allen, T. and D. Donaldson (2020). Persistence and path dependence in the spatial economy. *NBER Working Paper* (w28059).

Azariadis, C. and J. Stachurski (2005). Poverty traps. *Handbook of Economic Growth 1*, 295–384.

Bartelme, D. G., A. Costinot, D. Donaldson, and A. Rodriguez-Clare (2024). The textbook case for industrial policy: Theory meets data. *Journal of Political Economy*, forthcoming.

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

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human Resources*, 436–455.

Comin, D., D. Lashkari, and M. Mestieri (2021). Structural change with long-run income and price effects. *Econometrica 89*(1), 311–374.

Davis, D. R. and D. E. Weinstein (2002). Bones, bombs, and break points: The geography of economic activity. *American Economic Review 92*(5), 1269–1289.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica 70*(5), 1741–1779.

Eckert, F., A. Gvirtz, J. Liang, and M. Peters (2020). A method to construct geographical crosswalks with an application to US counties since 1790. Technical report, National Bureau of Economic Research.

Fishback, P. and J. A. Cullen (2013). Second World War spending and local economic activity in US counties, 1939–58. *The Economic History Review 66*(4), 975–992.

Fishback, P. V. and T. Jaworski (2016). World War II and US economic performance. In *Economic History of Warfare and State Formation*, pp. 221–241. Springer.

Fujita, M., P. R. Krugman, and A. Venables (2001). *The spatial economy: Cities, regions, and international trade.* MIT press.

Garin, A. and J. Rothbaum (2025). The long-run impacts of public industrial investment on local development and economic mobility: Evidence from world war ii. *The Quarterly Journal of Economics 140*(1), 459–520.

Harrison, M. (1998). The economics of World War II: an overview. *The economics of World War II: Six great powers in international comparison*, 1–42.

Hooks, G. and L. E. Bloomquist (1992). The legacy of world war II for regional growth and decline: The cumulative effects of wartime investments on US manufacturing, 1947–1972. *Social Forces 71*(2), 303–337.

Imbens, G. W. and J. M. Wooldridge (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature 47*(1), 5–86.

Jaworski, T. (2017). World War II and the industrialization of the American South. *The Journal of Economic History 77*(4), 1048–1082.

Jaworski, T. and D. Yang (2025). Did war mobilization cause aggregate and regional growth? *Explorations in Economic History*, 101685.

Juhász, R. (2018). Temporary protection and technology adoption: Evidence from the Napoleonic blockade. *American Economic Review 108*(11), 3339–3376.

Kline, P. (2011). Oaxaca-Blinder as a reweighting estimator. *American Economic Review 101*(3), 532–537.

Kline, P. and E. Moretti (2014). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority. *The Quarterly Journal of Economics 129*(1), 275–331.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy 99*(3), 483–499.

Lashkaripour, A. and V. Lugovskyy (2023). Profits, scale economies, and the gains from trade and industrial policy. *American Economic Review 113*(10), 2759–2808.

Li, Z. and D. Koustas (2019). The long-run effects of government spending on structural change: Evidence from Second World War defense contracts. *Economics Letters 178*, 66–69.

Lin, J. and F. Rauch (2022). What future for history dependence in spatial economics? *Regional Science and Urban Economics 94*, 103628.

Monte, F., S. J. Redding, and E. Rossi-Hansberg (2018). Commuting, migration, and local employment elasticities. *American Economic Review 108*(12), 3855–3890.

Murphy, K. M., A. Shleifer, and R. W. Vishny (1989). Industrialization and the big push. *Journal of Political Economy 97*(5), 1003–1026.

Nakamura, E. and J. Steinsson (2014). Fiscal stimulus in a monetary union: Evidence from US regions. *American Economic Review 104*(3), 753–792.

Newey, W. K. and D. McFadden (1994). Large sample estimation and hypothesis testing. *Handbook of Econometrics 4*, 2111–2245.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 693–709.

Rauch, J. E. (1993). Does history matter only when it matters little? The case of city-industry location. *The Quarterly Journal of Economics 108*(3), 843–867.

Redding, S. J. and E. Rossi-Hansberg (2017). Quantitative spatial economics. *Annual Review of Economics 9*(1), 21–58.

