War Mobilization and Economic Development: World War II and Structural Transformation in India *

Aneesha Parvathaneni and Dean Yang

December 1, 2024

Can temporary wartime mobilization change the long-run development trajectory of an economy? We study how mobilization for World War II in colonial India influenced its subsequent development. From 1939 to 1945, the British colonial government purchased massive amounts of war materiel within India. We study long-run impacts on structural transformation – the transition of employment from agriculture to modern sectors (industry and services) – in Indian districts. Causal identification takes a shift-share approach, exploiting variation across industries in war-related government orders, and variation across districts in their pre-war industrial structure. We find that World War II economic mobilization had positive and significant impacts on long-run development. More than six decades later, districts with higher procurement of war materiel saw greater structural transformation from agriculture towards industry and services, and higher consumption levels, urbanization rates, and in-migration. We find substantial spillovers on services sectors not directly subject to procurement. The majority of structural transformation effects are driven by procurement of heavy industrial goods.

JEL codes: H56, N15, N45, O14, O25
Keywords: war, industrial policy, structural transformation, India

---

*Parvathaneni: University of Michigan (aneeshap@umich.edu); Yang: University of Michigan, NBER and BREAD (deanyang@umich.edu). Zara Jacob, Jinghao Liu, and Ann Yang provided exemplary research assistance. We appreciate the research assistance of Rohan Bhargava, Yuqing Wang, and Lana Yang-Maccini in earlier stages of this project. We thank Prashant Bharadwaj, Roberto Bonfatti, Latika Chaudhary, Mark Dincecco, James Fenske, Rinchan Mirza, and Saumitra Jha for generously sharing data. We are thankful for thoughtful feedback from Dmitry Arkhangelsky, Manuel Arellano, Sam Asher, Chris Barrett, Laura Boudreau, Paula Bustos, Sebastian Calonico, Guillermo Caruana, Julieta Caunedo, Giacomo de Giorgi, Brian Dillon, Mark Dincecco, Esther Duflo, Gilles Duranton, Alfonso Flores-Lagunes, Andrew Foster, Tishara Garg, Jonas Gathen, Victor Gay, Andre Groger, John Hoddinott, Elise Huillery, Clement Imbert, Seema Jayachandran, Ruixue Jia, Joe Kaboski, Gaurav Khanna, Federico Kochen, Pushkar Maitra, Gianmarco Leon, Ben Marx, Craig McIntosh, Devashish Mitra, Melanie Morten, Petra Moser, Karthik Muralidharan, Suresh Naidu, Gerard Padro i Miquel, Kiki Pop-Eleches, Tommaso Porzio, Imran Rasul, Tristan Reed, Mark Rosenzweig, Paulina Rossi, Alex Rothenburg, Anisha Sharma, Sandip Sukhtanktar, Jan Svejnar, Alessandro Tarozzi, Liz Tennant, Gabriel Ulyssea, Eric Verhoogen, and Tom Zohar, from seminar participants at CEMFI, Columbia, Cornell, Michigan, and Syracuse, and from audience members at the conferences: Columbia industrial policy (Nov 2023), ISI (Dec 2023), AMES (Jan 2024), U. Chicago MWIEDC (Apr 2024), UCSD BREAD (May 2024), U. Michigan MDev (May 2024), and Barcelona Summer Forum (Jun 2024). Hardish Bindra and his team at Paradigm Data Services provided expert data entry. We are grateful for U. Michigan Economics Department MITRE Research Grant support. Dean Yang appreciates funding from Spain's State Research Agency under the Maria de Maetzu Unit of Excellence Program (project CEX2020-001104-M.)

1 Introduction

September 1, 1939 is one of the most famous dates in the 20th century, marking the German invasion of Poland and the start of World War II in Europe. The date two days later is less well-known. On September 3, 1939, the British Viceroy of India, Lord Linlithgow, made a brief address on All India Radio announcing that India was at war with Germany. The British brought India into World War II on the Allied side by fiat, without consulting Gandhi, Nehru, or any other Indian political leader (Raghavan, 2017). It could do so because India was a British colony, and would remain so until another famous date in history, August 15, 1947.

India would subsequently make major contributions to Britain's World War II effort. From 1939 to 1945, India served as a major arsenal and war materiel supplier for the British Empire in its war efforts worldwide. The total value of goods procured was phenomenal, amounting to one-sixth of India's pre-war (1938) GDP. This wartime procurement constituted the last major intervention of the British Raj in the Indian economy (Sinha and Khera, 1962).

Can temporary war mobilization change the long-run development trajectory of an economy? We study how the economic mobilization of colonial India for World War II – its supply of materiel for the war effort – influenced independent India's subsequent long-run economic development. We are interested in structural transformation (the transition of employment from agriculture to industry and services) in Indian districts that were exposed to varying degrees of World-War-II-related demand for war materiel.

For causal identification, we exploit variation across industries in the magnitude of World War II purchases in India by the British colonial government, combined with variation across Indian districts in the pre-war presence of industries producing those war-related products. We combine these sources of variation in a shift-share research design. Our analysis traces the dynamics of effects on structural transformation over a century, from 1911 to 2011.

This study is made possible by two innovations on the data front. First, we make use of a unique tabulation of procurement of World War II materiel in India by the British colonial government, which provides the total Indian rupee value of hundreds of distinct procured products. To our knowledge, this data source, Aggarwal (1947), has not previously been used in research in economics. Second, we have made substantial investments in digitizing district economic structure data from Indian Censuses from 1911 to 1981, which previously were not available in electronic form. With district-level employment at the detailed occupation level from the pre-war Census and product-level war procurement from Aggarwal (1947), we can construct our key right-hand-side shift-share variable. The Census data also provide our main dependent variable, structural transformation of the economy from agriculture to the modern (industry and service) sectors.

We find that demand for war materiel during World War II had a positive and statistically significant impact on long-run structural transformation in Indian districts. More than six decades later (through 2011), Indian districts more exposed to World War II procurement see greater transitions of their labor forces from agriculture to the industry and service sectors. The effect shows up very soon after World War II, by 1951, and persists for six decades, through 2011, with roughly the same magnitude. A one-standard-deviation increase in World War II orders per worker leads districts to have 6-7 percentage points higher share of workers in the modern sectors (over one-third standard deviation of the outcome variable).

Impacts are not limited to the specific industrial sectors that produced war-related goods. In particular, we find substantial spillovers of impacts to service sectors that were not directly subject to the World-War-II-related procurement. Growth of service-sector employment accounts for the majority of structural transformation effects, in both the short and longer run. Employment in "consumer" services (whose demand comes from consumers and households, e.g., personal services and retail trade) shows the largest response to war orders across services subsectors. These patterns are highly suggestive of coordination externalities, such as spillovers to other industries via increases in worker purchasing power (Murphy et al., 1989).

We also examine impacts of different types of war orders, in particular procurement in light versus heavy industry. We find that the majority of effects on structural transformation are driven by procurement of heavy industrial goods (which includes vehicles, munitions, and chemicals). This finding may reflect the importance of information externalities as a mechanism, for example learning-by-doing effects (Arrow, 1962) and knowledge spillovers across firms (Hausmann and Rodrik, 2003). Heavy industrial goods were further away from India's initial comparative advantage at the time of World War II, accounting for only a small minority of industrial employment (compared to light industrial sectors such as textiles and footwear). The fact that effects are driven by heavy industry procurement points strongly towards positive externalities due to learning-by-doing and learning spillovers across firms.

We address potential threats to causal identification. A pre-trend analysis from 1911-1931 establishes that districts experiencing higher World-War-II-related demand (as measured by our shift-share variable) were not already experiencing more rapid structural transformation in the pre-war period. We also show that our estimates are robust to controlling for time trends that are related to a wide range of baseline (pre-World-War-II) characteristics of districts (economic characteristics, historical conditions, and geographic features). In addition, we also show that variation across districts in military service of soldiers in the war is not driving the empirical results: estimates are highly robust to controlling for proxies for a district's population in World War II military service.

Our primary empirical analyses focus on outcomes related to the structural transformation of the economy – employment shares in the modern sectors. These data are available in the Indian Census from 1911-2011, allowing us to use panel data methods. But other development outcomes are of interest as well. We therefore also conduct cross-sectional regression analyses under the assumption that the shift-share variable is exogenous conditional on the rich set of 1931 (or time-invariant) district controls. We show that higher World War II procurement leads districts to have higher consumption per capita, lower poverty, higher urbanization, higher nighttime light intensity, higher in-migration, and lower out-migration in 2012. We also find that the increases in consumption and reductions in poverty are primarily urban phenomena; there are no effects on these outcomes among rural households.

Our work contributes to three research areas: the economic impacts of war mobilization, the economics of industrial policy, and the long-run consequences of British colonial policies in India.

Economic Impacts of Wartime Mobilization

We contribute to research on the economic consequences of wartime mobilization. Prior research on the impacts of war production and investment, mostly on the U.S., has found mixed results. Many studies argue that World-War-II-related demand had limited impact on post-war productivity growth (Rhode, 2003; Fishback and Cullen, 2013), for example due to inefficiencies from shifting between civil and military production (Higgs, 2004; Field, 2008; Rockoff, 2012; Jaworski, 2017; Field, 2022).

Other studies have documented positive effects of military spending and investment on both short- and long-run economic outcomes. Several studies find that World War II military spending had positive effects on productivity through the 1950s, owing to economies of scale, learning by doing, public R&D, and government provisioning of plant and equipment (Gordon, 1967; Ruttan, 2006; Ristuccia and Tooze, 2013; Gordon, 2017); Rose (2018) and Bose et al. (2022) study impacts on female labor force participation. Similar short-run effects have been noted in Japan, South Korea, and Taiwan during the Vietnam War (Naya, 1971; Stubbs, 1999). Moretti et al. (2021) find, among OECD countries in recent decades, that government defense-related R&D expenditures have positive spillovers on R&D and productivity growth in the private sector. Studies have also identified longer-run impacts of war mobilization. Garin and Rothbaum (2022) find that government investment in plants for World War II production had long-run positive effects on overall employment and high-wage manufacturing work in U.S. localities. U.S. public R&D investments in World War II have also been found to have long-run positive effects on patenting and high-tech employment in U.S. localities (Gross and Sampat, 2023).[1]

Historians have also argued that World War II stimulated subsequent Indian industrialization (Morris, 1983; Roy, 2016), although there are views to the contrary (Tomlinson, 1996; Kamtekar, 2002). McNeill (1982) (p. 356) also views World War II production as having given "special impetus to Indian industrialization".

We contribute with economic analysis of the impact of war mobilization in a context, India, that is more relevant for developing countries overall than prior research focusing on the U.S. or the OECD. Our work is also distinguished in its analysis of very long-run effects of war mobilization – over six decades since World War II.

Economics of Industrial Policy

Our research also sheds light on the impacts of industrial policy (policies aimed at changing the industrial structure of the economy). Wartime mobilization policies are a type of industrial policy, in that they aim to shift production towards industries that contribute to military capability. While not perfectly analogous, studying the effects of World War II mobilization in India offers valuable insights into the potential impacts of large-scale industrial policy interventions today. Both involve government efforts to shape the industrial structure of the economy, focusing on rapidly developing

[1] Relatedly, Bianchi and Giorcelli (2023) show that U.S. Marshall Plan aid had long-run effects on the development of Italian provinces. Studies in political science argue that war is conducive to long-run growth by fostering state-building and institutional development (Rasler and Thompson (1985), Stubbs (1999), Gupta et al. (2016), Dincecco et al. (2022)).

specific strategic sectors. Government procurement serves as a key policy lever in both cases, often complemented by subsidies, capital investments, and knowledge transfer initiatives. The World War II mobilization in India targeted key sectors like heavy industry and munitions, similar to how modern industrial policies often focus on high-tech or advanced manufacturing. It involved developing new industrial capabilities, overcoming coordination problems, and facilitating technology transfer – challenges that industrial policymakers continue to face today.

Since the beginnings of development economics, scholars have highlighted the potential for industrial policy to promote structural transformation from agriculture to industry (Rosenstein-Rodan, 1943; Nurkse, 1953; Hirschman, 1961). Industrial policy has been seen by many scholars as a key driver of economic development in a number of East Asian countries, such as South Korea and Taiwan (Amsden, 1989; Wade, 1990; Evans, 1995; Rodrik, 1995). Others have argued that industrial policy has been ineffective or even harmful for economic development (Baldwin, 1969; Krueger, 1990; Weinstein, 1995; Beason and Weinstein, 1996; Lee, 1996; Pack, 2000; Lederman and Maloney, 2012).

Justifications for industrial policy (as opposed to *laissez-faire*) point to a variety of market failures, such as information imperfections and the need for learning-by-doing (Arrow, 1962; Hausmann and Rodrik, 2003), coordination externalities (Buera et al., 2021), and labor-training externalities (Rodrik, 2007). In many models of economic growth, there can be low- and high-development equilibria, for example due to financial market incompleteness (Townsend, 1979; Greenwood and Jovanovic, 1990; Bencivenga and Smith, 1991; Acemoglu and Zilibotti, 1997), aggregate demand externalities (Murphy et al., 1989), or credit constraints on human capital investments (Galor and Zeira, 1993). In such growth models, industrial policy can move the economy from a low to a high equilibrium.

We contribute to an emerging literature that exploits historical natural experiments to understand the impacts of industrial policy. Recent such papers include empirical analyses of the South Korean 1970s heavy and chemical industry drive (Liu, 2019; Choi and Levchenko, 2021; Kim et al., 2021; Lane, forthcoming), Finnish World War II reparations (Mitrunen, forthcoming), import trade protection in France (Juhász, 2018), temporary input cost advantages in British shipbuilding (Hanlon, 2020), as well as historical industrial plant establishment in China (Fan and Zou, 2021; Heblich et al., 2022; Bo et al., 2023).[2]

Compared to the literature examining historical episodes of industrial policy, our work is distinguished by its geographic scope covering (nearly) all of India, and thus roughly one-sixth of world population. The Indian context, while distinct in its own ways, provides insights that may be of greater relevance to developing countries more broadly than existing research on historical industrial policy episodes in South Korea, Finland, France, or Britain.[3] The impacts of industrial policy may vary across countries with different initial levels of industrialization. The nature of such heterogeneity is ambiguous in theory; industrial policy could have either larger or smaller effects on subsequent development in initially less-developed places. Our research also takes a very long-run scope compared to most prior studies, over six decades from World War II to 2011. Only two papers examine effects of industrial policy over such a long time span: Juhász (2018) over seven decades in the 19th century, and Heblich et al. (2022) from 1950-2010.

British Colonialism in Indian Economic History

Finally, we contribute novel insights in the literature on the long-run impacts of British colonialism in India. Prior work examines the impacts of direct vs. indirect colonial rule (Banerjee et al. (2005), Iyer (2010)), colonial institutions (Banerjee and Iyer (2005), Gupta et al. (2016), Castelló-Climent et al. (2018), Lee (2019)), railroad infrastructure (Donaldson (2018), Chaudhary and Fenske (2022)), and the India-Pakistan partition (Bharadwaj and Fenske (2012), Bharadwaj et al. (2015), Bharadwaj and Mirza (2019)). Bonfatti and Brey (2023) study how reductions in imports due to World War I trade disruptions affect industrialization and support for the anti-colonial movement in Indian districts.

In this context, our work is unique in examining the impacts of war mobilization on long-run economic development. No prior research in Indian economic history has covered this ground.

[2]Our work is also related to research using frontier econometric techniques to study modern-era industrial policies, as opposed to policies enacted in a distant historical period (such as Nunn and Trefler (2010), Aghion et al. (2015), Alder et al. (2016), Rotemberg (2019), Criscuolo et al. (2019), Fan and Zou (2021), Manelici and Pantea (2021), Giorcelli and Li (2021), Cox (2023), Goldberg et al. (2024), Barwick et al. (2024), Bartelme et al. (2024)), as well as those using structural estimation (Kalouptsidi, 2018; Barwick et al., 2019). Earlier calibration-based analyses include Head (1994) and Irwin (2000). Harrison and Rodríguez-Clare (2010) provide a literature review. Research on place-based policies for local economic development in the U.S. are also relevant, such as Kline and Moretti (2014) and Aguilar et al. (2024). Dell and Olken (2020) on the persistent impacts of the colonial Dutch cultivation system in Java also has elements in common with the industrial policy literature, in highlighting how historical production investments can affect long-run structural transformation.

[3]Among prior historical studies, only Bo et al. (2023) and Heblich et al. (2022) examine a developing-country context (China), and only Heblich et al. (2022) documents (negative) effects persisting to the present day.

Our paper is structured as follows. We first provide an overview of World War II mobilization in India. Following that, we describe our empirical analyses, data, and results. We conclude by discussing the implications of our results for economic policy-making and potential future research directions.

2 World War II Mobilization in India

With the onset of World War II, the British colonial government of India initiated a wide-ranging set of policies to expand Indian production of goods needed for the war effort. Preparation for the World War II effort in India was minimal prior to the outbreak of hostilities. The crisis of the Munich settlement of September 1938 led the government to plan establishment of an Indian government ministry to centralize war procurement; it did not make plans public, but documentation from July 1939 indicates government intentions to establish the "Supply Department" to be in charge of war procurement in India. The government formally established the Supply Department on August 26, 1939, just prior to the September 1 German invasion of Poland (Aggarwal, 1947; Sinha and Khera, 1962).

During the early years of war, the British Empire established the Eastern Group Supply Council to coordinate demand for war supplies across different colonies and avoid duplication of war orders. However, with the loss of Malaya, Burma and other colonies to Japan by mid-1942, alongside difficulties in seaborne transport due to enemy submarine attacks, the British colonial government sought to have India become self-sufficient and produce nearly all war supplies needed for its own defense (Aggarwal, 1947).

War-related public procurement in India was massive: the total value of goods procured over 1939-1946 amounted to 17% of 1938 Indian GDP.[4] War procurement in India took some time to ramp up: of the total rupee value of World War II procurement, nearly two-thirds (64.4%) was procured after April 1942, following the appointment of Archibald Wavell as Viceroy of India. The British colonial government procured goods for the war effort from a wide variety of industries in India, to varying degrees. Our empirical analyses take advantage of this variation in the magnitude of purchases across industries, and the geographic variation in the location of pre-war industries (described below in Section 3).

[4]The total rupee value of war-related public procurement is calculated using Aggarwal (1947). The GDP figure for 1938 is from Appendix Table 6(d), Sivasubramonian (2000).

In addition, the set of government policies to support the war effort included measures such as credit subsidies, subsidies for capital investments, and direct establishment of state-owned firms in key industries. In some cases (such as munitions and machine tools), the government mandated production by private firms, coordinated production across firms (say, to ensure supplies of intermediate inputs), and facilitated knowledge transfer (e.g., via technical assistance missions by foreign experts). The government also supported research institutes to develop substitutes using local materials for goods that were scarce due to war-related trade disruptions.

