# Industrial clusters in the long run: Evidence from Million-Rouble plants in China*

Stephan Heblich  Marlon Seror  Hao Xu
Yanos Zylberberg

September 12, 2025

**Abstract**

We study the impact of large, successful manufacturing plants on other local producers in China, focusing on "Million-Rouble Plants" built in the 1950s during a brief alliance with the U.S.S.R. The ephemeral geopolitical situation and the locations of allied and enemy airbases provide exogenous variation in plant siting. We find a boom-and-bust pattern: Counties hosting these plants were 80% more productive than control counties in 1982 but 20% less productive by 2010. This decline reflects the performance of local establishments, which exhibit low productivity, limited innovation, and high markup. Specialization hindered spillovers, preventing the emergence of new clusters and local entrepreneurship.

JEL codes: R11, R53, J24, N95

*Heblich: University of Toronto, NBER, stephan.heblich@utoronto.ca; Seror: Université du Québec à Montréal, seror.marlon@uqam.ca; Xu: China Construction Bank, sysxuhao@163.com; Zylberberg: University of Bristol, CEPR, yanos.zylberberg@bristol.ac.uk. This work was part-funded by the Economic and Social Research Council (ESRC) through the Applied Quantitative Methods Network: Phase II, grant number ES/K006460/1, and a BA/Leverhulme Small Research Grant, Reference SRG171331. We are grateful to Sam Asher, Kristian Behrens, Sylvie Démurger, Christian Dustmann, James Fenske, Richard Freeman, Jason Garred, Ed Glaeser, Flore Gubert, Marc Gurgand, Ruixue Jia, Vernon Henderson, Matthew Kahn, Francis Kramarz, Sylvie Lambert, Yao Luo, Florian Mayneris, Alice Mesnard, Thomas Piketty, Simon Quinn, Steve Redding, Uta Schönberg, Arthur Silve, Helen Simpson, Jon Temple, and Liam Wren-Lewis for very useful discussions. We also thank conference and seminar participants at Bath, Bristol, CREST, DIAL, Geneva, Georgia Tech, Goettingen, Laval (Quebec), LSE, Ottawa, Oxford, PSE, Toronto, UCL, UQAM, the EUEA, the EEA, the UEA, the IOSE, CURE, the China Economy Seminar (Harvard), and Cities & Development (Harvard) for helpful comments. We thank Liang Bai, Matthew Turner, Andrew Walder, and Siqi Zheng for sharing data. We are also grateful to Ziyu Peng, Junhui Yang, and Kaixin Zhao for excellent research assistance. The usual disclaimer applies.

The structural transformation of agrarian economies entails a significant spatial concentration of economic activity (Kim, 1995; Henderson et al., 2001). Regions that attract successful industries during this process typically experience a boom followed by a bust, as illustrated by the decline of factory towns in the United States (Detroit and the "Rust Belt"), the United Kingdom (Manchester and other mill towns), the Ruhr region in Germany, or the Northeast of France. Explanations for this decline often hinge on external, macro-economic factors such as shifts in employment away from industry (Ngai and Pissarides, 2007; Desmet and Rossi-Hansberg, 2014), exposure to increased import competition (Autor et al., 2013), or changes in trade policy (Pierce and Schott, 2016).

This paper takes a different perspective and focuses on factors *within* the local economy that contribute to the bust. This perspective holds great relevance because here policymakers can effectively wield their influence to help revive these local economies.

Our analysis exploits an unprecedented industrial policy that involved the most comprehensive technology transfer in modern industrial history and served as the foundation of China's industrialization (Lardy, 1987; Naughton, 2007). As part of the *Sino-Soviet Treaty of Friendship, Alliance and Mutual Assistance*, the U.S.S.R. assisted China in constructing 149 "Million-Rouble Plants" (MRPs) in the 1950s.[1] These plants gave China access to frontier technology, and they were scattered across 98 counties in an attempt to jump-start their local economies. We follow locations that benefited from these initial investments and analyze the dynamics of the industrial clusters that emerged around them. Our research presents causal evidence that these industrial investments initially generated positive spillovers, creating a "boom." However, even as the MRPs continued to thrive in the long run, they eventually unleashed significant negative production spillovers, leading to a "bust."[2] This decline is independent of the local composition of industries and products. Consequently, we can dismiss *external* macroeconomic trends as drivers and shift our focus toward explaining the emergence of negative production spillovers *within* the local economy. Using firm-level data on production and innovative activities, we show that establishments in host economies are not productive, not competitive, and not innovative. Host regions are characterized by a specialized production structure that is oriented toward the MRPs. Technological spillovers are limited, and there is little room for other industries to emerge. Consistent with this, we observe that cohorts of college graduates who have benefited the most from Chinese growth

[1]These plants were part of the "156 Program," which aimed to build 156 large industrial projects, though not all were ultimately constructed.

[2]A recent contribution focusing on the nature of technological transfers studies a subset of MRPs that were in the steel industry and shows that the (full) reception of Soviet knowledge and equipment benefited (i) these plants in the long run; and (ii) connected firms (Giorcelli and Li, 2022). Our findings complement this insight, by considering general spillovers on indirectly-connected firms, which account for most of the aggregate local output.

are increasingly less likely to occupy top entrepreneurial positions in MRP counties, a pattern partly explained by the flight of prospective entrepreneurs starting businesses elsewhere.