Rhode, P. (2003). After the war boom: Reconversion on the US Pacific Coast, 1943-49.

Rhode, P. W., J. M. Snyder Jr, and K. Strumpf (2018). The arsenal of democracy: Production and politics during WWII. *Journal of Public Economics 166*, 145–161.

Rosenstein-Rodan, P. N. (1943). Problems of industrialisation of eastern and south-eastern Europe. *The Economic Journal 53*(210-211), 202–211.

Valentinyi, A. and B. Herrendorf (2008). Measuring factor income shares at the sectoral level. *Review of Economic Dynamics 11*(4), 820–835.

Wright, G. (1986). *Old South, New South: Revolutions in the southern economy since the Civil War.* Basic Books.

Online Appendix

Initial Conditions and the Biggest Push in U.S. History

Dominick Bartelme, Zhimin Li, Daniel Velasquez

August 2025

A.1 Data Used in Estimation and Simulation Samples

The data for WWII supply contracts come from volumes originally issued by the Civilian Production Administration (formerly the War Production Board), which record all defense contracts greater than $50,000 issued between June 1940 and September 1945. Our measure of government defense spending is the aggregate monetary value of war supply contracts in each county during that period.

The empirical analysis is based on county-level data for the years 1900 to 2000. Data on county characteristics are the same as those in Kline and Moretti (2014) and come from a variety of sources including the Census of Population and Housing and Agricultural Census. In particular, the data for the years 1900 to 1930 comes from the Historical, Demographic, Economic, and Social Data: The United States, 1790-2000, ICPSR 2896 (specifically Parts 20, 22, 24, 26, 29, and 85, corresponding to the 1900, 1910, 1920, and 1930 Censuses, as well as the 1930 Census Parts I and IV Families, and the 1910 Census of Agriculture). However, variables for manufacturing and agricultural employment were constructed from individual-level data sourced from a 1% sample extract of the 1900, 1910, 1920, and 1930 Censuses, available through IPUMS.

Variables for the period from 1940 to 1970 were developed using two primary data sources: the County and City Data Book [United States] Consolidated File: County Data, 1947-1977 (ICPSR 7736) and the Historical, Demographic, Economic, and Social Data: The United States, 1790-2000 (ICPSR 2896, specifically Parts 70, 72, 74, and 76, corresponding to the County Data Books for 1947, 1952, 1962, and 1972). The definitions of the variables were mostly consistent across these two data sources. In case of discrepancies, we use the definitions from ICPSR 2896. The data from 1980 to 2000 was derived from census extracts provided by the National Historical Geographic Information System.

The quality of some key variables is imperfect, particularly at the beginning of our sample period, where significant measurement errors are likely. In the early years, there is no direct data on workers' wages at the county level. As in Kline and Moretti (2014), we proxy the average wage in manufacturing by dividing the total annual wage bill in manufacturing by the estimated number of workers in the industry. Additionally, in some counties, the wage bill data is incomplete or missing. For agriculture, there is no available county-level wage bill data, making it difficult to compute an average agricultural wage.

Since we need wage information to invert the structural model, we impute missing wages for the simulation sample. Specifically, we impute missing wage values for county i in decade t using a dynamic panel regression specification as follows:

$$w_{it} = \alpha w_{it-1} + \beta X_{it} + \gamma t + \theta_i,$$

where we control for lagged wage w_{it-1}, county characteristics X_{it} (which include population, median house value, median rent, ratio of white population, agricultural employment share, urban population share, when possible), county fixed effects θ_i, and a linear time trend.[18]

To avoid issues with changes in county definitions over time due to splits and merges, we use the crosswalk provided by Eckert et al. (2020) to obtain a sample of harmonized

[18] If lagged wage is missing, we drop the first term from the regression equation.

counties between 1900 and 2000. We also exclude counties in Alaska or Hawaii as well as underpopulated counties with population fewer than 1,000 in any decade in the sample period or manufacturing employment fewer than 10 in 1930 from the empirical analysis.