We would expect that the extent of these other policies to stimulate production in different industries would be highly correlated with the amount of government procurement across industries. The amount of government procurement in an industry can thus serve as a measure that represents both the impact of government procurement *per se*, as well as the set of other policies that are aimed at stimulating production in the industry. Our analyses therefore focus on estimating the impact of the amount of government procurement.

Minimal fighting took place on Indian soil during World War II, but 2.5 million Indian soldiers fought on the Allied side in a number of war theaters, most importantly against the Japanese in Burma (Raghavan, 2017). In principle this military service could also have economic effects on soldiers' origin areas. In analyses below we show that including proxies for district-level participation in military service in World War II has no influence on the estimated effect of the shift-share variable. Military service does not appear to be a mechanism through which effects of our shift-share variable operate.

3 Empirical Analyses

We aim to shed light on the impact of war mobilization on Indian economic development in the long run. We present here analyses examining impacts on structural transformation – the shift of employment from agriculture to the modern sectors (industry and services).

The causal variation we exploit is variation across industries in the magnitude of World War II procurement by the British colonial government of India, combined with variation across Indian districts in the presence of those industries in the pre-war period. We combine these sources of variation to implement a shift-share research design, which we describe in Section 3.1 below.

The sample for analysis is a panel of Indian locations ("districts", numbering around 160, depending on the time period) observed from before to after World War II. In these analyses we take the magnitude of war-related government procurement as the measure of the extent of "war mobilization" across industries. One should therefore interpret our regression coefficients as representing the combined effect of the magnitude of procurement itself, as well as any concurrent government policies to stimulate supply that are correlated with the magnitude of government procurement.

Our primary analyses take the Indian district as the unit of analysis. In Section 3.6 we will also examine outcomes at the level of the national industry (in aggregate for all of India).

3.1 Empirical Approach

To estimate the causal impact of war mobilization on structural transformation of Indian districts, we take a shift-share approach (following Borusyak et al. (2022)) that exploits the district-level incidence of British colonial government World War II procurement across industries.

The intuition for the shift-share strategy is as follows. British wartime procurement varies across industries, with some industries (e.g., munitions) experiencing very high demand, some (e.g., footwear) seeing intermediate levels of demand, and others low or zero wartime demand (e.g., musical instruments, jewelry, pottery). Indian districts also vary in the pre-war presence of different industries, as measured by the distribution of a district's employment across industries. Some have relatively high shares of employment in industries that experienced war-related demand, such as munitions and footwear, while other districts have low such shares. Districts with higher pre-war employment shares in war-related industries should experience higher increases in demand (on a per worker basis) due to war-related government procurement. Our approach involves creating a shift-share variable quantifying the extent to which a district experienced World-War-II-related procurement. This variable will be the causal variable of interest in our analyses.

To account for border changes over time, we combine administrative districts so as to be able to track consistently-defined locations from before to after the war using the Dincecco et al. (2022) district concordance. We aggregate Census data appropriately to map to these combined locations. We continue to refer to these combined locations as "districts". Some districts cannot be included in our current analyses due to data limitations (e.g., pre-war data are absent for much of present-day Rajasthan and Gujarat). We exclude the Andaman and Nicobar Islands since these were occupied by Japan during the war and thus did not provide any war materiel. We also exclude from our analyses districts in the northeast region (which includes Bengal and Assam) because war procurement and production intentionally avoided that region due to the proximity to Japan's military advance in Burma (Raghavan (2017), p. 321).

The shift-share variable, $Shiftshare_d$, is predicted World-War-II-related government procurement per worker in district d:

$$Shiftshare_d = \sum_i S_i \times \omega_{id,1931} \tag{1}$$

The "shifts" in the shift-share are S_i, wartime procurement per worker in industry i: total World War II procurement in industry i divided by the total number of pre-war (1931) workers in industry i in British India (procurement is denominated in real 2011 Indian rupees, INR). This is a measure of the magnitude of war-related procurement across industries. In a subset of 45 "war-related" industries out of 195 industries, there is non-zero procurement.[5]

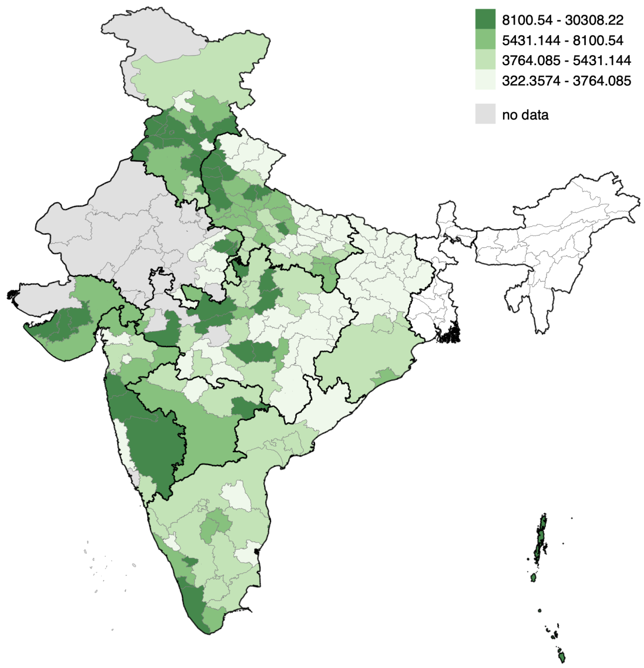
The "shares" in the shift-share are $\omega_{id,1931}$, employment in industry i in district d, as a share of all employed people in district d (measured in the closest pre-war Census year, 1931). $\omega_{id,1931}$ measures the "exposure" of district d to war-related procurement in industry i. We calculate these $\omega_{id,1931}$ shares for each industry for each district.

Taking the product of the shift S_i and the share $\omega_{id,1931}$ for each of a district's industries, and then summing across the district's industries, yields the shift-share variable $Shiftshare_d$: the predicted total value (in INR) of war-related procurement per worker in district d. The spatial distribution of the shift-share variable is shown in Figure 1.

We estimate the following regression equation:

[5] We exclude some parts of British India from the total count of workers in the denominator of S_i. We exclude workers in Burma as well as the Andaman and Nicobar Islands, since they were occupied by Japan during the war and thus did not provide any war materiel. We also exclude workers in the northeast region (which includes Bengal and Assam), since war procurement intentionally avoided that region due to fear of Japanese invasion from Burma (Raghavan, 2017).

Figure 1: Spatial Variation in Shift-Share Variable



Notes: Districts shown are consistent geographic units between 1931 and 2011 Census. Light grey lines demarcate district borders. Black lines demarcate larger-scale "regions" (author defined) of contiguous groups of districts (for estimation of region * time fixed effects). Green shading represents value of shift-share variable, expression (1), in real 2011 Indian rupees (INR). Grey shading indicates districts for which we cannot calculate the shift-share variable due to availability of Indian Census data. Districts in white (in northeast) are not included in analysis, due to proximity to Japanese military advance in Burma. Andaman and Nicobar Islands also not included in analysis since they were occupied by Japan during the war.

$$y_{dt} = \alpha_d + \beta(Shiftshare_d \times Post_t) + \gamma Post_t + \boldsymbol{\delta}(\boldsymbol{X_{d,1931}} \times Post_t) + \epsilon_{dt} \quad (2)$$

The sample for each regression we estimate will be a short two-period panel of districts in one pre-war year (1931) and one post-war year. Post-war periods will be decadal periods corresponding to Indian Census years: 1951 to 2011 inclusive. We will estimate a separate regression in the form of Equation (2) for each post-war period. This allows us to trace the temporal dynamics of effects, by observing how coefficient estimates change across post-war decades.[6]

[6]This approach also allows us to maximize the number of districts represented in each regression sample, since combining districts due to border changes only needs to consider border changes between 1931 and the corresponding post-war decade. If we were to aim to construct districts that had consistent borders across 1931 and *all* post-war decades, we would end up with substantially fewer (and larger) districts in the analysis, reducing the spatial variation used to estimate our effects.

Variables in Equation (2) are as follows. y_{dt} is the dependent variable, the share of employment in the modern sector (industry and services, or non-agriculture) of district d in year t.

$Shiftshare_d$ is the shift-share variable (expression (1)). This is interacted with $Post_t$, an indicator for post-war periods. For ease of interpretation of the regression coefficient, we normalize the shift-share variable to have mean zero and standard deviation one when including it in the regression.

α_d are district fixed effects, which account for any time-invariant differences across districts. $Post_t$ is the time fixed effect (an indicator for the post-war period), and accounts for any changes over time common to all districts between 1931 and post-war period t. ϵ_{dt} is a mean zero error term.

$\boldsymbol{X_{d,1931}}$ is a vector of 1931 characteristics of district d. These are interacted with the $Post_t$ dummy. First of all, the vector includes the "sum of shares" (sum of $\omega_{id,1931}$ across war-related industries within districts). This sum of shares varies across districts (and is never equal to 1), making this an "incomplete shares" case in the Borusyak et al. (2022) framework. Conceptually, the sum of shares represents the share of employment in some war-related industry; inclusion of this variable as a control interacted with $Post_t$ controls for differential trends related with a district's pre-war employment in war-related industries.

In addition, $\boldsymbol{X_{d,1931}}$ includes controls for baseline (pre-war or time-invariant) economic, historical, and geographic characteristics of districts. Interacting $\boldsymbol{X_{d,1931}}$ with $Post_t$ accounts for differential time trends associated with baseline characteristics of districts. Economic controls include share of employment in industry and share of employment in services (share of employment in agriculture is the omitted category). These controls account for any differences in trends across districts related to their pre-war economic characteristics (e.g., if areas that were already more industrialized prior to the war were on different time trends). In addition, economic characteristics include share of employment in heavy industry, years of prior railroad access, log population, share of population employed, and population density as key pre-war characteristics that may also be associated with differential time trends. Historic controls include share of population under British direct rule and historical conflict within 250 km (years 1000-1757), from Dincecco et al. (2022).

The vector $\boldsymbol{X_{d,1931}}$ also includes region fixed effects (for 11 regions); with this interacted with $Post_t$, estimates will be based only on variation in $Shiftshare_d$ within (and not across) regions. Geographic controls include temperature, precipitation, slope, elevation, land area, and caloric yield in agriculture. Finally, to assess whether correlations between military service and war materiel procurement confounds effect estimates, the vector includes World War II casualties per million, martial castes per thousand, and an indicator for non-missing military controls (from Jha and Wilkinson (2012)).

β is the coefficient of interest, and is interpreted as the causal impact of a one-standard-deviation increase in the shift-share variable on the share of employment in the modern sectors of the economy. It is identified from changes in the dependent variable for a district over time that are associated with the district's value of the shift-share variable, net of time trends associated with the vector of controls $\boldsymbol{X_{d,1931}}$.

In the Borusyak et al. (2022) shift-share approach, causal identification depends on the exogeneity of the shifts (shocks), rather than the shares. Our identification assumption is that World War II purchases from industry i are as good as randomly assigned (conditional on district-d-level pre-war controls). Shares $\omega_{id,1931}$ can actually be endogenous.

We provide a partial test of the identification assumption by showing a pre-trend ("placebo" or "false" experiment) regression analysis alongside the main regression results. This is analogous to tests of "parallel trends" in difference-in-difference research designs. The pre-trend test will show that the pace of structural transformation (the change in the share of employment in the modern sectors) was not faster in the pre-war decades in districts that *would in the future* experience higher World-War-II-related procurement (districts that would have higher $Shiftshare_d$.) This test rules out that government World War II procurement was targeted (intentionally or inadvertently) towards districts that were already on steeper economic growth trajectories prior to the war.

3.2 Data

Our most unique data source is the reference we use to construct our shift-share "shifts", S_i (government wartime procurement in each industry i). The data come from the book *History of the Supply Department* (Aggarwal, 1947). This source reports the value of World-War-II-related procurement by the British colonial government of India, in Indian rupees (INR), for 385 detailed products from 1939 to 1946. These so-called "supply orders" were placed by the Supply Department of the

Table 1: Major Product Categories

Product Category	Total Value of Orders (2011 INR)	% of Total
Food Items	156,294,371,446	21.02
Woollen Textiles (Other than factory production)	95,787,392,268	12.88
Cotton Textiles (Other than factory production)	93,055,582,789	12.51
Factory Items - Clothing	77,001,141,426	10.36
Engineering Stores	62,503,547,062	8.41
Vehicles	56,885,281,672	7.65
Miscellaneous	48,042,672,398	6.46
Factory Items - Munitions	43,317,151,048	5.83
Leather	42,808,598,866	5.76
Timber	37,440,779,350	5.04
Factory Items - Harness and Saddlery Items	24,270,857,059	3.26
Shipbuilding	6,186,038,960	.83
Total	743,593,410,560	

Notes: Data are from Aggarwal (1947). Total value of procured products by category are summed across 1939-1946, expressed in real (2011) Indian rupees (INR) (Sivasubramonian 2000).

colonial government of India, which was responsible for sourcing goods for the World War II effort from India. The supply order data are reported at the national (India) level, by product.

We display the top 20 of the 385 products by value of goods procured in Appendix Table A1. The largest product, cotton tent components, saw World War II procurement amounting to INR 63 billion. The other products in the top 20 illustrate the variety of products procured, across food products, clothing, footware, and heavy equipment. Table 1 aggregates products into major categories. Food items is the top category, amounting for 21.0% of total procurement by value. Three categories related to textiles and clothing each amount to 10-13% of the total. Heavier industrial categories are also present, such as engineering stores (8.4%), vehicles (7.7%), munitions (5.8%), and shipbuilding (0.8%).

Data on the shift-share "shares" $\omega_{id,1931}$ of employment by industry are from the 1931 Indian Census (the last Indian Census before World War II). We use data from this pre-war Census to ensure the shares are predetermined with respect to World War II. We create a concordance between the 385 products in Aggarwal (1947) and the 195 occupations in the 1931 Indian Census.

We display the value of the shift S_i (total purchases over 1939-1946 per worker in the industry, in real 2011 INR) for the top 10 industries in Table 2. The value of S_i is largest in the following three industries: making, assembling or repairing motor vehicles or cycles (INR 5,976,304); ship, boat, aeroplane builders (INR 3,223,575);

Table 2: Top 10 Industries by Orders per Worker

Census Occ Code	Census Occupations 1931	Order per worker (INR)
91	Making, assembling or repairing motor vehicles or cycles	5,976,304
93	Ship, boat, aeroplane builders	3,223,575
58	Makers of arms, guns, etc	2,161,173
66	Matches, fireworks and other explosives	2,134,918
177	Architects, surveyors, engineers and their employees	1,830,416
81	Others (Food industry)	1,611,488
70	Others (Chemical products)	1,044,478
94	Heat, light, electricity, motive power, etc	996,502
44	Jute pressing, spinning and weaving	807,801
46	Wool carding, spining and weaving	654,857

Notes: Calculated using product-level war procurement data from Aggarwal (1947), concorded to occupations in the 1931 Indian Census. Orders are summed across 1939-1946, in real (2011) Indian rupees (INR).

makers of arms, guns, etc. (INR 2,161,173); and manufacture of matches, fireworks, and other explosives (INR 2,134,918). On the other end of the scale, S_i takes very small values for potters and makers of earthenware (INR 940) and cabinet makers, carriage painters, etc. (INR 543), and is zero for other industries (e.g., jewelry, musical instruments).

To get a sense of the variation in the shares $\omega_{id,1931}$ (share of pre-war employment in industry i in district d), consider the cotton spinning, sizing, and weaving industry. The standard deviation of $\omega_{id,1931}$ for this industry across districts is 0.015. The maximum of $\omega_{id,1931}$ for this industry is 0.089, for a district consisting of Bijnor (Uttar Pradesh) and its surrounding rural area. Surguja district (Chhattisgarh) is at the median, with $\omega_{id,1931}$ of 0.0126. At the other extreme, the district of Dang (Gujarat) has an $\omega_{id,1931}$ of zero for this industry.

Since district borders change over time, we use the Dincecco et al. (2022) concordance to define districts that are consistent geographical units between 1931 and any post year that we consider in our regression analysis. We refer to these consistent geographical units as "districts". These are shown with grey borders in Figure 1 for the 1931-2011 sample. We also combine multiple districts to form geographically contiguous areas which we call "regions" (the areas surrounded by black borders in Figure 1). These 11 regions are the basis of the Region * $Post_t$ fixed effects included in the regression.

For data on our outcome variable (share of employment in modern sectors), as well as 1931 economic controls, we conducted data entry of tabulated district-level variables from the 1911, 1921, 1931, 1951, 1961, 1971, and 1981 Indian Censuses. (There was no census in 1941.) Creation of the exposure shares $\omega_{id,1931}$ also required us to conduct data entry for employment by industry from the 1931 census. Census data for 1991, 2001 and 2011 were already available in electronic form.

The summary statistics for key variables are shown in Table 3. The share of employment in the modern sectors (non-agriculture) rises between 1931 and 2011 by approximately 9 percentage points, indicating some structural transformation over the course of 80 years. There is considerable variation in the shift-share variable $Shiftshare_d$: it has mean INR 6,058 and standard deviation INR 4,957 (2011 INR). A point of reference for these magnitudes is wartime Indian GDP per capita, 12,133 (in 2011 INR); two-fifths of per capita GDP at the time.[7]

All data in Indian rupees (INR) in this paper are expressed in real 2011 units. Conversion to real 2011 units uses GDP deflators from Sivasubramonian (2000) and the World Bank's World Development Indicators.

3.3 Exogeneity of War Procurement

In the Borusyak et al. (2022) shift-share approach, causal identification is based on the exogeneity of the shifts (shocks), rather than on the exogeneity of the shares. Our shifts are industry-i war procurement per capita, S_i. The shares are $\omega_{id,1931}$, employment in industry i in district d, as a share of all employed people in district d.

Our identification assumption is therefore that the industry-level war procurement shocks S_i are as good as randomly assigned (conditional on district-d-level controls). The exposure weights (shares) $\omega_{id,1931}$ can actually be endogenous.[8]

An example of a failure of this assumption would be if an industry's war procurement S_i were correlated with industry characteristics such as its size, whether the industry is a heavy (vs. light) industry, or the industry's geographic concentration. These characteristics could be associated with subsequent industry growth, separately from any effects due to war procurement in the industry. We might then worry that districts with higher shares of pre-war workforce in industries with higher war procurement might have higher post-war growth because of these particular industry characteristics, and not because of the procurement *per se*. Our estimate of β in equation (2) could then be biased by any ongoing trends related to the pre-war

[7] This GDP per capita figure is the 1936-45 average, expressed in 2011 INR (Sivasubramonian, 2000).