The credibility of this analysis critically hinges on the identification of agglomeration spillovers. In our context, achieving this requires exogenous variation in the allocation of large plants across space. In an ideal setting, actual project sites would naturally have a set of counterfactual sites, similar to the list of candidate locations employed in Greenstone et al. (2010), along with exogenous variation in the selection process among these sites. We emulate this setup in two steps. We first rely on the *economic* criteria that policymakers used to select a set of suitable counties (Bo, 1991).[3] Next, we exploit the ephemeral *geopolitical* context—the short-lived Sino-Soviet Treaty of Friendship, Alliance and Mutual Assistance—to isolate temporary, exogenous variation in the probability of hosting a large factory, within the subset of suitable counties. After the Korean war, Chinese planners were keenly aware of the new factories' vulnerability to enemy bombing; this had a marked effect on location choices.[4] To exploit this variation, we combine novel information about airplane technologies with the location of enemy and allied air bases and derive a measure of vulnerability to aerial attacks from major U.S. bases for the period between 1950 and 1960. We use this measure as an instrument for the probability of hosting an MRP. To rule out possible confounding effects from later place-based policies and industrial investments, we condition our analysis on similar measures of vulnerability computed after the Sino-Soviet split in 1960.[5] One remaining concern is that the instrument may be coincidentally correlated with other geographic factors underlying recent economic growth, such as proximity to the coast, major trading ports, and booming industrial centers in the South. We present a variety of robustness checks to alleviate such concerns. Among others, we show that our findings are robust to flexible controls for first-nature determinants (travel time to the coast, local geography, pollution), second-nature determinants (proximity to booming regions), the secular decline of public enterprises, and controls for later spatial policies such as the Special Economic Zones, the Third Front Movement, or transport infrastructure.

[3]In stark contrast to the Third Front Movement, this program was efficiently implemented. The location choice was economically sound (e.g., based on market access and access to natural resources), and attention was paid to production efficiency, including material incentives for managers (Eckstein, 1977; Selden and Eggleston, 1979) and technology transfers (Giorcelli and Li, 2022).

[4]Senior generals were directly involved in siting decisions to protect the state-of-the-art factories from enemy airstrikes, using intelligence maps of the U.S. and Taiwanese air bases (Bo, 1991). Historical U.S.S.R. documents report the same strategy for locating Soviet Science Cities out of the reach of enemy bombers (Schweiger et al., 2022).

[5]After the Sino-Soviet split in 1960, the set of protected locations shrank, which led to directing new industrial investments to the interior during the Third Front Movement ("close to the mountains, dispersed, and hidden in caves").

Our paper combines county-level data on economic activity with a census of manufacturing firms, which we complement with patent applications, productivity measures, and markup calculations. We find that locations that received MRPs experienced rapid industrialization from 1953 to 1982: At the onset of the Opening and Reform era, these counties had productivity levels 80% higher than those observed in counties without such initial investments. Despite the head start provided by the MRPs and the subsequent investments of the planned economy era,—which we consider as our "treatment," reflecting the central role of planning in allocating resources during this period,—treated counties saw their relative advantage steadily erode over the following decades. By 2010, average productivity was *lower* in treated counties, even though the (large) MRPs remained productive and innovative. This is because other manufacturing establishments in treated counties were (i) less productive, (ii) less competitive, and (iii) less innovative than establishments in control counties, conditional on establishment age, ownership, and detailed industry or product category.

What explains the poor performance of these other establishments? To answer this question, we must understand how MRP(s) affected their local economies. This is challenging because each MRP draws on different factor markets and operates a distinct technology to produce different products. Such heterogeneity among MRPs rules out the application of a standard difference-in-differences approach, since we would need to identify sub-populations of firms *in control counties* that are comparable to the sub-populations of firms with possible links to MRP(s) in treated counties. To address this challenge, we exploit our baseline two-step identification procedure. In the first, "matching" stage, each treated county is paired with a control county, allowing us to hypothetically assign the MRP(s) of the treated county to its control counterpart. To infer linkages between the assigned MRPs and other establishments in terms of factor markets, production lines, or technological spillovers, we rely on the procedure developed in Imbert et al. (2022), which associates a precise product code with the textual product descriptions provided by manufacturing establishments.

We gain two important insights. First, we find that treated counties become concentrated in a large share of firms operating through the MRPs' production chain. However, those firms that are upstream and downstream of the MRP(s) are not productive, not competitive, and not innovative.[6] Outside this chain, production is scattered across small clusters. Second, we shed further light on the far-reaching externalities exerted by MRPs and look at *missing* production and entrepreneurs: Using micro-level census data,

[6]We find limited evidence of treatment heterogeneity depending on the position of MRPs in the production chain, the size of the initial investment, their industry (heavy or light, multi- or single- product), their reliance on natural resources, or their completion date. Among these various dimensions of treatment heterogeneity, MRPs operating upstream appear more detrimental to productivity and innovation.

we show that individuals who are potentially responsible for firm creation are less likely to occupy top entrepreneurial positions in treated counties—a finding partly explained by the fact that prospective entrepreneurs export their skills to other locations (reminiscent of Chinitz, 1961). One concern with these insights is that China is characterized by a unique set of institutions and policy changes. We provide a detailed discussion supported by empirical tests (covering, among others, the role of factor misallocation, urban disamenities, corruption and the local political environment, or public ownership of firms) that alleviates these concerns.

Our overall contribution is to study the evolution of production externalities over a long period of time. Our careful analysis of innovation, productivity, and production linkages at the firm level reveals *negative* long-run production externalities. The increasingly specialized production structure around the MRP(s) comes at a cost: Linked firms get locked in and face little incentive to innovate. This constrains the potential for between-industry spillovers that could otherwise benefit entrepreneurs in new industries, highlighting a fundamental trade-off between the short-run benefits of specialization and its long-run crowding-out effects.

Our first contribution is to show that the nature and direction of agglomeration economies can change as industrial clusters mature. There is ample evidence that agglomeration economies benefit the emergence of industrial clusters and local economic growth, as reviewed in Duranton and Puga (2014), and recent evidence sheds light on the positive economic effects of large-scale industrialization (Greenstone et al., 2010; Fan and Zou, 2021; Mitrunen, 2025; Garin and Rothbaum, 2025). Our findings complement these contributions in that we also observe positive effects during the initial boom phase, but, as the industrial clusters mature, they become more specialized (Kim, 1995; Henderson et al., 1995; Ciccone, 2002), and we observe that this increasing specialization gives rise to *negative* externalities that lead to a decline in the local economy (echoing a literature discussing the negative effects of specialization, see Duranton and Puga, 2001; Faggio et al., 2017).[7] This main finding casts doubt on the general effectiveness of place-based