A.2 Model Inversion

We infer TFP in manufacturing, agriculture, and services $(T_{M,i,t}, T_{A,i,t}, T_{S,i,t})$, and amenities $(H_{i,t})$ using observed data on population, sectoral employment, and wages. Given the data $\{w_{i,t}, l_{M,i,t}, l_{A,i,t}, l_{S,i,t}, L_{i,t}, G_t, \kappa_{i,t}\}_{t\in\mathcal{T}}$, where $\mathcal{T} = \{0, 1, ..., T\}$, and parameters $\{\sigma, \eta, \epsilon_s, \theta, \alpha, \tau_{i,n,t}, \xi_{i,n,t}, \zeta_t\}_{t\in\mathcal{T}}$, we invert the model and solve for $\{T_{M,i,t}, T_{A,i,t}, T_{S,i,t}, H_{i,t}\}_{t\in\mathcal{T}}$ for each location, up to a normalization for each variable.

1. The first step is to solve for $c_{M,i,t}, c_{A,i,t}$, and $c_{S,it}$ up to a normalization. Given the data, $c_{M,i,t}, c_{A,i,t}, c_{S,it}$ solves the implicit equations:

$$\begin{aligned}
w_{i,t} l_{M,i,t} &= \alpha \left[c_{M,i,t}^{1-\sigma} \left(\sum_{n\in N} \tau_{i,n,t}^{1-\sigma} \frac{E_{M,n,t} + \frac{1-\alpha}{\alpha} w_{n,t} l_{M,n,t}}{\sum_j c_{M,j,t}^{1-\sigma} \tau_{j,n,t}^{1-\sigma}} \right) + \kappa_{i,t} G_t \right], \forall i, t \\
w_{i,t} l_{A,i,t} &= \left(c_{A,i,t}\right)^{1-\sigma} P_{A,t}^{\sigma-1} \sum_{n\in N} E_{A,n,t}, \forall i, t \\
w_{i,t} l_{S,i,t} &= E_{S,i,t}, \forall i, t
\end{aligned}$$

where

$$\begin{aligned}
E_{s,i,t} &= \omega_{s,i,t} \left(w_{i,t} L_{i,t} - \tau_t L_{i,t} \right) \\
\omega_{s,i,t} &= \rho_s \left(\frac{P_{s,i,t}}{P_{i,t}} \right)^{(1-\eta)}, \forall s \in \{A, S, M\} \\
\omega_{A,i,t} &+ \omega_{S,i,t} + \omega_{M,i,t} = 1
\end{aligned}$$

and ρ_s are chosen such that market clearing is assured in all sectors for any unit cost vector $\{c_{M,i,t}, c_{A,i,t}, c_{S,it}\}$. Note that this equation is homogeneous of degree zero in $\{c_{M,i,t}, c_{A,i,t}, c_{S,it}\}$, and that all other variables are known from the data. There is a unique solution (up to scale) of this equation, $\frac{c_{M,i,t}}{c_{M,1,t}}, \frac{c_{S,i,t}}{c_{S,1,t}}, \frac{c_{A,i,t}}{c_{A,1,t}}$.[19]

[19] Our main data source are the Decennial Censuses. Since most of the contract expenditures happened between 1940 and 1950, and we do not observe any data point in those years, we assume $G_t = 0$ when backing out sectoral TFP.

2. The second step is to recover the price index $P_{M,n,t}^{1-\sigma}=\sum_j \frac{c_{M,j,t}^{1-\sigma}}{c_{M,1,t}^{1-\sigma}}\tau_{j,n,t}^{1-\sigma}$ and compute:

$$
\begin{aligned}
\frac{T_{M,i,t}f(l_{M,i,t-1})}{T_{M,1,t}f(l_{M,1,t-1})} &= \frac{c_{M,1,t}}{c_{M,i,t}}\frac{w_{i,t}^{\alpha}\left(\sum_j \frac{c_{M,j,t}^{1-\sigma}}{c_{M,1,t}^{1-\sigma}}\tau_{j,i,t}^{1-\sigma}\right)^{\frac{1-\alpha}{1-\sigma}}}{w_{1,t}^{\alpha}\left(\sum_j \frac{c_{M,j,t}^{1-\sigma}}{c_{M,1,t}^{1-\sigma}}\tau_{j,1,t}^{1-\sigma}\right)^{\frac{1-\alpha}{1-\sigma}}}, \forall i,t \\
\frac{T_{S,i,t}}{T_{S,1,t}} &= \frac{w_{i,t}/c_{S,i,t}}{w_{1,t}/c_{S,1,t}}, \forall i,t \\
\frac{T_{A,i,t}}{T_{A,1,t}} &= \frac{w_{i,t}/c_{A,i,t}}{w_{1,t}/c_{A,1,t}}, \forall i,t
\end{aligned}
$$