[8] In the Goldsmith-Pinkham et al. (2020) approach, the shares must be considered exogenous.

Table 3: Summary statistics for 1931-2011 sample

Variables	Mean	SD	N
Dependent Variables			
Share employed in modern sector (1931)	.291	.153	166
Share employed in modern sector (2011)	.377	.174	166
Shift-share Variables	6058.347	4957.208	166
$Shiftshare_d$			
Sum of shares ($\sum_i \omega_{id,1931}$)	.083	.053	166
Shiftshare (Industry)	5749.457	4288.539	166
Sum of shares (Industry)	.076	.039	166
Shiftshare (Light Industry)	4244.603	2703.163	166
Sum of shares (Light Industry)	.064	.035	166
Shiftshare (Heavy Industry)	1504.854	2734.063	166
Sum of shares (Heavy Industry)	.011	.008	166
Controls (1931)			
Population (000)	1410.131	1589.1	166
Share of employed in industry	.153	.07	166
Share of employed in services	.138	.106	166
Share of population employed	.463	.099	166
Population density	47.447	86.211	166
Direct British Rule	.73	.433	166
Historical conflicts within 250km (1000–1757)	.127	.138	166
Years of prior railroad access (to 1934)	47.849	20.494	166
Temperature	24.91	3.515	166
Precipitation	1186.036	554.793	166
Slope	1.324	2.635	166
Elevation	403.435	595.853	166
Total land area (sq km)	.049	.074	166
Avg max caloric yield (000s)	6655.125	1150.366	166
WWII casualties per million	206.224	472.84	166
Martial castes per thousand	19.99	51.053	166
War controls non-missing	.789	.386	166
Cross-sectional Outcomes (2011)			
Consumption Per Capita	21000	5392.213	162
Poverty Rate	.214	.119	162
Share of Pop. Urban	.278	.179	166
Log Mean Nighttime Lights	.339	.612	166
In Migration	.144	.079	166
Out Migration	.102	.056	166
Rural Consumption Per Capita	18000	3854.923	161
Rural Poverty Rate	.239	.126	161
Urban Consumption Per Capita	28000	5836.719	161
Urban Poverty Rate	.136	.074	161

Notes: is industry and services (i.e., non-agriculture). is equivalent to 1931 share of employment in war-related industries. Sum of shares and control variables are interacted with post-war indicator ($Post_t$) when included in regression to account for time trends associated with pre-war characteristics.

characteristics of industries that received higher war procurement.

It is plausible *a priori* that industry-level war procurement S_i was exogenous. From the standpoint of 1931, the pre-war "base" year of our analyses, World War II was likely unanticipated by British and other governments. While Japan's invasion of Manchuria in 1931 raised international concerns, there was little indication in that year that a global conflict on the scale of World War II was imminent. Hitler had not yet come to power in Germany, and the Nazi party's future dominance was not assured. Many believed the Treaty of Versailles and League of Nations, despite their flaws, would prevent another major European war. Even military planners were focused on colonial policing and local conflicts rather than preparing for a multi-continent total war resembling what would unfold after 1939 (Steiner, 2011). Our estimates are therefore unlikely to be clouded by anticipation of the future war procurement by firms or Indian colonial government policy-makers (i.e., there are plausibly no effects of an industry's being treated in the future on the industry's outcomes in our pre-treatment period). The fact that World War II was unanticipated from the standpoint of our 1931 base year makes it plausible to consider industry war procurement S_i exogenous.

Since exogenous variation in this framework derives from the S_i shifters, we test statistically for balance in S_i with respect to industry-i characteristics, as recommnded by Borusyak et al. (2022). Equation ((2)) includes extensive controls for pre-war district industrial structure (in particular, share of employment in industry and heavy industry), and variation in the shift-share variable is essentially all coming from variation in war procurement across industries (not from agriculture or services). Therefore, the relevant balance tests to conduct are across the 75 industries (occupations). With these 75 industries as observations, we regress industry-level war procurement S_i on pre-war industry-i characteristics.[9]

The industry characteristics we examine are all pre-shock (based on 1931 Census data). We first consider the aggregate scale of the industry in India: its share of all employment (column 1). The second characteristic (column 2) is an indicator for a "heavy" industry, which captures the industry's capital intensity.[10] In column 3 we examine an industry's geographic concentration, defined as the Herfindahl-Hirschman

[9]Following Borusyak et al. (2022), observations in these regressions are weighted by the industry's average employment share $\omega_{id,1931}$ across districts.

[10]This heavy industry indicator is the same variable included in the vector of pre-war controls $\boldsymbol{X_{d,1931}}$ in equation ((2)) and to determine light and heavy industries in the analysis of section 3.4.4.

Index of the industry's district-level employment share of all workers in the industry across India (higher values indicate the industry is concentrated in fewer districts across India). In a final regression, we include all of the above independent variables on the right hand side.

Regression results are in Appendix Table A2. We run these regressions for the industry characteristics defined over all of British India (Panel A), as well as for just the districts included in our empirical analyses (Panel B); the latter excludes Bengal and the Northeast, as well as districts that cannot be included due to missing data (the gray areas in Figure 1).

In either panel, we find no statistically significant relationships between pre-war industry characteristics and the level of war procurement S_i. We reject joint significance of the right-hand-side variables in the last column. These results provide support for the assumption that industry-i war procurement can be considered as-good-as-randomly assigned.

Another statistic that Borusyak et al. (2022) recommend calculating relates to whether the variation across shocks (industry-level war procurement, in our case) is sufficiently dispersed; otherwise, the asymptotic requirements underlying their shift-share estimator are unlikely to be satisfied. In our case, because the key variation is across only industrial sectors, the relevant statistic is the inverse Herfindahl-Hirschman Index (HHI) of the average employment share across the industrial occupations; this can also be thought of as the "effective sample size". The inverse HHI in our setting is 339. This effective sample size should be sufficiently large for the Borusyak et al. (2022) asymptotic approximation to be considered valid.

3.4 Results

We now present regression estimates of the impact of wartime mobilization on long-run structural transformation.

In Table 4, we present estimates for the 1931-2011 sample in the first five columns. The dependent variable is share of employment in the industry and services sectors (the modern sectors). We present β estimates from equation (2) with different sets of controls. β estimates are interpreted as the impact of one-standard-deviation higher shift-share variable on the fraction of workforce in the modern sectors.

Column 1 includes district fixed effects, period fixed effects, and 1931 economic controls interacted with $Post_t$. In column 2, we add interactions of historical con-

Table 4: Regression Results: Impact of World-War-II-Related Government Purchases on Structural Transformation in Indian Districts, 1931-2011

	1931-2011 Sample					1921-1931 Sample (Pre-Trend Test)	1911-1931 Sample (Pre-Trend Test)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Shiftshare $\times Post_t$	0.082***	0.070**	0.071**	0.064**	0.065**	-0.015	-0.021
	(0.032)	(0.032)	(0.034)	(0.031)	(0.032)	(0.011)	(0.029)
District F.E.	YES	YES	YES	YES	YES	YES	YES
Year F.E.	YES	YES	YES	YES	YES	YES	YES
1931 Economic Controls $\times Post_t$	YES	YES	YES	YES	YES	YES	YES
Historic Controls $\times Post_t$	NO	YES	YES	YES	YES	YES	YES
Region FE$\times Post_t$	NO	NO	YES	YES	YES	YES	YES
Geographic Controls$\times Post_t$	NO	NO	NO	YES	YES	YES	YES
Military Controls$\times Post_t$	NO	NO	NO	NO	YES	YES	YES
Num. Obs.	332	332	332	332	332	286	278

Notes: **Dependent variable** is employment in modern sectors (industry and services) as share of total employment. 166 districts observed in 1931 and 2011. All regressions include district and year fixed effects. Controls interacted with $Post_t$ are all from pre-WWII period or time-invariant. **Economic controls** (from 1931 Census) are log population, share population employed, industrial workers as share of employment, service workers as share of employment, heavy industry workers as share of employment, population density, years of prior railroad access, and shift-share "sum of shares" (share of workers in any war-related industry). **Historical controls** (from Dincecco et al. (2022)) are share of population under British direct rule and historical conflict within 250 km (years 1000-1757). **Region fixed effects** are for 11 regions. **Geographic controls** are mean temperature, mean precipitation, mean slope, mean elevation, land area, and maximum caloric yield in agriculture. **Military controls** (from Jha and Wilkinson (2012)) are WWII casualties per million, martial castes per thousand, and indicator for non-missing military controls. Standard errors are exposure-robust, accounting for correlation of shocks across districts, based on estimation of shock-level (industry-level) regressions (Borusyak et al., 2022). Significance levels: * 10%, ** 5%, *** 1%.

trols with $Post_t$. In column 3, we include region fixed effects interacted with $Post_t$, which allows regions to be on different time trends (capturing spatially-correlated time-variant factors such as weather shocks, or region-specific economic trends or government policies); with these included in the regression, the coefficient estimate exploits only variation in the shift-share variable across districts within the same region. In column 4, we add geographic controls interacted with $Post_t$. In column 5, we add controls proxying for military service in World War II interacted with with $Post_t$.

The coefficient on the shift-share variable declines in magnitude slightly between columns 1 and 2 (from 0.082 to 0.070), but remains relatively stable thereafter as additional controls interacted with $Post_t$ are added to the regression. The stability of the coefficient between columns 2 and 3, when region fixed effects (times $Post_t$) are added to the regression, suggests that cross-district spillovers (within regions) are not substantial. The fact that the the coefficient is stable between columns 4 and 5, when military service controls (times $Post_t$) are added to the regression, indicates that the coefficient on the shift-share variable is not also capturing (or confounded with) impacts of military service in the war.

In column 5, with all sets of controls interacted with $Post_t$ included, the coefficient indicates that a one-standard-deviation increase in the shift-share variable raises modern sector employment by 6.5 percentage points. The magnitude of this effect is not small, amounting to about two-fifths of a standard deviation of the outcome variable.

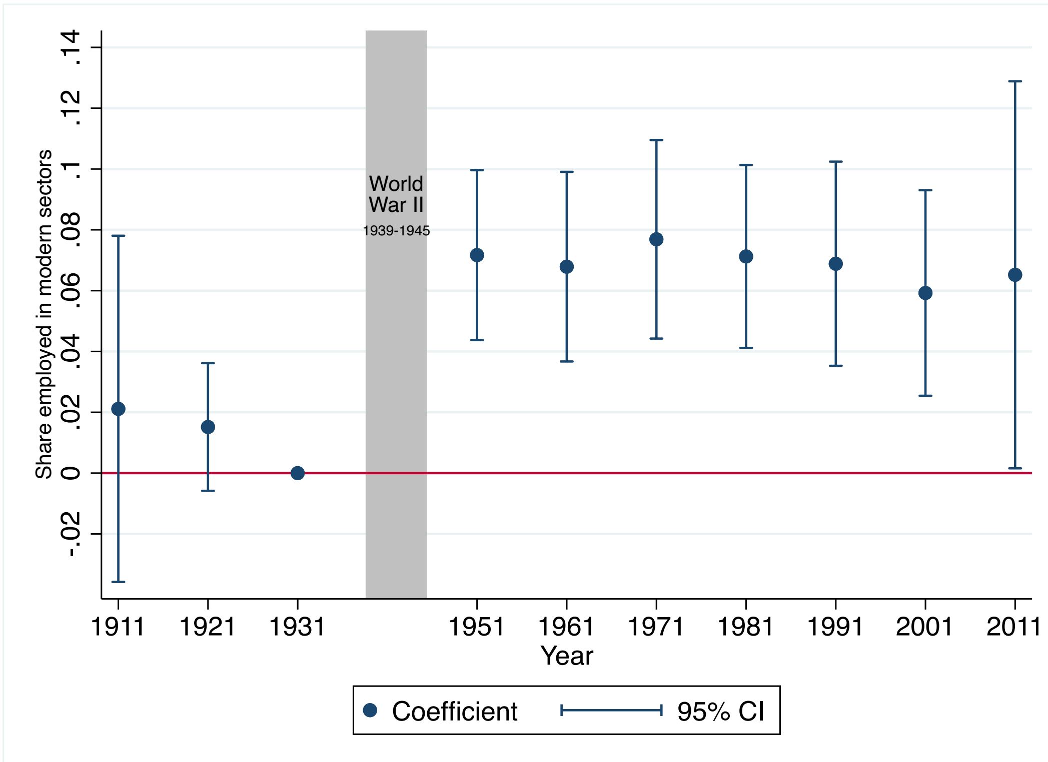
In columns 6 and 7, we present results of pre-trend test regressions ("placebo" or "false" experiments). These pre-trend tests ask whether districts that would *in the future* receive higher World War II product demand were on faster growth trajectories in the pre-war period. The regression specifications are the same as in column 5, but each district's data are from a pair of pre-war decades: 1921 and 1931 in column 6, and 1911 and 1931 in column 7. For these pre-trend tests, in each regression we let 1931 be the "false" post-treatment period (i.e., $Post_t = 1$ in 1931). The coefficient estimate in each of the pre-trend regressions is negative, small in magnitude, and not statistically significantly different from zero. We conclude from columns 6 and 7 that there is no evidence of differential pre-trends in the pre-war period related to districts' future value of the shift-share variable.[11]

3.4.1 Dynamics of Effect Over Time

We have also conducted similar analyses of treatment effects over other decadal time spans, corresponding with Indian Census rounds. Census outcome data (share of employment in industry and services) were already available electronically for 1991 and 2001, and we also conducted data entry for Census outcome data for 1951-1981. We run regressions analogous to those of column 5 of Table 4. For the post-war years, regressions take 1931 as the pre-war year and a decadal observation from 1951 to 2011 inclusive as the post-war year (the latter estimate will be identical to the estimate in column 5 of Table 4). We also show the pre-trend tests using data from 1911, 1921, and 1931 (where the reference year is taken to be 1931 for the purpose of this figure; the pre-trend coefficients are therefore identical to the coefficients in column 6 and 7 of Table 4, but opposite in sign).

[11] Due to missing 1911 or 1921 Census data, the sample sizes in the pre-trend test regressions are smaller than in the first five columns. Results in the first five columns are robust to restricting the sample to the same districts included in the pre-trend test regressions (the sample for either column 6 or 7).

Figure 2: Event Study: Impacts of War Procurement on Employment in the Modern Sectors (Industry and Services)



Notes: Coefficient estimates using specification of column 5, Table 4, for different time periods (with 95% confidence intervals). Point estimates displayed are effects of one-standard-deviation increase in shift-share variable. Dependent variable is district's share of employment in modern sectors (industry and services). For list of control variables and other details, see Table 4. In regressions for post-war periods, pre-war (reference) year is 1931, and post-war year is either 1951, 1961, 1971, 1981, 1991, 2001, or 2011. Figure also shows "pre-trend" tests using data from 1911, 1921 and 1931, with 1931 as reference year for purpose of this figure (coefficients are identical to but opposite in sign of coefficients in columns 6-7 of Table 4).

We present all these coefficient estimates in an event study diagram, Figure 2. The World War II years are depicted as a vertical gray rectangle. In all post-war time periods, the coefficient estimate is positive and statistically significantly different from zero at conventional levels. Districts that received one standard deviation higher orders per worker have 6-8 percentage points higher share of employment in the modern sectors over the entire post-war period. The effect size is quite stable over the decades: the peak is 7.7 percentage points in 1971, and the smallest coefficient is the 2001 value of 5.9 percentage points. The figure also makes clear the absence of a pre-trend in the pre-war period (1911-1931).

In Appendix Table A3, we show a full set of regressions for each post-war decade from 1951 to 2011 inclusive, that vary the set of control variables. (The last panel is identical to Table 4's 2011 regression results, and is included there to facilitate comparison). Figure 2's coefficient estimates are from column 5 of these regressions. Results in Appendix Table A3 indicate that the coefficient on the shift-share variable is robust to inclusion of different subsets of control variables across all post-war decades.

3.4.2 Effects on Industry and Services Employment

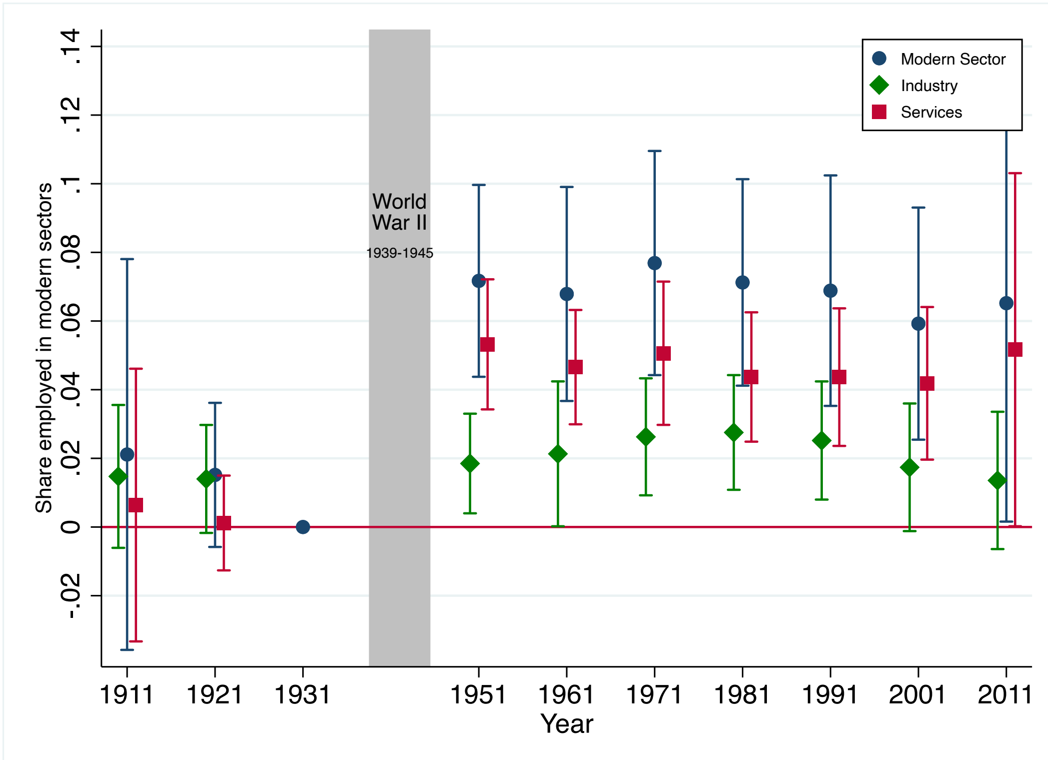
The estimates we have presented so far are effects on total modern sector employment (industry and services). It is also of interest to examine effects on industry and service sector employment separately. This analysis can shed light on cross-sector spillovers, since the vast majority of war procurement was in industrial (not services) sectors. Any such spillovers to the services sector could reflect heterogeneity in income elacticities of demand across sectors (nonhomotheticity in preferences), in particular high income elasticities for services, as shown empirically in U.S. data by Comin et al. (2021). Ahsan and Mitra (2016) highlight that most structural transformation (the transition of employment to modern sectors) in India in recent decades has been driven by growth in the services sector, which Fan et al. (2023) attribute to service-sector productivity growth.