[7] The difference between our findings and those from this recent literature, most notably Mitrunen (2025) and Garin and Rothbaum (2025), could be due to our respective settings (e.g., the geography of industrialization programs and the time horizon of estimated spillovers). For example, Garin and Rothbaum (2025) examine investments in locations that were "dispersed" or remote, at least at the time of investment; evidence from China's Third Front Movement indeed similarly suggests positive long-term effects in such a setting (Fan and Zou, 2021). Another possible difference is the time horizon: Mitrunen (2025) looks at manufacturing outcomes and census information around 1970 (20-25 years after treatment), which would correspond to what we define as short term—or the treatment period—in our context, and Garin and Rothbaum (2025) primarily use tax returns between 1969 and 1984 (30-40 years after treatment); the latter also use decennial censuses that suggest a partial return to the mean by 1990, e.g., in terms of the share of manufacturing employment. A last difference could be the nature of initial industries and their later, aggregate dynamics; entrepreneurial individuals in a location might be more likely to invest in an industry or a skill if this industry or skill does particularly well *at the aggregate level*. We think of our

policies in creating *self-sustaining* economic gains, integrating two overlooked insights from Greenstone et al. (2010) and Bleakley and Lin (2012). Greenstone et al. (2010) show that spillovers from "Million-Dollar plants" are highly heterogeneous and occur along labor pooling or technological linkages, which condition the positive effect that they observe on local economies in the short run, and might lead to a specialization of these economies. Bleakley and Lin (2012) show that portage towns characterized by a more diverse economy were more likely to adapt to changes in transportation technology and persist in the longer run. A contribution of this paper is to provide the first empirical evidence that large, productive firms might have a positive impact on local economies in the short run *and* a negative impact in the long run through the lack of economic diversity—when the latter effect dominates.

Our second contribution is to provide a detailed analysis of negative externalities arising from specialization. We use firm-level data to show that establishments operating upstream or downstream from the MRPs refrain from innovating, which limits technological spillovers and stifles the entry of new industries (Glaeser et al., 1992; Henderson et al., 1995) and entrepreneurs (Chinitz, 1961; Glaeser et al., 2015). We consider a range of competing channels and show that the decline of local economies is not due to: (i) high production costs and inefficient provision of factors (Duranton, 2011); (ii) distortions in the allocation of physical capital (Hsieh and Klenow, 2009; Song et al., 2011), human capital (Franck and Galor, 2021), land (Brueckner et al., 2017; Yu, 2024), or labor (Brandt et al., 2013; Tombe and Zhu, 2019; Mayneris et al., 2018); (iii) urban (dis)amenities such as pollution (Chen et al., 2022; Khanna et al., 2025) or sprawl; (iv) the local political environment (Chen et al., 2017; Wen, 2019; Fang et al., 2023; Harrison et al., 2019); or (v) the life cycle of firms, industries, and public establishments (Brandt et al., forthcoming). This strengthens our argument that production and technological spillovers are the key force driving the decline in hosting counties.

Finally, our work contributes to a growing body of research on place-based policies including Busso et al. (2013), Kline and Moretti (2014a), von Ehrlich and Seidel (2018), Schweiger et al. (2022), and Fajgelbaum and Gaubert (2020). Kline and Moretti (2014b) and Neumark and Simpson (2015) review this literature. A subset of this literature focuses on place-based policies in China: Fan and Zou (2021) analyze the agglomeration effects of industrial clusters that were established under the Third Front Movement; and Wang (2013), Alder et al. (2016), and Zheng et al. (2017) evaluate Special Economic Zones and industrial parks. Related to our work is a recent paper by Giorcelli and Li (2022), which focuses on MRPs in the steel industry and disentangles the effect of technology

contribution as (i) providing an example of when agglomeration effects of successful, local industries turn negative and (ii) discussing potential channels through which it may happen.

versus knowledge transfers on longer-run plant performance. Our paper focuses instead on the long-run development of the wider local economies *around* MRPs.

The remainder of the paper is organized as follows. Section 1 describes the historical context. Section 2 details the data and the empirical strategy. Section 3 presents empirical facts about the rise and fall of early-industrialized counties. Section 4 provides evidence for the mechanisms behind the relative decline of treated counties, and Section 5 briefly concludes.

## 1 Historical context and the "156" program

The Million-Rouble plants were built as part of the "156" program. This program represents a unique experiment to study agglomeration spillovers in the long run. First, the geopolitical context introduces exogenous variation in the decision to locate projects. The "156" program was unanticipated before 1950, and strategic considerations behind the opening and location of plants became irrelevant a few years later, after the Sino-Soviet Split. Second, the program constitutes a large push shock for an agrarian economy (Rawski, 1979), and factories were built across a wide range of sectors.

### 1.1 The historical context

This section provides a brief account of the historical context; a comprehensive description can be found in Appendix A.1.

**Sino-Soviet cooperation (1950–1958)** In 1949, after the Sino-Japanese and Chinese civil wars, Chinese leaders studied the possibility of international cooperation to transform China's agrarian economy and foster the development of an independent industrial system (Dong, 1999; Lüthi, 2010). The Chinese government engaged in economic cooperation with the Soviet Union for both ideological and geopolitical reasons. The possibility of economic cooperation with the U.S.S.R., which was not grounded in strong pre-existing economic ties, indeed became credible after the Sino-Soviet Treaty of Friendship and Alliance of 1950—already including a large loan. In August 1952, Chinese Premier Zhou Enlai visited Moscow to formalize the involvement of the U.S.S.R. in the First Five-Year Plan (1953–1957). The U.S.S.R. agreed to cooperate and assist China in the creation of state-of-the-art industrial sites, with the purpose of extending its influence in the region.

**Sino-Soviet Split (1958–1960)** Rapid ideological and geopolitical divergence however precipitated a Sino-Soviet split that ended the cooperation between the two countries.

The split formally unfolded in 1960 with an abrupt termination of industrial collaboration and heightened military tensions. The termination of industrial collaboration materialized in the sudden withdrawal of experts and engineers from China, the repatriation of Chinese students from the U.S.S.R., and the cancellation of ongoing industrial projects. The only remnants of the short-lived Sino-Soviet alliance were 149 plants that had been completed and were operational by 1960 or shortly thereafter.