3. The final step is to recover exogenous amenities. Armed with previous solutions, we compute price index $P_{i,t}^{(1-\eta)}=\rho_s\left(P_{s,i,t}\right)^{(1-\eta)}$ and use the migration equation:

$$
\begin{aligned}
L_{n,t} &= \zeta_t \sum_i \frac{U_{i,n,t}^{\theta}}{\sum_{j\in N} U_{i,j,t}^{\theta}} L_{i,t-1}, \forall n,t \\
U_{i,n,t} &= H_n \xi_{i,n,t}\frac{w_{n,t}-\tau_t}{P_{i,t}},
\end{aligned}
$$

which can be solved uniquely (up to scale) for $\frac{H_{n,t}}{H_{1,t}}$.

A.3 Additional Tables and Figures

Figure A.1. Cumulative Probability Distribution of Initial Manufacturing Employment

Notes: This figure shows the cumulative probability distribution of manufacturing employment in 1930 in the sample.

TABLE A.1. Summary Statistics of Pre-WWII Regional Characteristics in 1930

	Treated Counties		Control Counties	
	Mean	Std. Dev.	Mean	Std. Dev.
Log population	10.428	0.596	10.031	0.520
Log manufacturing employment	7.105	1.286	6.138	1.316
Log employment	9.445	0.622	9.015	0.539
Agricultural employment share	0.359	0.180	0.486	0.192
Manufacturing employment share	0.138	0.106	0.094	0.095
Log average manufacturing wage	2.738	0.296	2.631	0.362
% Urban	0.359	0.203	0.216	0.195
Log average farm value	6.815	0.652	6.602	0.619
Log median housing values	10.683	0.382	10.462	0.391
Log median contract rents	10.045	0.414	9.834	0.407
% White	0.880	0.166	0.852	0.203
% Illiterate	0.047	0.048	0.057	0.059
% households owning radio	0.310	0.166	0.247	0.173
% Urbanization Rate	0.033	0.018	0.025	0.018
% Whites foreign born	0.051	0.056	0.043	0.053
Max elevation (meters)	1,675.159	2,084.700	1,761.862	2,292.476
Elevation range (max-min)	1,177.056	1,964.030	1,186.685	2,006.357
Area (squared miles)	780.856	846.258	769.884	691.974

Notes: Treated counties correspond to those that received WWII war supply contract value per capita exceeds the median in logarithmic terms. To better balance treated and control samples, we drop counties whose pre-WWII characteristics yield extreme propensity scores of treatment higher than 99% or lower than 1%. All monetary values are in constant 2000 USD.

TABLE A.2. Contract Allocation and Initial Manufacturing Conditions

	(1) Treat	(2) ContractValue	(3) Treat	(4) ContractValue
Auto and Air Share	0.009	0.080		
	(0.004)	(0.024)		
Auto and Air Share Big			0.127	0.896
			(0.023)	(0.185)
Log Manuf. Emp. 1930	0.015	-0.017	0.025	0.074
	(0.058)	(0.565)	(0.060)	(0.596)
Controls	Yes	Yes	Yes	Yes
Obs.	1,777	1,777	1,780	1,780
Adj. R^2	0.377	0.530	0.382	0.532

Notes: This table shows the relationship between wartime contract allocation and initial manufacturing conditions in 1930 across counties. The dependent variables are treatment indicators and log contract value per capita. Key independent variables are log manufacturing employment levels, predicted log contract value, and employment shares in sectors highly related to the war effort including automobile factories, ship and boat building, wagon and carriage factories, and air transportation. All regressions adjust for the control variables described in the text. Treated counties correspond to those that received WWII war supply contract value per capita exceeds the median in logarithmic terms.