We run regressions analogous to those in Figure 2, but separately for share of employment in services and share of employment in industry. Figure 3 is the event study figure capturing these regression results. The coefficient (and 95% confidence interval) for employment in industry is displayed in green, and corresponding estimates for services are displayed in red. For comparison, the estimates for the total modern sector (industry plus services) are shown in blue (which replicates the results in Figure 2).

The majority of effects on modern sector employment are driven by the services sector. In each time period, the coefficient estimate for services employment is larger than the corresponding coefficient for industry employment. While this result is strongly suggestive of cross-sector spillovers – because the vast majority of war procurement orders are in the industrial sectors – it is possible to make this point more clearly by estimating the impact of war procurement in *only* the industrial sectors, excluding non-industrial sectors.

We therefore run analogous analyses but with an alternate shift-share variable that excludes war procurement in non-industrial sectors: services and agriculture,

Figure 3: Event Study: Impacts on Industry and Services Employment Separately



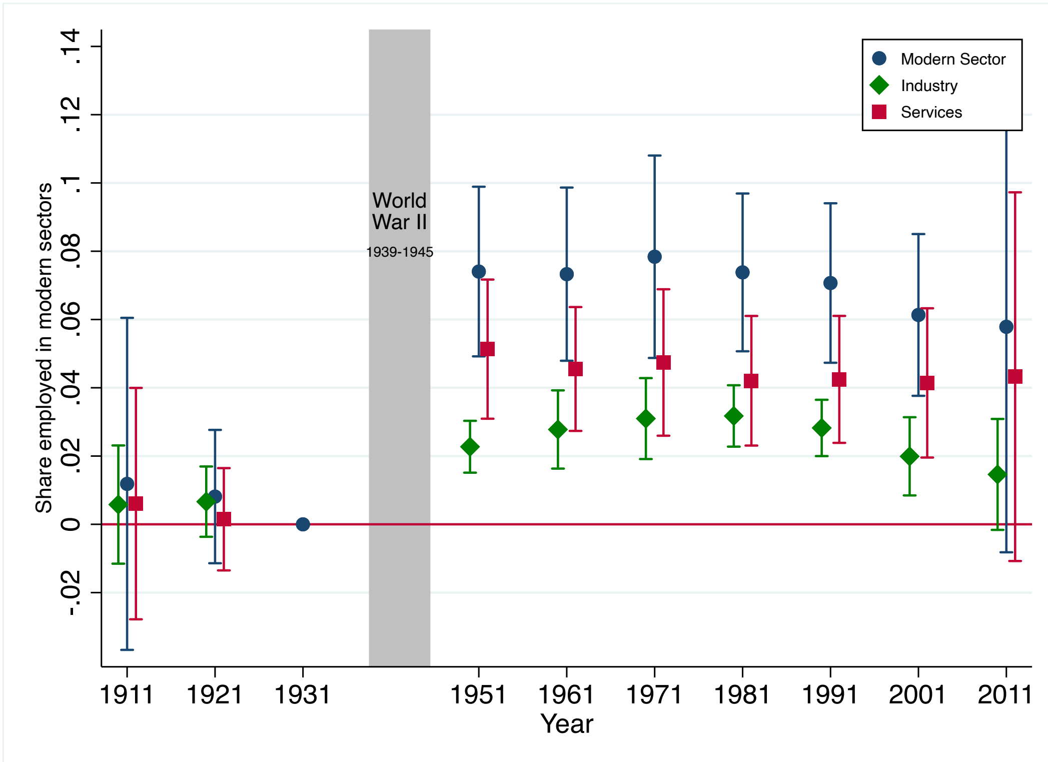
Notes: This figure replicates estimates of Figure 2 (coefficient estimates are blue circles), and adds coefficient estimates for separate regressions where the dependent variables are share of employment in industry (coefficient estimates are green diamonds) and share of employment in services (coefficient estimates are red squares). All other details are as in notes of Figure 2.

which account for 1.5% and 2.5% of war procurement, respectively.[12] Any estimated impacts of the industry-only shift-share measure on services employment would be clearly interpretable as a cross-sector spillover from industry to services employment.

We display the event-study coefficients from this analysis in Figure 4. The coefficient estimates for industry, services, and total modern sector employment are similar in magnitude and statistical significance levels to those in Figure 3. There is a positive effect on both services and industry employment in all periods, with the effect on services employment always larger than that on industry employment. These results reveal quite substantial spillovers of wartime industrial procurement to the service sector.

[12]The regression also includes on the right-hand-side an analogous shift-share variable measuring war orders in the non-industrial sectors (services and agriculture), to cleanly separate effects of industrial orders from non-industrial orders.

Figure 4: Impacts of Industry-Only Shift-share



Notes: This figure replicates estimates of Figure 3, but replacing the shift-share variable with an analogous shift-share variable that excludes procurement in non-industrial sectors (services and agriculture), respectively accounting for 1.5% and 2.5% of procurement by value. The regression also includes on the right-hand-side a shift-share variable measuring war orders in the non-industrial sectors (services and agriculture); coefficients not shown in figure. All other details are as in notes of Figure 2.

3.4.3 Impacts on Services Subcategories

The cross-sector spillovers from industrial war procurement to service sector employment are consistent with models that emphasize the importance of coordination externalities, such as Murphy et al. (1989). The key mechanism in the model of the "big push" in Murphy et al. (1989) is worker purchasing power: when firms coordinate to invest at the same time, aggregate worker purchasing power rises, fueling demand for goods and services broadly in the economy. The economy is thus able to sustain the high-investment equilibrium, associated with higher employment in the modern sectors. This mechanism has been highlighted recently by Goldberg and Reed (2023), who emphasize the importance of domestic market size in driving aggregate household demand in all sectors of the economy, including nontradables such as services. Amodio et al. (2024) find spillovers from industrial sectors to consumer-related service sectors in modern-day Ethiopia.

To shed light on whether household demand for services may be driving the cross-sector spillovers we have documented so far, we examine impacts of the industry-only shift-share variable on categories of services occupations. We examine as an outcome variable the share of employment in "consumer" services that are primarily demanded by households (e.g., retail trade, restaurants, personal services). We also examine "business" services that are mainly purchased by firms (e.g., transport, logistics, accounting, office clerks), as well as the services subsectors of health, education, and government.

Regression results for the 1931-2011 period are in Table 5. All regressions include the full set of control variables (as in column 5 of Table 4 and Figures 2 through 4). The coefficient estimates are all interpreted as the impact of a one-standard-deviation increase in the shift-share variable on the fraction of the workforce in the given services subcategory. The largest coefficient is in the regression for consumer services, 2.4 percentage points, and statistically significant at the 10% level. The coefficient for business services is also substantial (1.6 percentage points, significant at the 5% level), slightly smaller than for consumer services. Coefficients for health, education, and government services are small in magnitude and not statistically significantly different from zero.

Table 5: Regression Results: Impact of World-War-II-Related Government Purchases on Share of Workforce in Service Sub-Categories, 1931-2011

	(1) Consumer	(2) Business	(3) Health	(4) Education	(5) Government
Shiftshare $\times Post_t$	0.02404*	0.01575**	0.00039	0.00077	0.00456
	(0.01308)	(0.00638)	(0.00097)	(0.00158)	(0.00561)
Dependent Variable Mean - 1931	.062	.053	.002	.003	.019
Dependent Variable Mean - 2011	.09	.049	.009	.024	.024
Num. Obs.	332	332	332	332	332

Notes: Dependent variable is employment (as share of total employment) in services subcategories: consumer, business, health, education, and government. 164 districts observed in 1931 and 2011. Independent variable is industry-only shift-share. In the computation of this shift-share, procurement in services and agriculture are excluded. All regressions include district and year fixed effects. Each column includes the following interacted with $Post_t$: economic, historic, geographic and military controls, and region fixed effects (see Table 4 for the full list). Standard errors are exposure-robust, accounting for correlation of shocks across districts, based on estimation of shock-level (industry-level) regressions (Borusyak et al., 2022). Significance levels: * 10%, ** 5%, *** 1%.

The fact that consumer services shows the largest response to the shift-share variable is consistent with theoretical models such as Murphy et al. (1989) that emphasize coordination externalities operating via aggregate worker demand. It also appears that spillovers via firms' demand for business services are not insignificant.

3.4.4 Light- vs. Heavy-Industry Shift-share

It is also of interest to examine impacts of procurement of different types of industrial products, in particular light versus heavy industrial goods.[13] We can construct a shift-share variable representing only procurement of light industrial goods (e.g., clothing, textile products, footwear, food products) and separately a shift-share variable for heavy industrial goods (e.g., vehicles, munitions, engineering stores, and chemical products). The relative magnitude of effects of light vs. heavy industry procurement could reveal the importance of information or learning externalities.

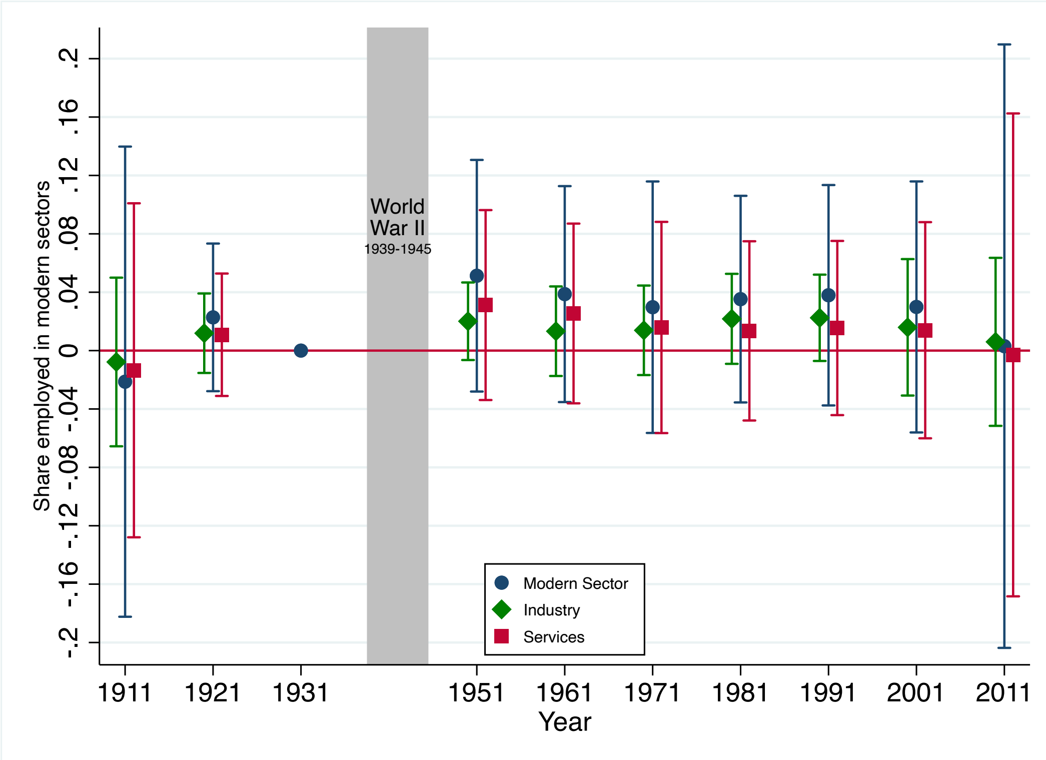
We run regressions for our main outcome variables (employment in modern sector, industry, and services) with shift-share variables defined for light industry only and heavy industry only.[14] We present event-study figures summarizing these results for the light-industry shift-share variable (Figure 5) and the heavy-industry shift-share variable (Figure 6). As in prior analyses, point estimates presented represent effects of a one-standard-deviation increase in the shift-share variable.

It is apparent from comparing Figures 5 and 6 that the largest share of effects we have seen in the prior analyses are driven by heavy industry procurement. Coefficients in Figure 5 reveal that the coefficient on the light-industry shift-share variable is small and never statistically significant in any decade for modern sector employment, services employment, or industry employment. By contrast, patterns of effects in Figure 6 reveal positive effects of heavy industry procurement on modern sector employment overall, and for services and industry employment separately. Coefficient estimates in the services regressions are always larger than coefficient estimates in the industry regressions. Coefficient estimates in the modern sector employment, services employment, and industry employment regressions are statistically significant at conventional levels in all decades, except for 2011 when these coefficients are marginally significant. There is no evidence of worrying pre-trends for any of the three dependent variables.

[13] Hasan et al. (2013) argue that post-war labor regulations and capital market imperfections biased Indian industrialization towards more capital-intensive industries.

[14] For clear separation of estimated effects, the light-industry and heavy-industry shift-share variables are both included on the right-hand-side in each regression. In addition, regressions also include on the right-hand-side the shift-share variable measuring war orders in the non-industrial sectors (services and agriculture), as in regressions of Figure 4.

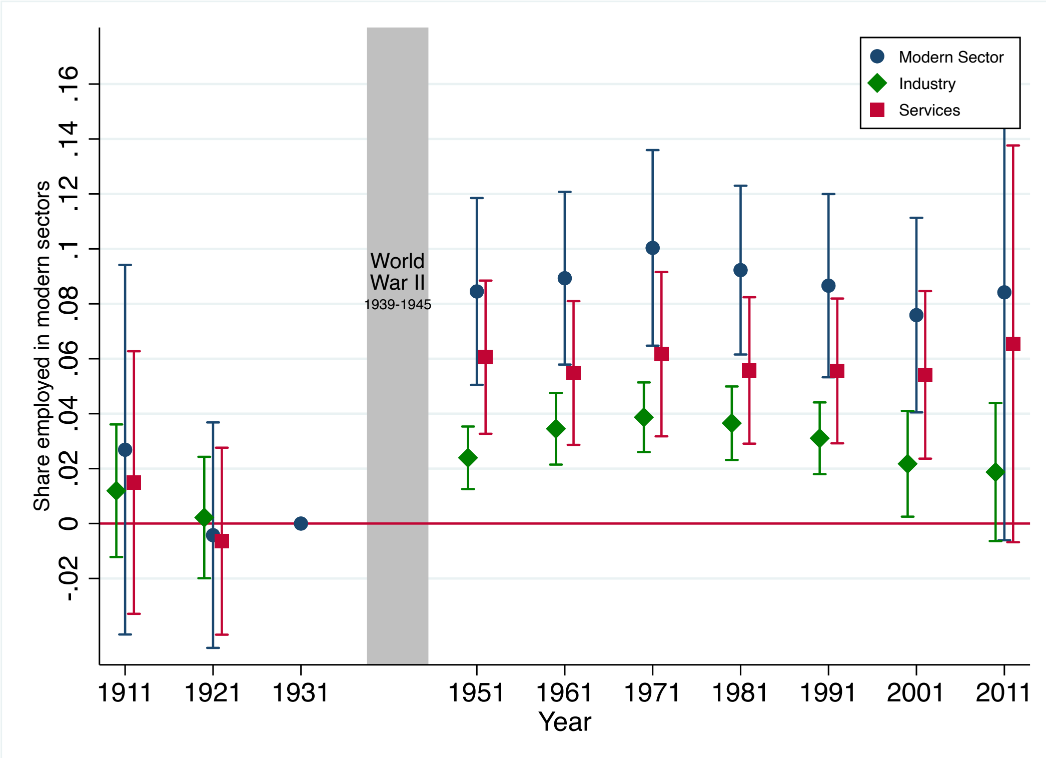
Figure 5: Light Industry Shiftshare



Notes: This figure replicates estimates of Figure 3, but replacing the shift-share variable with an analogous shift-share variable that includes procurement in light industries only. All other details are as in notes of Figure 2.

The fact that structural transformation is primarily driven by heavy industry procurement may reflect the importance of information externalities as a mechanism, for example learning-by-doing (Arrow, 1962; Moser and Voena, 2012) and knowledge spillovers across firms (Hausmann and Rodrik, 2003; Greenstone et al., 2010). Production techniques for heavy industrial goods would have been closer to the world technological frontier than for light industrial goods. The potential for learning-by-doing and knowledge spillovers across firms would therefore have been greater for heavy industrial goods. The fact that effects we find on modern-sector employment are driven by heavy industry procurement points towards these kinds of informational externalities.

Figure 6: Heavy Industry Shiftshare



Notes: This figure replicates estimates of Figure 3, but replacing the shift-share variable with an analogous shift-share variable that includes procurement in heavy industries only. All other details are as in notes of Figure 2.

3.5 Cross-Sectional Analyses of District-Level Outcomes

Our empirical analyses so far have used a two-way fixed effects regression specification with panel data on districts from before and after World War II. Such an approach of course requires district-level data from before and after the war. We are fortunate that Indian Census data allows us to examine a range of outcomes related to structural transformation in a panel setting over this very long time-frame: share of workforce in the modern sectors (industry and services), as well as detailed services subsectors.

Panel data approaches are of course limited to outcomes for which district-level data are available prior to World War II. That said, it is of interest to examine the impact of World War II procurement on a wider set of district-level outcomes.

In this section we examine a range of outcomes related to economic development that are only available in the post-war period. We examine impacts of a district's World War II procurement on consumption per capita, poverty rates, urbanization, nighttime light intensity, and migration outcomes. We also examine consumption and poverty outcomes separately for urban and rural populations in each district.[15]

For these outcomes we cannot use panel data methods, so we run regressions using a cross-sectional analog to our main regression equation (2). We estimate the following cross-sectional regression equation:

$$y_{d,post} = \rho + \theta Shiftshare_d + \boldsymbol{\phi X_{d,1931}} + \epsilon_d \tag{3}$$

The unit of observation is a district. The dependent variable $y_{d,post}$ is a post-war development outcome measured in a time period near the present day, such as consumption per capita, in district d. $Shiftshare_d$ is the shift-share variable (expression (1)); as before, this is normalized to have mean zero and standard deviation one when included in the regression. $\boldsymbol{X_{d,1931}}$ is a vector of 1931 characteristics of district d, and is identical to the corresponding vector used in the prior panel regression equation (2) (economic, historical, geographic, and military controls, plus region fixed effects). ρ is the constant term, and ϵ_{dt} is a mean zero error term.