### 1.2 The "156" program

This section summarizes the key features of the "156" program. We provide a detailed description of the program in Appendix A.2, a description of later place-based policies in Appendix A.3, and descriptive statistics about the "156" plants in Appendix A.4.

**An industrial collaboration** As part of the First Five-Year Plan (1953–1957), the Soviet Union committed to assisting China in the construction of 50 industrial sites. In May 1953, 91 new projects were agreed on and an additional 15 in October 1954, reaching a total of 156 after which the program was named. Eventually, some modifications were brought to the program and 149 state-of-the-art factories were approved, which would theoretically be constructed between 1953 and 1958; the factories were huge investments and benefited from economic and technological assistance from the Soviet Union.

The U.S.S.R. actively participated in the design and construction of these factories. First, the economic aid from the U.S.S.R. extended beyond large loans; the U.S.S.R. provided more than half of the required equipment. For instance, the last 15 projects agreed on in 1954 even benefited from state-of-the-art equipment that few Soviet factories enjoyed (Goncharenko, 2002). Second, the collaboration involved the exchange of information, human capital, and technology, which in some cases was the best in the world (see Giorcelli and Li, 2022). During the peak of the cooperation, 20,000 scientific, industrial, and technical experts from the Soviet Union worked in China to design the construction of factories and rationalize production (Zhang, 2001; Wang, 2003). As part of the technology transfer, 80,000 Chinese students were trained in Soviet universities and technological institutes. While some blueprints were destroyed, the existing technology could be imitated, representing a large shift in the technological frontier (Bo, 1991).

Chinese scholars credit the "156" program with laying the foundations for the development of other industries, boosting production capacity, shifting the technological frontier, and promoting an even spatial development by industrializing central and western provinces (Dong and Wu, 2004; Zhang, 2009; Shi, 2013; He and Zhou, 2007). While these factories are known as the "156" in China, we refer to them as the "Million-Rouble plants" (MRPs), since only 149 plants were finally approved and were operational at the time of the Sino-Soviet Split or shortly thereafter.[8]

**Location decisions** The MRPs were regarded as iconic factories, and planners put much thought in siting decisions. First, planners selected locations using economic criteria. These criteria, detailed in Bo (1991), are: (i) connection to the transportation network and access to markets, (ii) access to natural resources through existing roads and rail, and (iii) belonging to an agrarian province, since the investments were seen as an opportunity to foster economic development outside of the few developed areas. We will use these criteria to identify a relevant set of suitable counties.

Second, this period was an era of heightened geopolitical tensions that culminated in the Korean War—where U.S. soldiers and Chinese "volunteers" directly clashed. Planners were concerned that these large, key factories might become the target of enemy attacks. The decision process involved senior military officials to decide on factory locations, accounting for the locations of enemy air bases (in Japan, South Korea, and Taiwan) and allied air bases (in the U.S.S.R. and North Korea). Enemy air bases were remnants of U.S. air bases used during World War II and the Korean War, and bases used by the United States Taiwan Defense Command. Most of Chinese territory lay within the range of U.S. strategic bombers; the decision process thus heavily relied on the locations of *allied* air bases able to intercept them. The Sino-Soviet split made this criterion redundant for later industrial investment: While proximity to military U.S.S.R. air bases would help counter aerial attacks before the Sino-Soviet Split, U.S.S.R. air bases were considered a threat after the Sino-Soviet Split, thereby explaining the peculiar geography of later strategic decisions (e.g., the Third Front Movement).

**Million-Rouble plants and economic growth** For the first 30 years of their existence, the MRPs operated within a planned economy. The factories and their local economies were sustained by provisions of the plan. Factor mobility was restricted, and when more workers or capital could be productively employed, the plan reallocated resources. We therefore treat the command-economy era as a whole as the treatment: counties with MRPs enjoyed a head start at the onset of the reform period.

Reforms to deregulate the economy were introduced in the 1980s. Private firms could be set up, and a dual price system introduced market transactions alongside the old quota

[8] Of the 149 projects approved in the First Five-Year Plan, several plants were not yet fully operational in 1960 (Giorcelli and Li, 2022), although only 2 were completed after 1962 (Dong, 1999). Another batch of projects was approved during the Second Five-Year Plan (see Appendix A.2); 63 were completed and 35 abandoned due to the Sino-Soviet split. We could identify and geolocate all those plants using historical archives. As these plants differed from the "156" in terms of sectors and the level of Chinese technological inputs, we exclude them from our definition of the treatment in the main analysis. However, we control for the presence of plants completed as part of the 2nd Five-Year Plan in a robustness check.

system. In the 1990s, restrictions on labor mobility were gradually relaxed, and rural-urban migration began to rise as a major feature of Chinese economic growth. The MRPs successfully adapted to the market economy and remained leaders in their respective industries.[9] Many of them have diversified their activities, with products ranging from computer screens to carrier-based aircraft.

## 2 Data and empirical strategy

This section discusses the data, the empirical strategy, and descriptive statistics.

### 2.1 Data

One requirement for estimating the agglomeration effects of large plants is to collect data on the local economy, ideally covering production dynamics from their openings to the present day. We mobilize the following main data sources: (i) information on the Million-Rouble plants and their evolution over time, (ii) county-level data on population and production at scattered points in time between 1953 and 2010, and (iii) establishment-level data in recent years, linked with patents and other product-level information.

**The Million-Rouble plants** In order to characterize the local treatment induced by the presence of MRPs, we collect information on their location, timing of construction, initial investment, original industry, and evolution of production over time. This information is extracted primarily from Bo (1991) and Dong and Wu (2004), as well as from historical archives, while the recent activity of these factories is obtained using establishment-level data (see Appendix A.4).[10]

**County-level data** We rely on Population Censuses in 1953, 1964, 1982, 1990, 2000, and 2010, and on the 2015 1% Population Survey, nested at the county level.[11] The 1953

[9] We provide evidence for the continued success of MRPs in Appendix A.4; they are also thoroughly studied in Giorcelli and Li (2022). Note, however, that a small number of firms went bankrupt. Nine factories closed down, all coal or non-ferrous metal mines. Two other firms, a paper mill and a former military electronics plant, were partly restructured and continue to operate. When construction plans were made in the 1950s, most plants were built in the city center. As pollution issues and the need for expansion had not been anticipated, nine plants were moved to the suburbs, within the same counties.