TABLE A.3. Robustness Check Controlling for Industry Composition

	(1) Pop 1930-2000	(2) ManufEmp 1930-2000	(3) Pop 1930-1960	(4) ManufEmp 1930-1960	(5) Pop 1960-2000	(6) ManufEmp 1960-2000
Treat*LowM	0.026	0.036	0.025	0.083	0.027	−0.004
	(0.008)	(0.010)	(0.008)	(0.020)	(0.009)	(0.010)
Treat*HighM	0.009	0.012	0.019	0.060	0.004	−0.022
	(0.008)	(0.009)	(0.009)	(0.015)	(0.008)	(0.009)
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows robustness checks controlling for employment shares in sectors highly related to the war effort including automobile factories, ship and boat building, wagon and carriage factories, and air transportation. All regressions adjust for the control variables described in the text. Treated counties correspond to those that received WWII war supply contract value per capita exceeds the median in logarithmic terms.

TABLE A.4. Reduced-Form Estimation Based on Employment Shares in War-Related Sectors

	(1) Pop 1930-2000	(2) ManufEmp 1930-2000	(3) Pop 1930-1960	(4) ManufEmp 1930-1960	(5) Pop 1960-2000	(6) ManufEmp 1960-2000
AA Share Big*LowM	0.059	0.067	0.068	0.104	0.052	0.025
	(0.015)	(0.015)	(0.017)	(0.034)	(0.015)	(0.017)
AA Share Big*HighM	0.003	0.006	0.023	0.032	−0.009	−0.011
	(0.006)	(0.008)	(0.009)	(0.014)	(0.006)	(0.010)
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows reduced-form effects based on employment shares in sectors highly related to the war effort including automobile factories, ship and boat building, wagon and carriage factories, and air transportation. *AA Share Big* is an indicator variable for employment shares in war-related sector above the median level if counties with positive employment shares in those sectors. Treated counties correspond to those that received WWII war supply contract value per capita exceeds the 50th percentile in logarithmic terms. Standard errors are clustered at the state level.

TABLE A.5. Long-Term Impact on Growth Rate of Outcomes by Contract Types

	(1) Pop 1930-2000	(2) ManufEmp 1930-2000	(3) Pop 1930-2000	(4) ManufEmp 1930-2000	(5) Pop 1930-2000	(6) ManufEmp 1930-2000
Treat*LowM	0.020	0.037	0.053	0.075	0.042	0.058
	(0.006)	(0.008)	(0.019)	(0.019)	(0.014)	(0.013)
Treat*HighM	0.006	0.012	−0.000	0.002	0.001	0.006
	(0.005)	(0.007)	(0.007)	(0.008)	(0.006)	(0.007)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Contract Type	I	I	II	II	III	III

Notes: Treated counties correspond to those that received WWII war supply contract value per capita exceeds the median in logarithmic terms. Treatment status is classified into three categories of contract spending as follows: I) “footwear, textile, chemicals, miscellaneous” (columns 1–2); II) “metal” (columns 3–4); and III) “machinery, electronics, transportation” (columns 5–6). Robust standard errors are reported in parenthesis.

TABLE A.6. Relationship Between Post-WWII Military Spending and War Supply Contracts

	(1) MilitaryExp 1966	(2) MilitaryExp 2000	(3) MilitaryExp 1966	(4) MilitaryExp 2000	(5) MilitaryExp 1966	(6) MilitaryExp 2000
Contract value	0.002	0.027				
	(0.023)	(0.021)				
Treat			0.096	0.438		
			(0.413)	(0.380)		
Treat (median)					-0.093	0.271
					(0.404)	(0.393)
Obs.	50	50	50	50	50	50

Notes: This table shows coefficients of regressing post-WWII military spending in 1966 and 2000 at the state level on war time contract spending or treatment status. Treatment status is defined based on whether the log contract value is above 8 (columns 3 and 4) or the median value (columns 5 and 6). Robust standard errors are reported in parenthesis.

TABLE A.7. Data Sources for Structural Estimation

Variable	Source
Total employment, $L_{i,t}$	Decennial Census
Manufacturing employment, $l_{M,i,t}$	Decennial Census
Agricultural employment, $l_{A,i,t}$	Decennial Census
Wages, $w_{i,t}$	Decennial Census
Government expenditure in manufacturing, $\kappa_{i,t}$	Digitized WWII supply contracts from volumes originally issued by the Civilian Production Administration
Distance between locations (in miles)	NBER Distance Database

Notes: Our sample corresponds to the period 1930-2000.