θ is the coefficient of interest: the causal impact of a one-standard-deviation increase in World War II procurement per capita on the outcome variable. It is identified from variation in the dependent variable associated with the district's value of the shift-share variable, conditional on the control variables $\boldsymbol{X_{d,1931}}$. The identification assumption is that the shift-share variable is exogenous (uncorrelated with the error term) conditional on the full set of economic, historical, geographic, and military controls, and the region fixed effects.

Regression results for a range of present-day development outcomes are in Table 6. In Panel A we present regressions for a range of outcomes for the full district. In Panels B and C we display regression results for urban and rural populations, respectively, for the consumption and poverty outcomes (other outcomes are not available separately for rural and urban populations).

Coefficient estimates in Panel A reveal that World War II procurement has positive long-run effects on a range of measures of economic development. Higher values of the shift-share variable lead to higher 2012 consumption per capita in 2012 Indian rupees (significant at the 5% level) and log consumption per capita (significant at

[15] All these outcomes except nighttime lights are obtained from the 2012 Socioeconomic and Caste Census (SECC) via the SHRUG Development Data Lab (Asher and Novosad 20XX, etc. etc.). Data on nighttime light intensity come from the Earth Observation Group and are for 2023.

Table 6: Regression Results: Impact of World-War-II Related Government Purchases on Modern Development Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Consumption Per Capita	Log Consumption Per Capita	Poverty Rate	Share of Pop. Urban	Log Nighttime Lights	In Migration	Out Migration
Panel A: District Outcomes							
Shiftshare	1,693**	0.060*	–0.005	0.083***	0.271***	0.039**	–0.018**
	(832)	(0.036)	(0.017)	(0.021)	(0.056)	(0.015)	(0.009)
Dep. Var. Mean	20764	9.91	.214	.278	.339	.144	.102
Num. Obs.	162	162	162	166	166	166	166
Panel B: Rural Outcomes							
Shiftshare	260	0.009	0.003				
	(575)	(0.032)	(0.019)				
Dep. Var. Mean	17986.8	9.776	.239				
Num. Obs.	161	161	161				
Panel C: Urban Outcomes							
Shiftshare	1,960**	0.068**	–0.020**				
	(879)	(0.029)	(0.010)				
Dep. Var. Mean	27687.78	10.208	.136				
Num. Obs.	161	161	161				

Notes: **Dependent variables** are consumption per capita, log consumption per capita, poverty rate, share of district population that is urban, log mean nighttime lights, in migration and out migration. Panel A, B and C report the regression results at the district, urban and rural areas respectively. All regressions include district and year fixed effects. Each column includes economic, historic, geographic and military controls. Standard errors are exposure-robust, accounting for correlation of shocks across districts, based on estimation of shock-level (industry-level) regressions (Borusyak et al., 2022). Significance levels: * 10%, ** 5%, *** 1%.

the 10% level). The coefficient in the poverty rate is negative, but not statistically significantly different from zero at conventional levels. The magnitude of the effect on consumption is not small: one standard deviation higher war procurement per capita leads to INR 1,693 higher consumption per capita, which is 8.2% of the mean.

Effects on additional outcomes in Panel A further bolster the conclusion that World War II procurement bolstered modern sector growth. There is a positive effect on urbanization rates in 2012: one standard deviation higher war procurement leads the urban population share to be 8.3 percentage points higher (statistically significant at the 1% level), which is about three-tenths of the mean urbanization rate. Effects are also visible from outer space: districts with one standard deviation higher shift-share have higher log nighttime light intensity (statistically significant at the 1%

level), and the effect is very large (27.1 log points).

If World War II procurement leads to higher long-run economic development, we might expect migrants to be attracted to districts with higher shift-share values, and to migrate less out of such districts. This would lead to positive effects on in-migration (share of a district's population born outside the district), and negative effects on out-migration (share of people born in a district currently living outside of it). Of course, effects on out-migration could theoretically be positive if economic development loosens financial constraints that would otherwise inhibit migration (Bazzi, 2017; Mahajan and Yang, 2020; Khanna et al., 2022).

We examine migration outcomes in the remaining columns of Panel A. Districts with one standard deviation higher shift-share have 3.9% higher in-migration rates. The effect is statistically significantly different from zero at the 5% level, and is large in magnitude, amounting to about one-quarter of the sample mean in-migration rate. Due to the nature of the Census data, the out-migration rate measure is imperfect: it only captures the share of people born in a district who are currently living in other districts *in the same state*. This understates the total out-migration rate. We find a reduction in this measure of out-migration: the coefficient is negative (-1.9 percentage points), statistically significantly different from zero at the 5% level, and large (about one-fifth of the sample mean).

In Panels B and C we examine the effects on consumption per capita and poverty rates in (respectively) rural and urban areas of districts separately.[16] Effects on consumption per capita and poverty rates are prominent in urban areas, and absent from rural areas: in Panel B (rural areas) coefficients are close to zero and not statistically significant at conventional levels, while coefficients in Panel C (urban areas) indicate statistically significant increases in consumption and reductions in poverty rates. The positive development effects of World War II procurement appears to be an entirely urban phenomenon (including an increase in urbanization).

3.6 Industry-Level Analyses

Our prior analyses have emphasized cross-sector spillovers from industry to services. It is also of interest to examine cross-sector spillovers solely between industrial sectors. In particular, we seek evidence of upstream and downstream linkages – indirect

[16] In each panel the sample size is one smaller than in Panel A, because one district is entirely rural (Dangs) and one is entirely urban (Chennai), and so each of these is only included in one of the two panels.

effects extending from industries experiencing World-War-II-related demand to other industries that either supply the directly affected industry (upstream linkages) or that demand intermediate goods from the directly affected industry (downstream linkages). Such linkages were first emphasized by Hirschman (1961) as a rationale for industrial policy, and empirical evidence for such linkages has been found by Choi and Levchenko (2021) and Lane (forthcoming).

To test for upstream and downstream linkages, we conduct analyses at the level of the national industry in India, because upstream and downstream linkages can operate nationally (via trade in inputs), not just within districts. To construct upstream and downstream links between industries, we use the input-output table from Saluja (1972). This table provides information on input-output linkages for 144 industries for the year 1964-1965. We restrict our analysis to the 72 Census occupations that we are able to concord with the industries covered in Saluja (1972). Our outcome variable is log employment in the occupation (nationally), as recorded in the Census.

Let i and j denote the producing and consuming sectors respectively. α_{ij} is the input from sector i to sector j and denotes the raw materials consumed by j from i in the production process. We first construct γ_{ij} to measure the relative importance of input i in production of j.

$$\gamma_{ij} = \frac{\alpha_{ij}}{\sum_i \alpha_{ij}} \tag{4}$$

Using total war orders received by industry j (V_j), we construct the upstream and downstream linkages per worker in industry i as

$$Upstream_i = \frac{\sum_j \gamma_{ij} \times V_j}{\textit{Total workers in industry } i \textit{ in 1931}} \tag{5}$$

$$Downstream_i = \frac{\sum_j \gamma_{ji} \times V_j}{\textit{Total workers in industry } i \textit{ in 1931}} \tag{6}$$

$Upstream_i$ represents "upstream" linkages, which are effects on (upstream) industry i resulting from war procurement in (downstream) industries j that demand inputs from industry i. $Downstream_i$ represents "downstream" linkages: effects on (downstream) industry i resulting from war procurement in (upstream) industries j that supply inputs to industry i.

We estimate the following regression on a national-level panel of industries (two observations per industry, for 1931 and 2011):

$$y_{it} = \theta_i + \gamma_t + \mu(S_i \times Post_t) + \beta(Upstream_i \times Post_t) + \\ \delta(Downstream_i \times Post_t) + \boldsymbol{\tau(y_{i,pre}} \times Post_t) + \epsilon_{it} \quad (7)$$

y_{it} is log workers in industry i in year t. θ_i are industry fixed effects and γ_t are period fixed effects. S_i is orders per worker in industry i (the "shift" in the shift-share variable, expression (1) above).

$Upstream_i$ and $Downstream_i$ are as defined in equations (5) and (6). S_i, $Upstream_i$ and $Downstream_i$ are interacted with the indicator for the post-war period $Post_t$. $y_{i,pre}$ refers to pre-WWII (1931) log workers in industry i, making this an ANCOVA specification. ϵ_{it} is the mean zero error term.

Table 7: Shift, upstream and downstream linkages (1931-2011)

VARIABLES	(1) Log Workers	(2) Log Workers
Shift × Post	0.361*	
	(0.208)	
Upstream × Post	0.017	
	(0.307)	
Downstream × Post	0.305	
	(0.302)	
Log Shift × Post		0.057*
		(0.029)
Log Upstream × Post		-0.020
		(0.040)
Log Downstream × Post		0.089
		(0.100)
industry FE	YES	YES
Year FE	YES	YES
No. of industries=	72	72
Obs=	144	144

Notes: Coefficient estimates from equation (7). Dependent variable is log workers in industry i. S_i are wartime procurement per worker in industry i. Upstream and downstream linkages per worker are as defined in expressions (5) and (6).

Regression results from estimating equation (7) are shown in Table 7. In column 1, we show results for a regression where we normalize S_i, $Upstream_i$, and $Downstream_i$ to have mean zero and unit standard deviation, so that each coefficient is interpreted as the impact of a one-standard-deviation increase. In column 2 we show regression results where the right-hand-side variables of interest are logged (non-normalized) S_i, $Upstream_i$, and $Downstream_i$, making this a log-log specification.

In both regressions, we find positive direct effects of war orders on industry size. The coefficient on S_i in column 1 and on log S_i in column 2 are each positive and statistically significantly different from zero (each at the 10% level). Column 1's estimate indicates that one-standard-deviation higher war orders per capita leads industry employment to be larger in 2011 by 36 log points. The estimate in Column 2 (the log-log specification) indicates that 10% higher orders per worker in industry i leads to 5.7% more workers in that industry in 2011.

Coefficients on the upstream and downstream variables are not statistically significantly different from zero in either regression. Coefficients on the upstream variables are small in magnitude, while those on the downstream variables are positive and larger, but imprecise. All told, we find no strong evidence here of upstream and downstream linkages between industrial sectors. That said, the coefficients are somewhat imprecise: 95% confidence intervals admit relatively large effects, particularly of downstream linkages.

4 Conclusion

Mobilization for war is one of the most prominent and costly activities undertaken by governments. Public decisions to mobilize for war must take into account a wide range of considerations – in, for example, the ethical, political, and social realms. This research deepens our understanding of the *economic* consequences of mobilizing for war. We shed light on these issues in an important context: the world's most populous country, India. That India is a developing country is also important, as there is very little empirical research on the consequences of war mobilization in developing countries. Our analysis also reveals very long-run impacts, over several decades. Our findings revealing the long-run economic impacts of war mobilization in a developing country can help guide debates and decision-making about participation in war in developing countries around the world.

Economic policies to mobilize for war have substantial overlap with "industrial policies" undertaken by governments, in that they both seek to shape the industrial composition and output of an economy. Our study therefore also contributes to our understanding of the long-run impacts of a type of industrial policy on economic development: in particular, policies that seek to promote the development of industrial sectors. Policy-makers should take account of our findings that temporary policies that seek to promote industrial sectors in the short run can have quite lasting impacts in the long run, persistently altering the industrial structure of the economy.

There are several important avenues for future research. It is of interest to examine a set of other district-level outcomes beyond the ones we have examined, for example related to firm productivity and investments, as well as international trade. In addition, it should be interesting to investigate whether post-independence economic policies may have responded to World War II economic mobilization, either magnifying or attenuating the impacts of the World War II procurement in subsequent decades. Future work could also conduct analogous analyses of war procurement in different historical and geographical contexts, to gauge the broader generalizability of these results.

References

Acemoglu, D. and F. Zilibotti (1997): "Was Prometheus Unbound by Chance? Risk, Diversification, and Growth," *Journal of Political Economy*, 105, 709–751.

Aggarwal, S. C. (1947): *History of the Supply Department (1939-1946)*, Manager of Publications.

Aghion, P., J. Cai, M. Dewatripont, L. Du, A. Harrison, and P. Legros (2015): "Industrial Policy and Competition," *American Economic Journal: Macroeconomics*, 7, 1–32.

Aguilar, J. A., R. Akee, and E. Mykerezi (2024): "Place-Based Economic Development and Long-Run Firm Employment and Sales: Evidence from American Indian Reservations," Working paper, National Bureau of Economic Research.

Ahsan, R. N. and D. Mitra (2016): "Can the Whole Actually be Greater than the Sum of its Parts? Lessons from India's Growing Economy and its Evolving Structure," in *Structural Change, Fundamentals, and Growth: A Framework and Case Studies*, ed. by M. McMillan, D. Rodrik, and C. Sepúlveda, International Food Policy Research Institute, 39–80.

Alder, S., L. Shao, and F. Zilibotti (2016): "Economic Reforms and Industrial Policy in a Panel of Chinese Cities," *Journal of Economic Growth*, 21, 305–349.

Amodio, F., E. Benveniste, H. Pham, and M. Sanfilippo (2024): "The Local (Informal) Multiplier of Industrial Jobs," Working paper, N/A.

Amsden, A. H. (1989): *Asia's Next Giant: South Korea and Late Industrialization*, Oxford University Press, USA.

Arrow, K. J. (1962): "The Economic Implications of Learning by Doing," *The Review of Economic Studies*, 29, 155–173.

Baldwin, R. E. (1969): "The Case against Infant-industry Tariff Protection," *Journal of Political Economy*, 77, 295–305.

Banerjee, A. and L. Iyer (2005): "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India," *American Economic Review*, 95, 1190–1213.

Banerjee, A., L. Iyer, and R. Somanathan (2005): "History, social divisions, and public goods in rural India," *Journal of the European Economic Association*, 3, 639–647.

Bartelme, D., A. Costinot, D. J. Donaldson, and A. Rodriguez-Clare (2024): "The Textbook Case for Industrial Policy: Theory Meets Data," *Journal of Political Economy*.

Barwick, P. J., M. Kalouptsidi, and N. B. Zahur (2019): "Chinaâs Industrial Policy: An Empirical Evaluation," Working paper, National Bureau of Economic Research.

Barwick, P. J., H.-s. Kwon, S. Li, Y. Wang, and N. B. Zahur (2024): "Industrial Policies and Innovation: Evidence from the Global Automobile Industry," Working Paper.

Bazzi, S. (2017): "Wealth Heterogeneity and the Income Elasticity of Migration," *American Economic Journal: Applied Economics*, 9, 219–55.

Beason, R. and D. E. Weinstein (1996): "Growth, Economies of Scale, and Targeting in Japan (1955-1990)," *The Review of Economics and Statistics*, 286–295.

Bencivenga, V. R. and B. D. Smith (1991): "Financial Intermediation and Endogenous Growth," *The Review of Economic Studies*, 58, 195–209.

Bharadwaj, P. and J. Fenske (2012): "Partition, migration, and jute cultivation in India," *Journal of Development Studies*, 48, 1084–1107.

Bharadwaj, P., A. I. Khwaja, and A. Mian (2015): "Population exchange and its impact on literacy, occupation and gender–evidence from the partition of India," *International Migration*, 53, 90–106.

Bharadwaj, P. and R. A. Mirza (2019): "Displacement and Development: Long Term Impacts of Population Transfer in India," *Explorations in Economic History*, 73, 101273.

Bianchi, N. and M. Giorcelli (2023): "Reconstruction Aid, Public Infrastructure, and Economic Development: The Case of the Marshall Plan in Italy," *Journal of Economic History*, 83, 501–537.

Bo, S., C. Liu, and Y. Zhou (2023): "Military investment and the rise of industrial clusters: Evidence from China's self-strengthening movement," *Journal of Development Economics*, 161.

Bonfatti, R. and B. Brey (2023): "Trade disruption, Industrialisation, and the Setting Sun of British Colonial Rule in India," *Journal of the European Economic Association*.

Borusyak, K., P. Hull, and X. Jaravel (2022): "Quasi-experimental Shift-share Research Designs," *The Review of Economic Studies*, 89, 181–213.

Bose, G., T. Jain, and S. Walker (2022): "Womenâs labor force participation and household technology adoption," *European Economic Review*, 147, 104181.

Buera, F. J., H. Hopenhayn, Y. Shin, and N. Trachter (2021): "Big Push in Distorted Economies," Tech. rep., National Bureau of Economic Research.

Castelló-Climent, A., L. Chaudhary, and A. Mukhopadhyay (2018): "Higher Education and Prosperity: From Catholic Missionaries to Luminosity in India," *The Economic Journal*, 128, 3039–3075.

Chaudhary, L. and J. Fenske (2022): "Did Railways Affect Literacy? Evidence from India," Working paper, University of Warwick Department of Economics.

Choi, J. and A. A. Levchenko (2021): "The Long-term Effects of Industrial Policy," Working paper, National Bureau of Economic Research.

Comin, D., D. Lashkari, and M. Mestieri (2021): "Structural Change with Long-Run Income and Price Effects," *Econometrica*, 89, 311–374.

Cox, L. (2023): "The Long-term Impact of Steel Tariffs on US Manufacturing," Working paper.

Criscuolo, C., R. Martin, H. G. Overman, and J. Van Reenen (2019): "Some Causal Effects of an Industrial Policy," *American Economic Review*, 109, 48–85.

Dell, M. and B. Olken (2020): "The Development Effects of the Extractive Colonial Economy," *Review of Economic Studies*, 87, 164–203.

Dincecco, M., J. Fenske, A. Menon, and S. Mukherjee (2022): "Pre-colonial Warfare and Long-run Development in India," *The Economic Journal*, 132, 981–1010.

Donaldson, D. (2018): "Railroads of the Raj: Estimating the impact of transportation infrastructure," *American Economic Review*, 108, 899–934.

Evans, P. B. (1995): *Embedded Autonomy: States and Industrial Transformation*, Princeton University Press.

Fan, J. and B. Zou (2021): "Industrialization from scratch: The "Construction of Third Front"' and local economic development in China's hinterland," *Journal of Development Economics*, 152, 102698.

Fan, T., M. Peters, and F. Zilibotti (2023): "Growing Like Indiaâthe Unequal Effects of Service-Led Growth," *Econometrica*, 91, 1457–1494.

Fenske, J., B. Gupta, and C. Neumann (2022): "Missing women in colonial india," .

Fick, S. E. and R. J. Hijmans (2017): "WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas," *International journal of climatology*, 37, 4302–4315.

Field, A. J. (2008): "The Impact of the Second World War on US Productivity Growth," *The Economic History Review*, 61, 672–694.

——— (2022): *The Economic Consequences of US Mobilization for the Second World War*, Yale University Press.