[10] Archival documents were primarily accessed through the National Library in Beijing and the Beijing City Archives. Additional information on the history of the plants was obtained from Tsinghua Library (Beijing) and the Universities Service Center (Hong Kong), e.g., form the gazetteers that many MRPs kept and published. The information from Dong and Wu (2004), e.g., locations, initial investments, and original sector of activity, is highly consistent with these archival documents. We also used the Baidu online encyclopedia to geolocate Second Five-Year Plan investments and check the identification of MRPs within the establishment-level data.

[11] Data collected by statistical offices—censuses, surveys, and yearbooks—rely on official administrative divisions at the time of data collection. County boundaries are subject to frequent and sometimes sub-

data only provide population and household counts, but subsequent censuses capture the agricultural status of households. At the time of the command economy, household registration (*hukou*) type faithfully reflects both activity and residential environment. This piece of information allows us to track the evolution of urbanization and economic sectors from 1964 onward. Additional county-level information is available in the 1982 census, most notably a disaggregation of employment by broad sectors and measures of output. In 1990, detailed information is collected on sector, type of employment, and occupation, as well as on housing characteristics. The 2000 and 2010 Censuses further include information on the place of residence five years earlier, timing of the last migration spell, reason for migrating, and place and type of household registration. The 2015 1% Population Survey, a nationally representative survey based on stratified multistage cluster sampling, also contains detailed textual descriptions of occupations, allowing us to capture the production and migration decisions of prospective entrepreneurs.

**Establishment-level data** We rely on the National Bureau of Statistics (NBS) "above-scale" firm data (sometimes called the *Annual Survey of Industrial Firms*, or *ASIF*), which constitute a longitudinal census of all state-owned manufacturing enterprises (SOEs) and of all non-state-owned manufacturing establishments, as long as their annual sales exceed RMB 5 million, over the period 1998–2007.[12] We use the establishment data to: (i) infer linkages between establishments and create measures of product concentration; (ii) estimate factor productivity; (iii) observe technological innovations; and (iv) create measures of markups to capture product competition. We first rely on text analysis of product descriptions to associate a 6-digit product code to each establishment (following Imbert et al., 2022, see Appendix B.4). We further complement the establishment data with product-level information, in particular a benchmark input-output matrix (United States, 2000), measures of technological closeness based on U.S. patenting (Bloom et al., 2013), and measures of factor intensity (Shirotori et al., 2010). We use the production functions derived in Imbert et al. (2022) to measure factor productivity (see Appendix B.5). We use the links from He et al. (2018) to match establishments with patent applications in three

stantial changes in China. To deal with this issue, we rely on yearly county maps between 1949 and 2015 (China Data Center, University of Michigan), use the 2010 administrative map of China as our benchmark, and re-weight the data collected in other years to match the 2010 borders. More precisely, we overlay the 2010 map with the map for every other year $y$ and create a new map with all the polygons defined by the 2010 and year-$y$ divisions. We then compute the area-weighted value of the variable of interest for each polygon and collapse the values at the level of the 2010 counties. We provide a description of the main county-level outcomes in Appendix B.3.

[12]These data contain accounting information at the level of "legal units." A legal unit can be a subsidiary of a firm, but has its own name and is financially independent (Brandt et al., 2014). Nearly 97% of legal units in our data correspond to single plants; we will refer to these units as establishments. We construct a panel spanning the period 1998–2007 thanks to the algorithm designed by Brandt et al. (2014).

categories (utility, invention, and design), and we rely on the procedure of De Loecker and Warzynski (2012) to estimate markups (see Appendix B.6).

### 2.2 Empirical strategy

We now describe the two steps of our empirical strategy. The first step consists of selecting a sample of counties based on their suitability for hosting a plant; and the second step consists of extracting exogenous variation in the allocation of Million-Rouble plants within these counties.

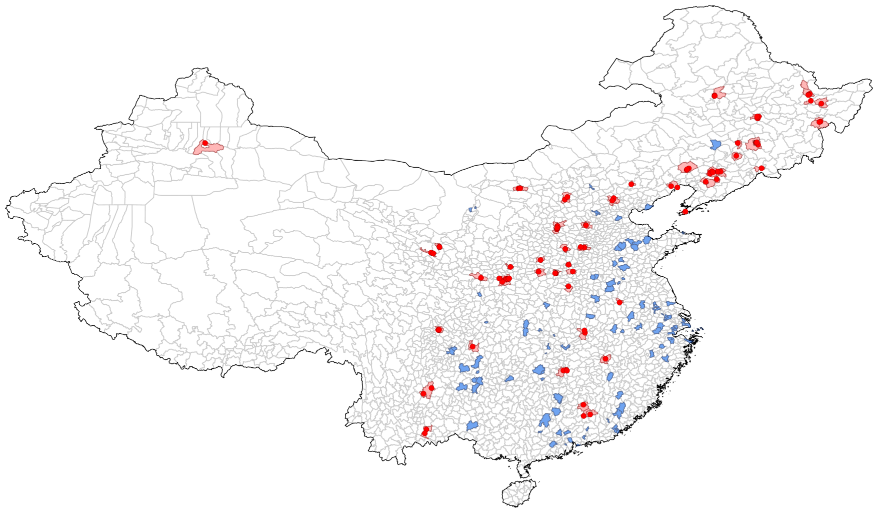
**Isolating suitable locations** We first isolate a group of suitable counties by implementing a one-to-one propensity-score matching based on the eligibility criteria described in Bo (1991). This step is not crucial for identification, and our findings are robust to its exact specification. It however disciplines the empirical strategy by reducing the sample of control counties to a set of most likely candidate locations and allows us to compare the subset of firms linked to MRPs in treated counties to a similarly defined subset in control counties.