Fishback, P. V. and J. A. Cullen (2013): “Second World War Spending and Local Economic Activity in US Counties, 1939â58,” *Economic History Review*, 66, 975–992.

Galor, O. and Ö. Özak (2016): “The agricultural origins of time preference,” *American economic review*, 106, 3064–3103.

Galor, O. and J. Zeira (1993): “Income Distribution and Macroeconomics,” *The Review of Economic Studies*, 60, 35–52.

Garin, A. and J. Rothbaum (2022): “The Long-Run Impacts of Public Industrial Investment on Regional Development and Economic Mobility: Evidence from World War II,” *NBER Working Paper*.

Giorcelli, M. and B. Li (2021): “Technology Transfer and Early Industrial Development: Evidence from the Sino-Soviet Alliance,” Working paper, National Bureau of Economic Research.

Goldberg, P., R. Juhasz, N. Lane, G. Lo Forte, and J. Thurk (2024): “Industrial Policy in the Global Semiconductor Sector,” Working paper, National Bureau of Economic Research.

Goldberg, P. and T. Reed (2023): “Presidential Address: Demand-Side Constraints in Development: The Role of Market Size, Trade, and Inequality,” *Econometrica*, 91, 1915–1950.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020): “Bartik Instruments: What, When, Why, and How,” *American Economic Review*, 110, 2586–2624.

Gordon, R. (2017): *The Rise and Fall of American Growth: The US Standard of Living since the Civil War*, Princeton University Press.

Gordon, R. J. (1967): “Problems in the Measurement of Real Investment in the US Economy,” *Ph.D. Dissertation, Department of Economics, MIT.*

Greenstone, M., R. Hornbeck, and E. Moretti (2010): “Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings,” *Journal of Political Economy*, 118, 536–598.

Greenwood, J. and B. Jovanovic (1990): “Financial Development, Growth, and the Distribution of Income,” *Journal of Political Economy*, 98, 1076–1107.

Gross, D. P. and B. N. Sampat (2023): “America, Jump-Started: World War II RD and the Takeoff of the US Innovation System,” *American Economic Review*, 113, 3323–3356.

Gupta, B., D. Ma, and T. Roy (2016): “States and Development: Early Modern India, China, and the Great Divergence,” *Economic History of Warfare and State Formation*, 51–69.

Hanlon, W. W. (2020): “The Persistent Effect of Temporary Input Cost Advantages in Shipbuilding, 1850 to 1911,” *Journal of the European Economic Association*, 18, 3173–3209.

Harrison, A. and A. Rodríguez-Clare (2010): “Trade, Foreign investment, and Industrial Policy for Developing Countries,” *Handbook of Development Economics*, 5, 4039–4214.

Hasan, R., D. Mitra, and A. Sundaram (2013): "What explains the high capital intensity of Indian manufacturing," *Indian Growth and Development Review*, 6, 212–241.

Hausmann, R. and D. Rodrik (2003): "Economic Development as Self-Discovery," *Journal of Development Economics*, 603–633.

Head, K. (1994): "Infant Industry Protection in the Steel Rail Industry," *Journal of International Economics*, 37, 141–165.

Heblich, S., M. Seror, H. Xu, and Y. Zylberberg (2022): "Industrial Clusters in the Long Run: Evidence from Million-Rouble Plants in China," Working paper, National Bureau of Economic Research.

Higgs, R. (2004): "Wartime Socialization of Investment: a Reassessment of US Capital Formation in the 1940s," *The Journal of Economic History*, 64, 500–520.

Hirschman, A. O. (1961): *The Strategy of Economic Development*, Yale University Press.

Irwin, D. A. (2000): "Did Late-nineteenth-century US Tariffs Promote Infant Industries? Evidence from the Tinplate Industry," *Journal of Economic History*, 60, 335–360.

Iyer, L. (2010): "Direct versus Indirect Colonial Rule in India: Long-term Consequences," *The Review of Economics and Statistics*, 92, 693–713.

Jaworski, T. (2017): "World War II and the Industrialization of the American South," *Journal of Economic History*, 77, 1048â–82.

Jha, S. and S. Wilkinson (2012): "Does Combat Experience Foster Organizational Skill? Evidence from Ethnic Cleansing during the Partition of South Asia," *American Political Science Review*, 106, 883–907.

Juhász, R. (2018): "Temporary Protection and Technology Adoption: Evidence from the Napoleonic Blockade," *American Economic Review*, 108, 3339–3376.

Kalouptsidi, M. (2018): "Detection and Impact of Industrial Subsidies: The Case of Chinese Shipbuilding," *The Review of Economic Studies*, 85, 1111–1158.

Kamtekar, I. (2002): "A Different War Dance: State and Class in India 1939-1945," *Past & Present*, 187–221.

Khanna, G., E. Murathanoglu, C. Theoharides, and D. Yang (2022): "Abundance from Abroad: Migrant Income and Long-Run Economic Development," *NBER Working Paper No. 29862*.

Kim, M., M. Lee, and Y. Shin (2021): "The Plant-level View of an Industrial Policy: The Korean Heavy Industry Drive of 1973," Working paper, National Bureau of Economic Research.

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, 275–331.

Krueger, A. O. (1990): "Government Failures in Development," *Journal of Economic Perspectives*, 4, 9–23.

Lane, N. (forthcoming): "Manufacturing Revolutions: Industrial Policy and Industrialization in South Korea," *Quarterly Journal of Economics*.

Lederman, D. and W. Maloney (2012): *Does what you Export matter?: In Search of Empirical guidance for industrial policies*, World Bank Publications.

Lee, A. (2019): "Land, State Capacity, and Colonialism: Evidence from India," *Comparative Political Studies*, 52, 412–444.

Lee, J.-W. (1996): "Government Interventions and Productivity Growth," *Journal of Economic Growth*, 1, 391–414.

Liu, E. (2019): "Industrial Policies in Production Networks," *The Quarterly Journal of Economics*, 134, 1883–1948.

Mahajan, P. and D. Yang (2020): "Taken by Storm: Hurricanes, Migrant Networks, and U.S. Immigration," *American Economic Journal: Applied Economics*.

Manelici, I. and S. Pantea (2021): "Industrial Policy at Work: Evidence from Romaniaâs Income Tax Break for Workers in IT," *European Economic Review*, 133, 103674.

McNeill, W. H. (1982): *The Pursuit of Power: Technology, Armed Force, and society since AD 1000*, University of Chicago Press.

Mitrunen, M. (forthcoming): "War Reparations, Structural Change, and Intergenerational Mobility," *Quarterly Journal of Economics*.

Moretti, E., C. Steinwender, and J. Van Reenen (2021): "The Intellectual Spoils of War? Defense R&D, Productivity and International Spillovers," *NBER Working Paper*, 26483.

Morris, M. D. (1983): "The Growth of Large-scale Industry to 1947," *The Cambridge Economic History of India*, 2, 553–676.

Moser, P. and A. Voena (2012): "Compulsory Licensing: Evidence from the Trading with the Enemy Act," *American Economic Review*, 102, 396–427.

Murphy, K. M., A. Shleifer, and R. W. Vishny (1989): "Industrialization and the Big Push," *Journal of Political Economy*, 97, 1003–1026.

Naya, S. (1971): "The Vietnam War and some Aspects of its Economic Impact on Asian Countries," *The Developing Economies*, 9, 31–57.

Nunn, N. and D. Trefler (2010): "The Structure of Tariffs and Long-term Growth," *American Economic Journal: Macroeconomics*, 2, 158–194.

Nurkse, R. (1953): *Problems of Capital Formation in Underdeveloped Countries*, B. Blackwell.

Pack, H. (2000): "Industrial Policy: Growth Elixir or Poison?" *The World Bank Research Observer*, 15, 47–67.

Raghavan, S. (2017): *Indiaâs War: The Making of Modern South Asia 1939-1945*, Penguin UK.

Rasler, K. and W. R. Thompson (1985): "War and the Economic Growth of Major Powers," *American Journal of Political Science*, 513–538.

Rhode, P. W. (2003): *After the War Boom: Reconversion on the Pacific Coast, 1943â1949*, Stanford University Press.

Ristuccia, C. A. and A. Tooze (2013): "Machine Tools and Mass Production in the Armaments Boom: Germany and the United States, 1929–44," *The Economic History Review*, 66, 953–974.

Rockoff, H. (2012): *America's Economic Way of War: War and the U.S. Economy from the Spanish-American War to the Persian Gulf War*, Cambridge University Press.

Rodrik, D. (1995): "Getting Interventions Right: how South Korea and Taiwan Grew Rich," *Economic Policy*, 10, 53–107.

——— (2007): *Industrial Policy for the Twenty-first Century*, Princeton University Press.

Rose, E. K. (2018): "The Rise and Fall of Female Labor Force Participation During World War II in the United States," *Journal of Economic History*, 78, 673–711.

Rosenstein-Rodan, P. (1943): "Problems of Industrialisation of Eastern and South-eastern Europe," *Economic Journal*, 202–211.

Rotemberg, M. (2019): "Equilibrium Effects of Firm Subsidies," *American Economic Review*, 109, 3475–3513.

Roy, K. (2016): *India and World War II: War, Armed Forces, and Society, 1939–45*.

Ruttan, V. W. (2006): *Is War Necessary for Economic Growth?: Military Procurement and Technology Development*, Oxford University Press.

Saluja, M. (1972): "Structure of Indian Economy: 1964-65 Input-output Relations among 144 Sectors," *Sankhyā: The Indian Journal of Statistics, Series B*, 433–462.

Sinha, N. C. and P. N. Khera (1962): *Indian War Economy: Supply, Industry & Finance*.

Sivasubramonian, S. (2000): *The national income of India in the twentieth century*.

Steiner, Z. (2011): *The Triumph of the Dark: European International History 1933-1939*, Oxford History of Modern Europe, Oxford: Oxford University Press.

Stubbs, R. (1999): "War and Economic Development: Export-oriented Industrialization in East and Southeast Asia," *Comparative Politics*, 337–355.

Tomlinson, B. R. (1996): *The Economy of Modern India, 1860-1970*, vol. 3, Cambridge University Press.

Townsend, R. M. (1979): "Optimal Contracts and Competitive Markets with Costly State Verification," *Journal of Economic Theory*, 21, 265–293.

Wade, R. (1990): *Governing the Market: Economic Theory and the Role of Government in East Asian Industrialization*, Princeton University Press.

Weinstein, D. E. (1995): "Evaluating Administrative Guidance and Cartels in Japan (1957-1988)," *Journal of the Japanese and International Economies*, 9, 200–223.

Online Appendix for

War Mobilization and Economic Development: World War II and Structural Transformation in India

Aneesha Parvathaneni
University of Michigan

Dean Yang
University of Michigan

December 1, 2024

Contents

A Appendix Tables

Table A1: Top Specific Products

Product	Product Category	Value of Orders (INR, thousands)
Cotton Tent Components (Flies, Walls, etc.)	Cotton Textiles	63,382,578
Ghee	Food Items	40,767,138
Timber - General Purpose	Timber	32,590,300
Vehicles and Chassis	Vehicles	30,918,373
Sand Bags	Woollen Textiles	26,502,801
Atta	Food Items	19,551,383
Batteries & Cells, various sizes and types	Engineering Stores	17,830,618
Rice	Food Items	16,179,445
Boots Ankle	Leather	16,083,376
Vehicles and lorry bodies	Vehicles	15,914,896
Hessian Cloth	Woollen Textiles	12,361,841
Cement	Engineering Stores	12,305,738
Socks Woollen	Woollen Textiles	12,286,220
Sugar	Food Items	10,752,845
E.D. Parachute Cotton Com-18, diameter	Clothing	9,874,211
Blankets Bawaok	Woollen Textiles	9,440,144
Serim Hessian Strips	Woollen Textiles	8,654,438
Bitumen	Engineering Stores	8,630,717
Web Equipment	Harness and Saddlery Items	8,155,310
Snap Fasteners	Miscellaneous	7,293,739

Notes: Data are from Aggarwal (1947). Orders by product are summed across 1939-1946, expressed in thousands of real (2011) Indian rupees (INR) (Sivasubramonian 2000).

Table A2: Falsification Tests

	Dependent Variable: War procurement per worker			
	(1)	(2)	(3)	(4)
Panel A				
Share of national employment	–1.626			–1.366
	(1.216)			(1.267)
Heavy Industry		0.094		0.067
		(0.084)		(0.090)
Geographic Concentration			0.256	0.101
			(0.561)	(0.576)
	75	75	75	75
Joint F-Test P-value				.49
Panel B				
Share of national employment	–1.467			–1.230
	(1.101)			(1.144)
Heavy Industry		0.094		0.066
		(0.084)		(0.090)
Geographic Concentration			0.250	0.118
			(0.424)	(0.439)
	75	75	75	75
Joint F-Test P-value				.48

Notes: Dependent variables are share of total employment in industry i, number of workers in industry i, binary indicator for whether an industry is heavy industry and geographic concentration of industry i in 1931. Geographic concentration is the Herfindahl-Hirschman index of share of total workers in industry i employed in district d. A high HHI for industry i indicates that it is very geographically concentrated in a certain district. The relevant districts here are geographic units that stay consistent between 1931 and 2011. Panel A shows the results for industry level falsification tests for a sample of workers in different industries spanning modern day India. These regressions are weighted by share of workers in industry i (normalized to sum up to 1). Panel B shows these industry level results for a sample that parallels our district-level regression sample excluding Rajputana agency, North east, Bengal and certain princely states in Central India Agency.

Table A3: 1931 vs. Different Post-War Decades

	(1)	(2)	(3)	(4)	(5)
Panel A: 1931 vs 1951					
Shiftshare $\times Post_t$	0.0821***	0.0799***	0.0723***	0.0720***	0.0717***
	(0.0141)	(0.0130)	(0.0143)	(0.0146)	(0.0143)
Num. Obs.	320	320	320	320	320
Panel B: 1931 vs 1961					
Shiftshare $\times Post_t$	0.0684***	0.0640***	0.0660***	0.0697***	0.0679***
	(0.0177)	(0.0166)	(0.0190)	(0.0164)	(0.0159)
Num. Obs.	310	310	310	310	310
Panel C: 1931 vs 1971					
Shiftshare $\times Post_t$	0.0906***	0.0851***	0.0806***	0.0786***	0.0769***
	(0.0198)	(0.0177)	(0.0181)	(0.0169)	(0.0167)
Num. Obs.	310	310	310	310	310
Panel D: 1931 vs 1981					
Shiftshare $\times Post_t$	0.0855***	0.0795***	0.0751***	0.0721***	0.0712***
	(0.0202)	(0.0180)	(0.0183)	(0.0155)	(0.0153)
Num. Obs.	316	316	316	316	316
Panel E: 1931 vs 1991					
Shiftshare $\times Post_t$	0.0834***	0.0761***	0.0743***	0.0688***	0.0688***
	(0.0228)	(0.0193)	(0.0188)	(0.0170)	(0.0171)
Num. Obs.	318	318	318	318	318
Panel F: 1931 vs 2001					
Shiftshare $\times Post_t$	0.0839***	0.0757***	0.0690***	0.0608***	0.0592***
	(0.0220)	(0.0185)	(0.0203)	(0.0172)	(0.0172)
Num. Obs.	320	320	320	320	320
Panel G: 1931 vs 2011					
Shiftshare $\times Post_t$	0.0818***	0.0696**	0.0708**	0.0636**	0.0652**
	(0.0317)	(0.0318)	(0.0337)	(0.0313)	(0.0325)
Num. Obs.	332	332	332	332	332
District F.E	YES	YES	YES	YES	YES
Year F.E	YES	YES	YES	YES	YES
1931 Economic Controls $\times Post_t$	YES	YES	YES	YES	YES
Historic Controls $\times Post_t$	NO	YES	YES	YES	YES
Region FE $\times Post_t$	NO	NO	YES	YES	YES
Geographic Controls $\times Post_t$	NO	NO	NO	YES	YES
Military Controls $\times Post_t$	NO	NO	NO	NO	YES

Notes: **Dependent variable** is employment in modern sectors (industry and services) as share of total employment. All regressions include district and year fixed effects. Controls interacted with *Post* are all from pre-WWII period or time-invariant. **Economic controls** (from 1931 Census) are log population, share population employed, industrial workers as share of employment, service workers as share of employment, heavy industrial workers as share of employment, population density, years of prior railroad access, and shift-share "sum of shares" (share of workers in any war-related industry). **Historical controls** (from Dincecco et al. (2022)) are share of population under British direct rule and historical conflict within 250 km (years 1000-1757). **Region fixed effects** are for 11 regions. **Geographic controls** are mean temperature, mean precipitation, mean slope, mean elevation, land area, and maximum caloric yield in agriculture. **Military controls** (from Jha and Wilkinson (2012)) are WWII casualties per million, martial castes per thousand, and indicator for non-missing military controls. Standard errors are exposure-robust, accounting for correlation of shocks across districts, based on estimation of shock-level (industry-level) regressions (Borusyak et al., 2022). Significance levels: * 10%, ** 5%, *** 1%.

B Data Appendix

In this Appendix section, we provide details about data sources and data processing procedures we have used to create the dataset for this paper's analyses. We also provide information on data locations to facilitate data access and replication of analyses. We refer below to files in a replication package. Upon publication of the paper, we will make the replication package available on our academic webpages.

B.1 Census Data

We assembled occupation and population data for Indian Census years spanning 1911-2011. For the years 1911-1981, the district level census data were extracted from the state level PDFs. For the years 1991-2011, the district level census data are digitally available and can be downloaded from the Indian Census website.[17]

Unless otherwise specified, we define **agriculture** workers as those working as agricultural laborers, cultivators, and those engaged in livestock, forestry, fishing, hunting and plantation. We define **industry** workers (often referred to in the Indian Census as "production" workers) as those engaged in mining and quarrying, manufacturing, processing and construction. We define **service** workers as those engaged in trade, transport, communications, public and administration services and other miscellaneous services.

For each of the Census years 1911-1951, we also remove "unproductive" workers (e.g., inmates of prisons, beggars, pensioners) from services, since in those Census years they are listed in the service sector. For pre-independence census years (1911-1931), three large residual industrial occupation groups ((i) Manufacturers, businessmen and contractors otherwise unspecified, (ii) mechanics otherwise unspecified and (iii) labourers and workmen otherwise unspecified) are classified by the Census as a part of "miscellaneous services". For these census years, we consider workers in these three groups to be part of the industrial sector, not the service sector.

We define employment as the sum of workers in the agriculture, industry, and service sectors. We define the "modern sector" as the sum of workers in the industry and service sectors. In the rest of this section, we provide information on the source (Census tables) of population data, occupation data, and data on unproductive workers (where relevant) for different Census years, and describe how we define **workers** in each year.

We contracted with Paradigm Data Services of Chandigarh for most of the data entry required to convert Census PDFs to data files. Prashant Bharadwaj kindly shared data with us from his team's Census data entry, which provided some of the data for 1921-1951. We also appreciate the help of Roberto Bonfatti, who shared some district level data on employment for the 1911 Census from Bonfatti and Brey (2023).

[17] https://censusindia.gov.in/census.website/

B.1.1 1911

The occupational data for the 1911 census was extracted from Table XV (Occupation or Means of Livelihood) from each of the Indian provinces, states and agencies. For the census year 1911, we define workers to be “actual workers” (“Total Workers and Dependents” minus “Dependents”). We define workers in sub-class IX (Persons living on their own income) and sub-class XII (Unproductive) from this table to be unproductive workers. The data for population is extracted from Table I (Area, Houses and Population) from each of the province censuses.

B.1.2 1921

The occupational data for the 1921 census was extracted from Table XVII (Occupation or Means of Livelihood, Part B - General Table – Details for Districts) from each of the provincial census PDFs. The data on population is from Table I (Area, Houses and Population). We define persons principally living on their own income (order 51), inmates of jails, asylums and almshouses (order 54), beggars, vagrants, prostitutes (order 55) and other unclassified non-productive industries (order 56) from table XVII as unproductive workers and remove them from the employment figures. We define workers to be “Actual workers” in a manner similar to 1911.

B.1.3 1931

The occupational data for the 1931 census was extracted from Table X (Occupation or Means of Livelihood) from each of the Indian provinces, states and agencies. For the census year 1931, we define “workers” as the sum of principal earners and earning dependents. The 1931 census report describes the sum of principal earners and earning dependents in effect to be equivalent to the workers as recorded by 1921 census. We define workers in occupation groups 185 (Persons living principally on their income), 192 (Inmates of jails, asylums and alms houses), 193 (Beggars and vagrants) and 195 (Other unclassified non-productive industries) as unproductive workers and remove them from employment figures. We consider those working as procurers and prostitutes (group 194) as employed. (In the previous two censuses, this occupation was not reported separately.)

B.1.4 1951

For the census year 1951, we extract data on livelihood classes and population from Economic Table B-I (Livelihood Classes and Sub-classes) and data on unproductive workers from the table in the preamble preceding Economic Table III for each of the Indian states. The introduction before Table BIII provides district-level information on various types of unproductive workers. In these tables, the livelihood classes were divided into three sub-classes, namely self-supporting persons, non-earning dependents and earning dependents. For the year 1951, we define workers as the sum of self-supporting persons and earning dependents. For this year, data in table B-I is reported for eight livelihood classes (e.g., Agricultural Class I: cultivators of land wholly owned; Agricultural Class II: cultivators of land wholly or mainly unowned; etc.) "Primary industries" (division 0) such as fishing, hunting etc are classified by the census as included as a part of the industry sector, so we re-assign these workers to the agriculture sector. We extract district level data on division 0 from Table B-III (Employers, Employees and Independent Workers in Industries and Services by Division and Sub-divisions).[18] For the year 1951, data on Jammu and Kashmir districts are not reported and are hence not a part of our analysis for the year 1951.

B.1.5 1961

For the year 1961, we extract data on workers and population from Table B-1 (Workers and Non-workers Classified by Sex and Age Group.) for each of the Indian states apart from Madras and North-Eastern Frontier Agency. For these two states, we were not able to locate PDFs containing the B-I economic table. Therefore, for these two states, we extracted information from Union Primary Census Abstracts. In the 1961 census, workers and non-workers as aggregate categories are reported but no distinction is made between principal earners and working dependents. We use the definition of "workers" in the 1961 census. From 1961 census onwards, non-workers are reported separately. No further adjustments are required to remove unproductive workers.

Similarly to 1951, primary industries like fishing, hunting and so on are reported along with mining, quarrying and hunting. We re-assign the workers in fishing, hunting, and the like (referred to in the census as Division 0) from industry to agriculture. We extract the data on division 0 workers from table B-IV Part C in state censuses. (Industrial Classification by Sex and Class of Worker of Persons at Work in Non-household Industry, Trade, Business, Profession or Services). For Madras, we extract the data for workers in division 0 from Table B-V (Occupational Classification by Sex of Persons at Work Other than Cultivation).

B.1.6 1971

For the year 1971, we extract data on population and workers from Table B-I Part A (Workers and Non-workers According to Main Activity Classified by Sex and Age Groups) for each of the Indian states (except Mysore). For Mysore, we were not able to locate the General Economic Tables PDF, and hence extract data from the primary census abstract. The data on primary industries is reported on its own from the 1971 census onwards and no further adjustments are necessary.

[18] Table B-III providesinformation on self-supporting persons only, and no data on earning dependents.

B.1.7 1981

From 1981 census and onwards data on two types of workers are reported, namely main workers and marginal workers. Main workers are broadly defined as those who work for over 6 months in a year and marginal workers are defined as those who work for fewer than 6 months in a year. We define workers in a certain sector as the sum of both marginal and main workers engaged in that sector. For the 1981 census year, the data on population and main workers is extracted from table B3 (Main Workers Classified by Industrial Category, Age and Sex) and data on marginal workers is extracted from table B7 (Marginal Workers by Industrial Category, Age and Sex) for each of the Indian states (apart from Assam and Arunachal Pradesh.[19])

B.1.8 1991

From 1991 census year onwards, the data tables are available in electronic format on the Indian census website, and so did not require data entry. For 1991, we download state-wise tables B-2(S) (Main Workers Classified by Industrial Category, Age and Sex), B-2(F) (Main Workers Other Than Cultivators And Agricultural Labourers Classified By Industrial Category, Age And Sex) and B-6(F) (Marginal Workers Classified By Industrial Category Of Work And Main Activity) from the census website. B-2(S) tables provide information on population, total main workers, main workers in agriculture, cultivation, manufacturing in household industry and sum of main workers in other industrial categories. B-2(F) tables provide us more detailed information on main workers in other industrial categories like mining and quarrying, construction and so on. For the census year 1991, data on Jammu and Kashmir districts is not reported. We do not include Jammu and Kashmir in our analysis for the year 1991.

B.1.9 2001

For 2001, we download data tables B-04 (Main Workers Classified by Age, Industrial Category and Sex (All)) for data on main workers, B-06 (Marginal Workers Classified by Age, Industrial Category and Sex) for data on marginal workers and district-wise population data from the Indian census website for population data.

B.1.10 2011

For the census year 2011, we download tables B-04 (Main Workers Classified by Age, Industrial Category and Sex (Total)) for data on main workers and B-06 (Marginal Workers Classified by Age, Industrial Category and Sex (Total)) for data on marginal

[19] We were not able to locate the General Economic Tables for these states. In any case, these states are not included in our analyses due to their proximity to the Japanese advance in Burma and the subsequent orientation of war production away from the northeast and Bengal.

workers for each of the Indian states from the Indian census website. We also download table B-18 (Industrial Classification of Main Workers and Marginal Workers other than Cultivators and Agricultural Labourers by Sex and by Section, Division and Class) for detailed data on workers in employed in different occupations. We use this data to determine workers in various service sub-categories as well as get national level estimates of workers engaged in different occupations.

B.2 World War II procurement by industry

Data on World War II procurement is taken from the table "Major Items of Supply Expenditure" in Appendix III of Aggarwal (1947). This table provides information on value of supply orders for 385 product categories. To calculate the real value of the supply orders, we convert the nominal value of orders to real 2011 INR terms using data from Sivasubramonian (2000).

B.2.1 Conversion of nominal value of war procurement to 2011 terms

We convert the nominal value of World War II procurement between 1939-1946 to 2011 INR terms. There is no single series that provides data on GDP deflators for the period spanning 1939-2011 for India. We proceed in three steps, drawing data from multiple sources to convert World War II procurement to 2011 INR real terms. Appendix table 6(g) in Sivasubramonian (2000) provides GDP deflators for 1935-1949 (with 1948-1949 as base year). Appendix table 9(b) in Sivasubramonian (2000) provides GDP deflators for 1946-2000 (with 1948-1949 as base year). The estimates from table 6(b) are for pre-independence India (which include modern day Pakistan and Bangladesh) and estimates from table 9(b) are for modern-day India. Both of these tables also report GDP deflators for 1946-1947, with 1948-1949 as base year. Table 6(b) reports the GDP deflator for 1946-47 to be 73.6 and table 9(b) reports 76. We make the estimates across these two tables consistent by multiplying the GDP deflator for pre-independent India by 1.036 (76/73.6). The GDP deflator after this adjustment allows us to convert nominal war-related orders received during the war years to 2000 INR. Finally, to convert these to 2011 INR, we use the GDP deflator for 2001 with 2011 as the base year from World Bank data.[20] Using these GDP deflators, in the final step we calculate the (real) 2011 INR value of World War II procurement in India.

B.2.2 Mapping supply orders to Census occupations

We mapped each product listed in Appendix III of Aggarwal (1947) to the 1931 Census occupations based on the product description and 1931 Census occupation description. Out of 385 products, we are able map 120 products one-to-one with 1931

[20] https://data.worldbank.org/indicator/NY.GDP.DEFL.ZS.AD?end=2011&locations=US-IN&name_desc=false&start=2000

Census occupations. In the rest of the cases we map the product to more than one census occupation. In such cases, we assign the INR amount of orders received to the Census occupation in proportion to the relative size of the occupation. We define the size of the occupation as the number of workers in that occupation across all of India excluding Burma, Bengal, Andaman and Nicobar and Assam.[21] For example, we map the product "25 Pounder Cartridge" to 1931 occupations "Makers of arms, guns etc" (Census occupation group 58) and "Manufactures of matches, fireworks and other explosives" (Census occupation group 66). 28% of the value of orders of the "25 Pounder Cartridge" product is assigned to "Makers of arms, guns etc" and the rest to "Manufacturers of fireworks and other explosives", in proportion to their relative numbers of workers in 1931. We then take the sum of product orders allocated to each of the 1931 census occupations to predict the total orders received by that occupation during the war years. This generates the variable S_i, wartime procurement per worker in industry i (the "shifts" in the shift-share variable).

B.3 Occupational crosswalks and classifications

B.3.1 Occupation crosswalk across different census years

To enable occupational or district-level pre-trend analyses, we build an occupational concordance between the 1911, 1921 and 1931 Censuses. In addition, to study the effect of wartime procurement on upstream and downstream industries we build a concordance between 1931 and 2011 occupations.

1911, 1921 and 1931 occupational concordance: We concord the occupations across these three census years using the description of the occupations provided in the Indian Census. There is not much change in occupational descriptions across these three censuses. Data are reported for a total of 195 census occupations in 1931, 191 in 1921 and 169 in 1911. For example, we map "Fruit, flower, vegetable, betel etc growers" in 1911 and 1921 to "Market gardners, flower and fruit growers" in 1931. Similarly, we map "Harbour works, dockyards and pilots" in 1911 to "Persons (other than labourers) employed in harbours and docks, including pilots" in 1921 and 1931. In a few cases, occupations reported in a more aggregate manner in 1911 and 1921 are reported with more detail in 1931 (and vice versa). For example, the census reports the total workers in tea, coffee, cinchona, rubber and indigo plantations for 1921 and 1911. In the 1931 census, the plantation workers in cinchona, coconut, coffee, ganja, rubber and tea are reported separately. In such cases, we aggregate the workers across these occupations to match with the occupation group in 1921 and 1911.[22]

1931 and 2011 occupational concordance: We also concord the occupations in 1931 and 2011. We use occupation descriptions provided in the 1931 census and

[21] We exclude workers in Burma as it was not a part of India during World War II. We exclude workers in Bengal and Assam as war-related orders were not sent here due to the threat of occupation by the Japanese army via Burma. We exclude Andaman and Nicobar Islands as this was under Japanese occupation during World War II.

[22] We provide the exact crosswalks in our replication package (*OccupationMapping1921-1931.xlsx* and *OccupationMapping1931-2011.xlsx*).

detailed occupational descriptions from the 2008 National Industrial Classifications (NIC) for India.[23] Since occupations changed substantially between 1931 and 2011, depending on feasibility we map 1931 occupations with either 2-digit broader occupations or more specific 3 or 4-digit occupations. For example we map 1931 occupation "Bank managers, money lenders, exchange and insurance agents, money changers and brokers and their employees" (occupation code 115) with three 2-digit NIC occupations in 2011, namely "Financial service activities, except insurance and pension funding" (occupation code 64), "Insurance, reinsurance and pension funding, except compulsory social security" (occupation code 65), and "Other financial activities" (occupation code 66). We map "Salt, saltpetre, and other saline substance" (occupation code 40) in 1931 with "Extraction of salt" (occupation code 0893) in 2011. This serves as an example of where we are able to map a 1931 occupation to a more specific 4-digit occupation code in 2011.[24]

B.4 Heavy vs Light Industrial Classification

We separately examine the effect of a heavy industry shift-share variable and light industry shift-share variable. In addition, in our regression, we also include a control for share of employment in heavy industries. In order to do this we classify the occupations in industry into heavy and light industrial occupations. Where possible we refer to capital intensities for Indian industries provided in Hasan et al. (2013). We classify occupations broadly fitting into industries (as mentioned in Hasan et al. (2013)) with capital intensity less than 10 as being light industries, and classify the rest as heavy industries. Some examples of heavy industries are "Smelting, forging and rolling of iron and other metals", "makers of arms, guns etc", "manufacturing of mineral oils" and "ship, boat, airplane builders". Examples of light industries include textiles (e.g. cotton ginning, cleaning and pressing), brewers and distillers, carpenters, and leather. We use the occupation mapping files mentioned in sub-section B.3.1 to assign 1921 and 1911 occupations to heavy and light industries.[25]

B.5 Services sub-sector classification

We also examine the effect of wartime procurement on five different sub-categories of services, namely consumer services, business services, health services, education services and government services. In 1931 and 2011, we assign all service occupations to one of these sub-categories. For 1931 we use the 1931 Census occupation descriptions and for 2011 we use the NIC occupation descriptions to determine who the services are aimed towards. Examples of service occupations classified as consumer related services in 1931 are private motor-car drivers and cleaners, general store keeps, hawk-

[23] In the 2011 Census, number of workers are reported for various 4-digit 2008 NIC occupations.

[24] For the exact crosswalk, please refer to *OccupationMapping1931-2011.xlsx* in our replication package.

[25] The exact assignment of 1931 census occupations to heavy and light industries is provided in file *capital intensity.xls* in our replication package.

ers of drink and food-stuffs and domestic servants. Examples of consumer related services in 2011 are food and beverage service activities, retail sale of various goods and hair-dressing and beauty treatment services. Examples of business related services in 1931 are cashiers, accountants, book-keepers, laborers employed on roads and bridges, lawyers and postal services. Examples of business related services in 2011 are wholesale sale of various goods, data processing activities, postal services and financial service activities.[26]

B.6 Input-Output (IO) Table - 1931 Census mapping

To study the effects of war-related orders on upstream and downstream industries, we use the Input-Output (IO) table from Saluja (1972). This provides us input-output linkages for 144 sectors for 1964-65. We map each of these 144 sectors to the 1931 census occupations based on the sector description in the IO table and the 1931 Census occupation description. In some cases, we are able to map an IO sector to one 1931 census occupation. For example, we map "Power Machinery except electric motors" from the IO table to "Heat, light, electricity, motive power" in 1931 Census occupations. Similarly, we map "Motor Cycles, Scooters and Bicycles" to "Persons engaged in making, assembling or repairing motor vehicles". In some other cases, we map an IO sector to more than one census occupation. For example, we map "Aircrafts" from the IO table to census occupations "ship, boat, aeroplane builders" and "persons connected with aerodomes and aeroplanes". We map different types of machinery like 'Chemical machinery, e.g.," to 6 census occupations, namely, (1) "smelting, forging, rolling of iron and other materials", (2) "makers of arms, guns etc", (3) "workers in brass, copper and bell metal", (4) "workers in other metals", (5) "workers in mints, die-sinks etc". and (6) "blacksmiths, other workers in iron, makers of implements". This mapping allows us to construct an IO table with 1931 census occupations as the consuming and producing sectors.[27]

B.7 District concordances

We use the crosswalk between modern Indian sub districts and historic districts constructed in Fenske et al. (2022) to form 'districts' that stay consistent between any two census decades.[28] We make minor changes to this crosswalk for the purposes of matching it with census as detailed below. Henceforth, we refer to this crosswalk as the "geographic boundaries concordance". For each census year, we map the districts as reported in census data with the districts reported in the geographic boundaries concordance. This allow us to assign economic data to geographically consistent districts over time.

[26] For more information on the exact assignment of these services, please refer to *ServicesSubcategories1931.xlsx* for 1931 mapping and *Services-sub-categories-complete-list.xlsx* for 2011 mapping in our replication package.

[27] For more details on the exact mapping, please refer to *Concordance_IOTable_census.xlsx* in our replication package.

[28] This crosswalk (gadmtohist_indiaonly.xlsx) is publicly available and can be downloaded from https://warwick.ac.uk/fac/soc/economics/staff/jefenske/data/.

Districts that were under the rule of non-British colonial powers (Portugal and France) are not present in any pre-independence censuses and hence do not match with the geographic concordance. Examples of such districts are Yanam, Pondicherry, Mahe and Karaikal, which were a part of French India before 1950. Other such districts are Goa, Dadra and Nagar Haveli and Daman and Diu, which were under Portuguese rule until 1961.

Districts that were a part of India before independence but are in modern day Pakistan or Bangladesh also drop out since the geographic boundaries concordance is only for modern day Indian districts. Similarly, we did not find census data for small princely states in Central India Agency for 1911 and 1921. As a result these are not included in our pre-trend sample. If the census data is missing for a district or a part of district, then it drops out of our sample and is depicted in grey in maps in Section C.

B.7.1 Modifications to geographic boundaries concordance

We rename districts that appear in different states to have an abbreviation of the state as a suffix. For example, if two different districts are called Bilaspur, we rename these districts to Bilaspur (cp berar) and Bilaspur (Punjab) to reflect that these are two separate districts in two distinct provinces.