We consider a specification at the county level (indexed here by $i$), and we predict the probability to receive the treatment (i.e., being within 25 kilometers of an MRP), $P_i = P(\mathbf{H}_i)$, by a set of observed location determinants, $\mathbf{H}_i$. In this step, our set of "explaining" variables only includes (economic) factors that were most explicitly mentioned by decision makers. The first suitability criterion is market access, which we measure with: (log) average travel cost to the nearest provincial capital at the time of the First Five-Year Plan using the existing road and railroad network around 1950; (log) population at baseline, as measured by the 1953 Census; and (log) county area. The second criterion is access to resources, which we proxy with (log) measures of travel cost to coal, ore, and coke deposits through the transportation network around 1950 (see Appendix C.1). The third criterion is (the lack of) early county-level industrial development, which we proxy with the (log) number of industrial firms over the period 1840–1937 (Du, 2014, 2019). In practice, we estimate the propensity score with a logit model, we use a one-to-one nearest neighbor matching using the odds ratio, and we limit the analysis to observations within the common support in terms of propensity score. In the baseline matching specification, we further impose that matched control counties be selected outside the vicinity of the treated counties, in order to avoid spillover effects on the control group.[13] This allows us to identify a set of suitable control locations $C = \{c_1, \dots, c_N\}$

[13]More specifically, we exclude counties whose centroids lie within $r = d/2$ of at least one treated county's centroid—where $d = 206.96$ kilometers is the 2.5% percentile in bilateral distance between any pairs of counties. This threshold excludes about a third of the 2,300 non-treated counties from the baseline matching procedure. We provide a more comprehensive description of the matching procedure in

matched with the treated counties $\mathcal{T} = \{t_1, \ldots, t_N\}$, giving us $N = 98$ matched pairs $\{(t_1, c_1), \ldots, (t_N, c_N)\}$, where each treated county is assigned to the nearest control county based on the propensity to receive an MRP.[14]

**Figure 1**. Treated counties and the group of control counties.



Notes: This map shows the treated group (red), i.e., counties with at least one "156"-program factory within 25 km of their borders, and the control group of counties (blue). The control group is selected through a one-to-one nearest neighbor (logit) matching based on the odds ratio and limited to the common support in terms of propensity score. Details about the matching procedure and the geography of control counties under alternative matching choices are available in Appendix C.1.

The geographic dispersion of treated and control counties is shown in Figure 1. Most treated and control counties are located along a "Second Front" crescent, with treated locations stretching from the Northeast region to Gansu/Shaanxi. The historical development of the railway network and the location of natural resources indeed induce a crescent of counties prone to receive large industrial infrastructure (see Appendix Figure C1). This crescent, located a few hundred kilometers from the Eastern coasts and borders, may be interpreted as a Second Front of industrialization; the later Third Front Movement will go deeper into the hinterland—a decision that will be rationalized by our empirical strategy.

Appendix C.1, where we show the distribution of propensity scores, the balance of matching variables within the selected sample of suitable counties, the geography of control counties under alternative matching choices, and the distributions of (i) bilateral distance between any pairs of counties and (ii) minimum distance to treated counties across all other counties. In a sensitivity analysis reported in Appendix D, we report our main estimates under alternative matching procedures, notably varying the set of matching variables (e.g., adding travel time to ports, or other proxies for initial industrial production), the size of the control set (with many-to-one matching), and the restrictions on spatial distances between matched counties with alternative exclusion zones.

[14]Our procedure only considers 98 treated counties, from 149 Million-Rouble plants, and thus identifies 98 control counties, because the average county hosts around 1.5 MRPs.

The rationale for using a rather parsimonious set of "propensity factors," $\mathbf{H}_i$, is to replicate the allocation process as explicitly described in archival material at the time of siting decisions. However, and although they do not feature among the list of explicit determinants, other geographical and economic factors may have more implicitly entered siting decisions, and such factors could affect the dynamics of economic development. Our primary approach to addressing these threats is to condition our baseline specification on a larger set of factors and to extract explicit exogenous variation in the allocation of treatment within these candidate counties.

**Empirical strategy and instrument** In a second step, we estimate the following TSLS specification on the sample of suitable counties $i \in \{t_1, \ldots, t_N, c_1, \ldots, c_N\}$:

$$Y_i = \alpha_0 + \alpha_1 T_i + \mathbf{X}_i \beta + \varepsilon_i, \tag{1}$$

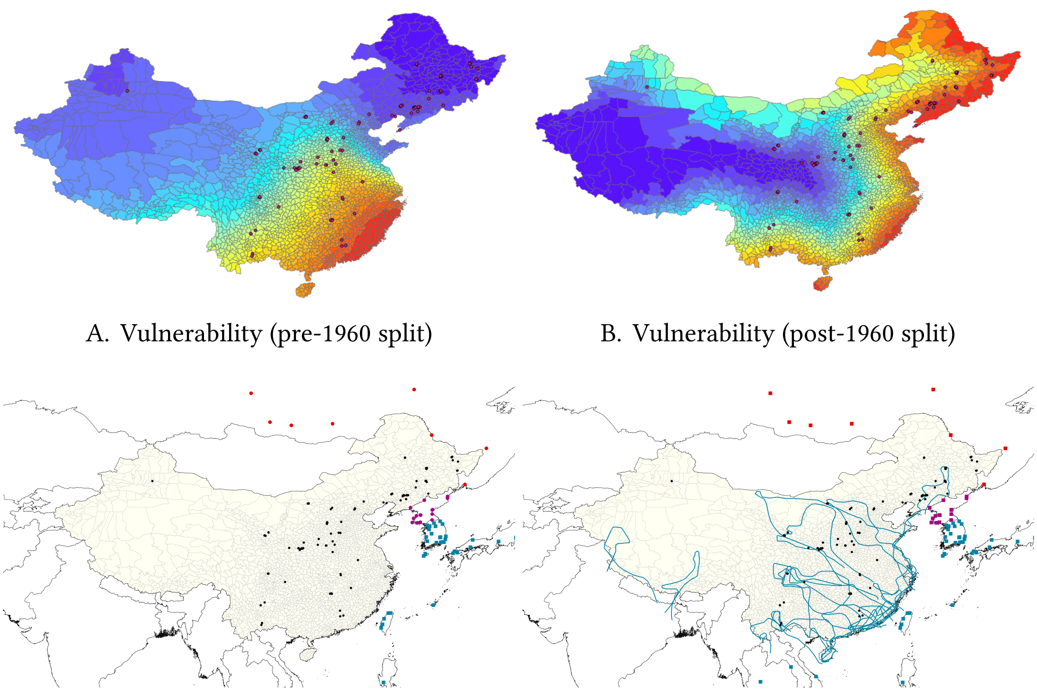
where: $Y_i$ is a measure of economic activity at the county level; the treatment, $T_i$, is defined as being in the close neighborhood of an MRP (within 25 kilometers); controls $\mathbf{X}_i$ include, among other variables which will be described below, the previous propensity factors $\mathbf{H}_i$ and a set of matched-pair dummies (stratifying the sample into 98 county pairs); and standard errors are clustered at the level of (34) projected cells with sides parametrized to be 400 kilometers.