Mapping princely states to political agencies. Under British rule, there were various political agencies. These are autonomous or semi-autonomous units and are mostly comprised of princely states. The geographic concordance provides information on these princely states and how they map to pre- and post-independence Indian districts. However, the Census reports information on political agencies (an aggregate unit that comprises of these princely states). To know how the princely states map to political agencies in pre-independence India, we visually compare 1911, 1921 and 1931 maps from the Indian Census atlas with census province maps (mainly for Bombay presidency) to concord these princely states with political agencies. For the 1911 Census, while the political agencies are directly reported, for 1921 and 1931 Censuses, information for some of the main princely states is reported and data for the other princely states is reported as "Rest of political agency". So for 1911 we concord the princely states with political agencies. For 1921 and 1931, we concord princely states (which are not directly reported in the census) with "Rest of the appropriate political agency". In 1911, 1921 and 1931 some princely states are in the Indian atlas but they are missing in the geographic boundaries concordance. In these cases, by visual comparison we assign these princely states to political agencies in pre-independence years and map them to 2011 districts by relying on district census handbooks and official district webpages. We provide more detailed explanations below.

1911:

- Balasinor, Lunwara, Sunth, Bariya, Chota Udepur, Pandu Mewas and Rajpipla in the geographic boundaries concordance map to Rewa Kantha Agency in the 1911 census. Kadana, Sanjeli and Sankheda Mewas are in Census atlas but not in geographic boundaries concordance. We consider these to be a part of Rewa Kantha Agency in 1911. We map these three princely states to modern day Indian districts by relying on census webpages. According the Indian webpage for Dahod district, Sanjeli is one of the sub-districts of Dahod. Similarly, according to Indian government webpage for Mahisagar district (carved out of Kheda and Panch Mahals), Kadana is one of the sub-districts of Mahisagar and according to the 2001 census handbook, Sankheda is one of the sub-districts of Vadodara. All these districts map to the consistent district unit dist_id dist0101.

- Santalpur, Wardhi, Radhanpur, Wao, Tharad, Diodar, Kankrej and Palanpur map to Palanpur Agency
- Porbandar, Junagarh, Bhavnagarh, Palitana, Lathi, Jasdan, Kotra Sangani, Bagasra, Jetpur, Gonda, Nawanagar, Dhrol, Khairasra, Rajkot, Morvi, Wankaner, Malia, Dhrangadra, Eastern Kathiawar Agency and Jafrabad map to Kathiawar agency. Bantwa, Bagasra and Virpur are in Census atlas but not in geographic concordance. We map these to Kathiawar agency in 1911. According to the 2001 district census handbook for Junagadh, Bantwa princely state is in Junagadh. Similarly, we see that there is a village called Virpur in Junagadh princely state. We geographically confirm the location by visually comparing Virpur princely state in 1911 with modern day Junagadh district. The 2011 district handbook of Amreli district mentions that the district comprises of former princely state Bagasara.
- Jat maps to Bijapur Agency.
- Mewas maps to Khandesh Agency.
- Oundh and Phaltan map to Satara Agency.
- Bansdra, dang and Dharampur map to Surat Agency.
- Jamkhandi, Mudhol, Ramdurg, Sangli, Miraj map to Southern Marath Jagirs.
- Athmallik, Pal Lahara, Kalahandi, Sonpur, Baud, Daspalla, Baramba, Dhenkanal, Hindol, Talcher, Rairakhol, Bamra, Keonjhar, Mayurbhanj, Nilgiri and Patna (different from Bihar capital) map to Orissa feudatory states. Ranpur, Nayangarh, Khandpara, Narsinghpur, Tigiria, Athgarh are present in the census atlas but missing in the geographic boundaries concordance. In 1911, we map these to be orissa tributary states. In 2011, we rely on district gazattes and handbooks to map these princely states. According to the Orissa District Gazetteer from 1993, Narsinghpur, Athgarh and Tigria are mentioned as sub-divisions of the Cuttack district and according to the 2011 census handbook, Ranpur, Nayagarh and Khandapada (Khandpara) are blocks (administrative sub-divisions) in Nayagarh district.

- Gangpur, Bonai, Saraikela, Kharsawan map to Chota Nagpur States.

1921:

- Athmallik, Pal Lahara, Kalahandi, Sonpur, Baud, Daspalla, Baramba, Dhenkanal, Hindol, Talcher, Rairakhol, Bamra, Keonjhar, Mayurbhanj, Gangpur, Nilgiri and Patna (different from Bihar capital) map to Orissa states. Similar to 1911, some princely states (Ranpur, Nayangarh, Khandpara, Narsinghpur, Tigiria, Athgarh) are in the Indian census atlas but missing from the geographic boundaries concordance. We treat them in same manner as 1911.
- Santalpur, Wardhi, Radhanpur, Wao, Tharad, Diodar, Kankrej map to Rest of Palanpur Agency.
- Balasinor, Lunwara, Sunth, Baria, Chota Udepur, Pandu Mewas map to Rest of Rewa Kantha Agency.
- Bansda, Dangs, Dharampur map to Surat Agency.
- Porbandar, Palitana, Lathi, Jasdan, Kotda Sanghani, Bagasra, Jetpur, Dhrol, Khairasra, Rajkot, Morvi, Wankaner, Malia, Dhrangadra, Eastern Kathiawar Agency and Jafrabad map to Rest of Kathiawar. Bantwa, Bagasra and Virpur are missing in the Indian census atlas but missing from the geographic concordance. We treat them in the same manner as 1911.
- Jamkhandi, Mudhol, Ramdurg, Miraj map to other Southern Maratha Country states.
- We map the remaining unassigned princely states to "Other states in the presidency proper" row in census. The princely states that are mapped in this manner are Oundh, Phaltan, Janjira, Jawhar, Surgana, Jath and Akalkot.

1931:

- Deodar, Kankarej, Santalpur map to Rest of Banaskantha.
- Pandu Mewas maps to Rest of Reva Kantha.
- Saraikela maps to Chota Nagpur States.
- Athmallik, Pal Lahara, Kalahandi, Sonpur, Baud, Daspalla, Baramba, Dhenkanal, Hindol, Talcher, Rairakhol, Bamra, Keonjhar, Gangpur, Mayurbhanj, Nilgiri and Patna (different from Bihar capital) map to Orissa feudatory states. Ranpur, Nayangarh, Khandpara, Narsinghpur, Tigiria, Athgarh are in the Indian census atlas but missing from the geographic concordance. We treat them in the same manner as 1911.

1991: In 1991, states are coded incorrectly for Raigarh in the geographic boundaries concordance. There are two Raigarh districts in India, one in Maharashtra and the other in Madhya Pradesh. For 1991 only, Raigarh in Madhya Pradesh is erroneously coded to be from Maharashtra. We make this correction in the geographic boundaries concordance to correctly map 1931 districts to 1991 districts.

B.8 Modification and missing data in Censuses

We make minor modifications to censuses for different years which we document in this section. We also make spellings of different districts consistent between the census data and the geographic boundaries concordance.[29]

1931: Miraj appears as Miraj (junior) and Miraj (senior) in census data. Dewas appears as Dewas (junior) and Dewas (senior) in census data. We sum up the occupation and population data for Miraj (junior) and Miraj (senior) data to obtain the data for Miraj as a whole and similarly for Dewas. We have to do this because we do not see separate entries for senior and junior in the geographic boundaries concordance.

After visually comparing map of Bombay Presidency from pre-independence census reports and the census atlas, we determine Wardhi and Varahi to refer to the same region. We map "d. s. vala mulu surang of jetpur (pithadia)" and "d. s. vala rawat ram of jetpur (bilkho)" in the 1931 census with Jetpur in the 1931 geographic boundaries concordance.

In addition to territories under French or Portuguese rule, we do not map a few districts that are in the pre-independence geographic concordance to the relevant census year either because we could not locate relevant detailed census data for these districts or detailed occupational data was not collected for these districts. Examples of such districts for 1931 are given below.

- Rampura (Indore)
- Kukshi (Indore)
- Nemawar (Indore)
- Rann of Kutch
- Naga Tribes
- Singhpos
- Districts from Rajputana Agency
- Lakshadweep Islands

In addition, for 1911 and 1921, we were not able to locate census data on districts in Central India Agency. As a result, these districts drop out of our pre-trend analysis.

1981. For the year 1981, we were not able to locate census data for Assam and Arunachal Pradesh states. However, districts in these states are not a part of our main sample.

2010 and 2011 Concordance. For the years 2010 and 2011, we map all of Delhi National Capital Territory region to 1931 Delhi. In the geographic boundaries concordance, 1931 Delhi is mapped to just West Delhi in 2010 and 2011.

[29] For details, see our census-Fenskeconcordance crosswalk files.

B.9 Controls

In this section we explain the data sources and calculation of economic, historic, geographic and military controls.

B.9.1 Economic controls

- To construct 1931 economic controls (i.e. share of employed in agriculture, industry, heavy industry, services, log population etc.,) we use data from the 1931 Census. Similarly, to construct 1911 and 1921 economic controls, we use data from the 1911 and 1921 Censuses respectively.
- To construct the population density variable, we divide population in the baseline year (obtained from the Census) with the area of that district (in square kilometers).
- To construct years of railway availability we use data on construction of first railway line from Dincecco et al. (2022). For a certain district, if we do not see year of railway construction, we assume a railway line was built after 1931. We use the year of railway construction variable to calculate years of railway availability by subtracting year of first railway line from the relevant decadal census year. For 1911 and 1921 pre-trend test, we control for years of railway availability up to 1911 and 1921 respectively. For our other main analysis, using 1931 as the baseline, we control for years of railway availability up to 1931. We use the maximum of railway availability for different constituent districts when constructing this control variable for consistent geographic units.

B.9.2 Geographic controls

For each of the analysis years, using the Fenske et al. (2022) geographic boundaries crosswalk, we construct a new map with geographic units (districts) such that boundaries stay consistent between 1931 and the analysis year.[30] We then use these maps to construct geographic controls as described below.

- **Total Area:** For the consistent geographic units, in each of the analysis panels, we calculate total area in square kilometers using the new maps (shapefiles).
- **Precipitation and Temperature:** Precipitation and temperature data are downloaded from the GAEZ data portal in the form of raster files.[31] The precipitation raster provides information on annual precipitation in millimeters (for 1960) using CRUST32 climate model. The temperature raster provides information on mean annual temperature in degrees Celcius (for 1960) estimated using CRUST31 climate model. Using this data, we calculate mean temperature and mean annual precipitation for each of the consistent geographic units.

[30] In our data, these larger geographic units are identified by dist_id.

[31] The link is https://gaez.fao.org/pages/data-viewer.

- **Average maximum caloric yield:** The maximum caloric yield per pixel is constructed taking inspiration from Galor and Özak (2016). The yield per pixel measures estimate potential yield under low levels of inputs and rain-fed agriculture, reflecting the agricultural methods and techniques used in India prior to the Green Revolution. The data for potential yield are from the Global Agro-Ecological Zones (GAEZ) project of the Food and Agriculture Organization (FAO). Potential yields under low level of inputs and rain-fed agriculture are calculated for 31 crops. To calculate maximum caloric yield, we then map the caloric content of these crops using SR28 reports from the U.S. Department of agriculture. Data on caloric content of different crops can be accessed at https://www.ars.usda.gov/.
- **Elevation and slope:** Elevation and slope data are from Fick and Hijmans (2017).[32] These data provide elevation information in meters at the 30 arc-second resolution (approximately at the 1 km square level near the equator). The elevation measure is constructed using NASAs SRTM satellite images.[33] Using this data, we calculate mean slope and elevation for each of the consistent geographic units.

B.9.3 Historic controls

We use data from Dincecco et al. (2022) to obtain variables on historic controls.[34] These data provide information on conflict within 250 km up to 1757 AD, being under direct British rule, and the year first railway line was available (up to 1934) for modern Indian districts. We concord this this with 2011 Indian districts[35]. Using the 1931-2011 geographic boundaries concordance described in B.7 we map these controls to 1931 districts. If multiple 2011 districts are mapped to 1931 district, we take the mean of the conflicts to assign to the 1931 district. If half or more than half the 2011 districts that are mapped to a 1931 district are under direct British rule, then we assign the 1931 district to be under direct British rule. If less than half of the 2011 districts mapped to a 1931 district are under direct British rule, we code the 1931 district as not being under direct British rule.

B.9.4 Military controls

We obtained the data for military participation from Jha and Wilkinson (2012). This dataset provides data on casualties per 1951 district (also mapped to 1931 districts) and males of "good martial race" in 1931. We use this data to construct variables

[32] Downloaded from https://www.worldclim.org/data/worldclim21.html.

[33] Available at http://www2.jpl.nasa.gov/srtm/.

[34] These data are publicly available and can be downloaded from https://warwick.ac.uk/fac/soc/economics/staff/jefenske/data/.

[35] See crosswalk_GeographicConcordance_HistoricRailwayControls.xlsx from our replication package.

capturing number of World War II casualties (per million) and number of males from martial castes (per thousand) which allow us to control for Indian military participation. We then map the districts from this dataset to the 1931 geographic boundaries concordance and construct a binary indicator for whether a census district has military participation controls.

Finally, when we collapse the 1931 districts to form consistent geographic units with a different year, we weight the dummy for being under direct British rule, conflicts, and dummy for war controls by the population of that district relative to the population of the larger 'consistent' geographic unit.

B.10 Other outcomes

B.10.1 Nighttime lights

The data for 2023 VIIRS nighttime lights is the Earth Observation Group.[36] These data provide information on average nightlight intensity per pixel measured in nanowatts per square centimeter per steradian. We calculate the average nighttime light intensity measure for each consistent district unit in the 1931 and 2011 panel and use the logarithm of the average as our dependent variable.

B.10.2 Consumption and poverty rates

Data on consumption and poverty rates are downloaded from 2012 Socioeconomic and Caste Census (SECC) module from SHRUG data development lab.[37] We use data on estimated per capita monthly consumption (in 2012 INR) and estimated share of households consuming less than 33 INR per day in 2012 as our dependent variables. SHRUG provides data on these variables separately for rural and urban areas and provides these estimates for 2011 census districts. We then map these districts to our geographic boundaries concordance and calculate the population-weighted estimates for each consistent geographic unit between 1931 and 2011. We also estimate "district level" consumption per capita and poverty rates for each consistent geographic unit by averaging (population weighted) urban and rural measures.

B.10.3 In migration

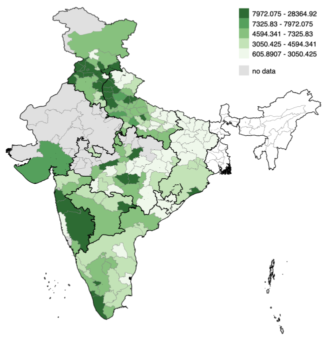
Data for migration are from 2011 Census tables D-01 (Population Classified by Place of Birth and Sex) and D-11 (Persons born and enumerated in Districts of the State). Table D-01 provides information on total population in the district of enumeration, number of people born in that district, people born in the state in which the district is from, people born in India and so on. Using this information we calculate the number of people living in a certain district born out of India, born outside the

[36] Downloaded from https://eogdata.mines.edu/products/vnl/#annual_v2.

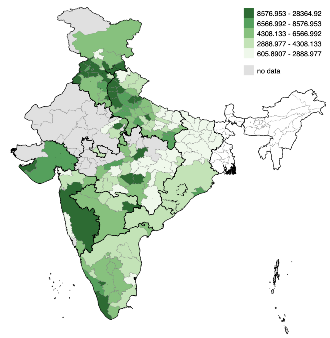
[37] This dataset can be accessed from https://www.devdatalab.org/shrug_download/.

district, born outside the state (including and excluding those born outside of India). For each consistent geographic unit between 1931 and 2011, we then calculate the share of population in that district that was born out of that district to construct our measure of in-migration. Table D-11 provides information on number of persons that are enumerated in a district born in other districts of that state. For each consistent geographic unit, we then calculate the share of population in a district born in that district but living in another district (in that state). Delhi is considered both a district and its own state (Union Territory) in the 2011 Census. As a result our out-migration measure for Delhi would be 0. We therefore treat Delhi as part of Haryana to construct our out-migration measure as Delhi borders Haryana state. Therefore for Delhi (only), our out-migration measure captures those who are born in Delhi but are living in Haryana as a share of Delhi's 2011 population.

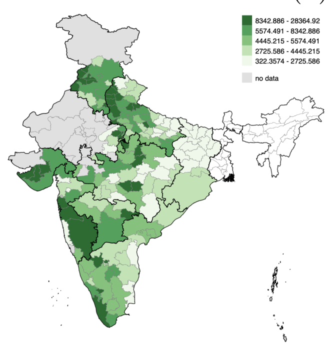
C Additional Maps



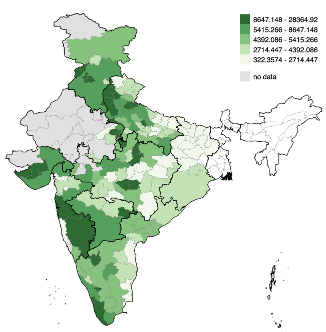
(a) 1911-1931-2011



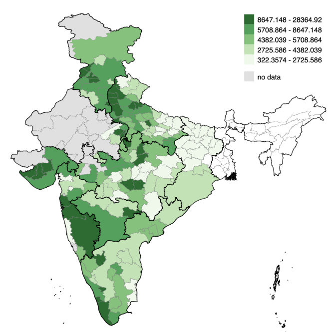
(b) 1921-1931-2011



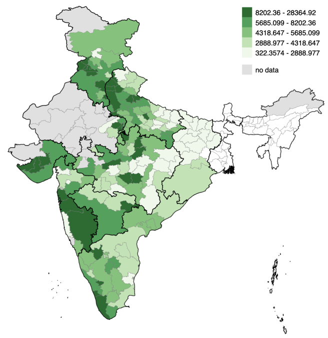
(c) 1931-1951



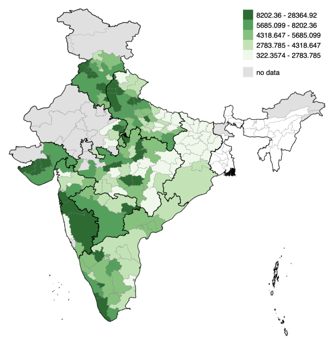
(d) 1931-1961



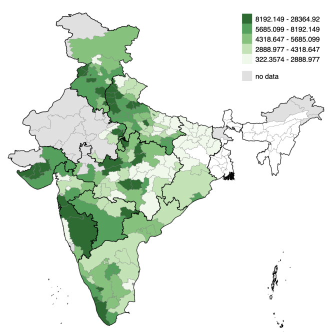
(e) 1931-1971



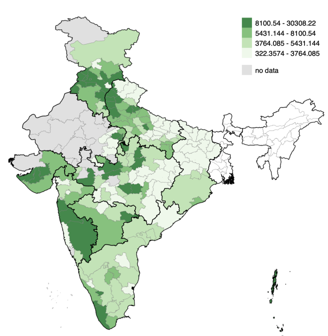
(a) 1931-1981



(b) 1931-1991



(c) 1931-2001



(d) 1931-2011