In specification (1), the decision to allocate an MRP to a location $i$ might correlate with unobserved factors, and we will isolate exogenous variation in the siting decision using an instrument based on vulnerability to airstrikes from U.S. and Taiwanese air bases, while accounting for the shield provided by allied bases. To this end, we geolocate major U.S.S.R. and North Korean air bases, active U.S. Air Force bases, Taiwanese military airfields, and Chinese military air bases. In the baseline specification, we assume that (i) active U.S. and Taiwanese airfields are direct threats throughout the period; (ii) U.S.S.R. and North Korean air bases act as allied air bases with interception capabilities at the time of siting decisions, and as direct threats after the Sino-Soviet split; and (iii) Chinese military airfields have no credible interception capabilities. Under these assumptions, we compute a measure of local flying cost that accounts for the vicinity of U.S.S.R. and North Korean bases. The procedure, discussed in greater detail in Appendix C.2 and disciplined by the technical characteristics of enemy jet fighters at the time, produces a map of local flying costs for enemy airplanes covering any given point in Chinese territory.[15] Our instrument (vulnerability to aerial attacks in 1953) is defined, for each

[15]We relax the assumption that North Korean bases have interception or attacking capabilities before or after 1960 in a sensitivity analysis (see Appendix C.2, and Appendix D), and probe the robustness of our main findings to variations in the range of enemy jet fighters or interception capabilities.

county, as the minimum—across enemy bases—of total flying cost from the base to the county centroid. We illustrate the spatial variation in vulnerability to aerial attacks in Figure 2. Before the Sino-Soviet split, military concerns favored the Northeast at the expense of East China (see Panel A showing vulnerability in 1953). The set of suitable and protected locations, however, became much smaller after the split, and investment during the Third Front had to be targeted toward interior provinces (see Panel B showing vulnerability in 1964). The paths of surveillance flights between 1963 and 1965 (Panel D) provide external validation for the decision to shield the MRPs and the later Third Front factories. After the Sino-Soviet split, the "Second Front" lost its location advantage, and U.S. reconnaissance aircraft appeared to explicitly target these factories.

**Figure 2**. Vulnerability to airstrikes, bases, and surveillance flights.



A. Vulnerability (pre-1960 split) B. Vulnerability (post-1960 split)

C. Bases (pre-1960 split) D. Surveillance flights (post-1960 split)

Notes: Panels A and B represent the flying (travel) cost from enemy airfields (red: low, blue: high; the color gradient corresponds to deciles) in 1953 and in 1964. Panel C shows the distribution of enemy and allied air bases in 1953. U.S. air bases are indicated by green rectangles; North Korean air bases are indicated by purple dots; Soviet air bases are indicated by red dots; the locations of MRPs are indicated by dark dots. Panel D adds the paths of U.S. surveillance flights between 1963 and 1965.

Our empirical strategy uses the pre-split measure of county $i$'s vulnerability to aerial attacks at the time of the MRP siting decisions, $V_i$, as an instrument for siting decisions, $T_i$, in specification (1), conditioning on the post-split measure, thereby leveraging the ephemeral China-U.S.S.R. alliance as the source of identification. Our baseline set of controls $\mathbf{X}_i$ thus includes matching factors, matched-pair dummies, (log) travel cost to major ports to control for the later propensity to grow as part of an exporting economy, and (log) distance to Chinese military airfields and the post-split vulnerability to air strikes, in order to control for later military factors affecting the strategic placing of future industrial capabilities.

**Identification assumption(s)** A condition for identifying the treatment effect $\alpha_1$ in specification (1) is that the vulnerability to aerial attacks should predict the allocation of treatment. This relationship forms the first stage of our empirical specification, and Table 1 shows that vulnerability to aerial attacks is quantitatively important in explaining siting decisions.[16] A one-standard deviation increase in flying cost from enemy bases rises the propensity to host MRPs by about 37 percentage points within a matched treated-control pair. The average difference in flying cost between treated and control counties is about 90% of a standard deviation; our instrument thus accounts for about $90\% \times 0.37 \approx 33\%$ of the allocation of MRPs among suitable counties.[17] Importantly, while there is a strong relationship between our treatment—a place-based policy between 1953 and 1958—and the vulnerability to aerial attacks in 1953, we show in Appendix D.1 that the treatment is only weakly correlated with vulnerability measures computed after the Sino-Soviet split. To ensure that the geography of our place-based policy does not simply capture the geography of later investments, all our specifications will control for vulnerability in 1964.

The key identifying assumption underlying the empirical strategy however, is that the instrument affects outcomes of interest only through the location of the Million-Rouble plants, conditional on the set of controls $\mathbf{X}_i$. We now discuss supporting evidence and possible concerns with this assumption. First, we provide a more explicit illustration of the *residual* TSLS variation in Appendix C.2, illustrating more clearly the nature of the excluded variation used for identification and formalizing the comparison between Panels A and B of Figure 2 across suitable counties $\{t_1, \dots, t_N, c_1, \dots, c_N\}$ and for alternative modeling choices. Second, the respective locations of military bases could have influenced investment at later stages of China's structural transformation. Conditioning on later vulnerability to aerial attacks—after the Sino-Soviet Split and the start of the Vietnam War re-balanced the geographic distribution of military power in the region—should reduce this concern. We also provide sensitivity analyss and control

[16] Appendix C.2 provides a visual representation of the relationship between the vulnerability to aerial attacks and factory location choices.

[17] Table 1 displays three specifications, one without any controls (column 1), one with matched pair fixed effects only (column 2), and one with the full set of controls (baseline specification, column 3). All specifications are restricted to the set of treated and control counties defined by matching on access to natural resources, market access, and pre-treatment industrial facilities ($\mathbf{H}_i$). The *extended* set of controls is used to condition the analysis on characteristics that may directly affect outcomes of interest in the second stage; it is however reassuring that the predictive power of our instrument is not heavily dependent on their inclusion.

**Table 1**. Treatment and penalized distance to enemy airbases (1953).

| Treatment | (1) | (2) | (3) |
|---|---|---|---|
| Penalized distance | 0.197 | 0.342 | 0.375 |
| | (0.040) | (0.050) | (0.042) |
| Observations | 196 | 196 | 196 |
| Matching-pair fixed effects | No | Yes | Yes |
| Extended controls | No | No | Yes |

Notes: Standard errors are clustered at the level of 400 × 400-kilometer cells. The unit of observation is a county. *Penalized distance* is the normalized travel cost from U.S. and Taiwanese military airfields penalized by proximity to U.S.S.R. and North Korean airfields ($V_i$). Extended controls include all matching controls, i.e., (log) travel cost to the provincial capital, (log) population in 1953, (log) county area, (log) travel cost to resources (coal, coke, ore), (log) number of industrial firms over the period 1840–1937, and additional controls, i.e., matching-pair fixed effects, (log) travel cost to major ports through the river network, (log) distance to military airfields, and penalized flying cost to enemy airfields in 1964.

for other policies with spatial impacts (e.g., Third Front Movement, Cultural Revolution, Special Economic Zones and industrial parks, etc.). Third, vulnerability may correlate with county amenities, which could explain both the decision to locate factories and later patterns of economic growth. We will control for elevation, ruggedness, soil quality, expected crop yield, and pollution in robustness checks. Fourth, vulnerability may correlate with the general geography of recent growth in China. For instance, China's Southeast, considered vulnerable, widely benefited from the opening of Chinese ports to trade in the reform era, at the expense of the Chinese Rust Belt in the North. Such a violation of the exclusion restriction could induce a spurious negative correlation between economic growth and the presence of MRPs. To deal with this concern, we will run a series of robustness checks, most notably excluding counties within a buffer of the Pearl River Delta, excluding all Chinese counties south of a certain latitude, controlling more carefully for or matching on pre-treatment industrial output to account for the early industrialization of Northern provinces, conditioning on the secular dynamics of public enterprises (disproportionately affecting the Chinese Rust Belt), or controlling for/matching on travel time to the coast and to the main trading ports.

### 2.3 Descriptive statistics

The MRPs expanded and modernized the Chinese industry across a wide range of sectors, though with a bias toward heavy, extractive, and energy industries (e.g., coal mining and power plants—see Appendix A.2 for more details on the characteristics of the MRPs at the time of the "156" program).[18] For most MRPs, construction started between 1953 and 1955, and was achieved by 1959–1960, although a subset of plants was only completed after the Sino-Soviet split. The MRPs constituted very large investments for an agrarian economy such as China in the 1950s: The average planned investment by factory was 130,000,000 yuan, equivalent to 20,000,000 Soviet roubles in 1957 ($160,000,000 in 2010 U.S. dollars); total investment was about one-fourth of annual production in 1955.

Our identification strategy does not rely on a direct comparison of treated and control counties; instead, identification comes from isolating exogenous variation in the allocation of treatment from the transitory vulnerability to enemy airstrikes between 1953 and 1960. However, they constitute our baseline sample of interest, and we provide key descriptive statistics for the treatment and control groups in Table 2. About 5% of Chinese counties are defined as being treated, and another 5% serve as suitable control counties in our baseline specification. As expected from a context of heightened international tensions following the Korean War, treated counties are located much farther from U.S.A.F. and Taiwanese bases. The difference in vulnerability between control and treated counties is about 90% of a standard deviation. However, control and treated counties differ less markedly in their exposure to enemy raids *after* the Sino-Soviet split (see Panel A).

These two groups differ in terms of post-treatment population (Panel B) and in industrial development (Panel C). Treated counties are more populated in 1982 and in 2010; and the employment share in industry is much larger in treated counties in 1982, but this difference reversed by 2010.

While treated and control counties are mechanically similar in terms of matching variables (Panel D), treated counties tend to have better access to coal mines, but not to coke or ore deposits. Treated and control counties are, however, not selected to be similar in terms of travel time to major trading ports or to military airfields, at least in the baseline specification, and treated counties indeed tend to be farther from ports than control counties (Panel E)—justifying the inclusion of those variables as controls in specification (1). Finally, the unconditional allocation of treatment appears somewhat correlated with later economic shocks and placed-based policies (Panel F), justifying a comprehensive sensitivity analysis of these confounding factors.

[18] The "156" program followed the "Russian model" of industrialization (Rosenstein-Rodan, 1943), with coordinated and large investments across industries to modernize agrarian economies. These upstream factories were expected to irrigate the economy downstream.

**Table 2**. Descriptive statistics (control and treated counties).

| | Mean | Std dev. | Treated | Control |
|---|---|---|---|---|
| *Panel A: Vulnerability to air strikes* | | | | |
| Penalized distance (1953) | -0.03 | 1.06 | 0.42 | -0.47 |
| Penalized distance (1964) | -0.27 | 0.89 | -0.12 | -0.42 |
| *Panel B: Population* | | | | |
| Population (1982, log) | 12.96 | 0.89 | 13.05 | 12.87 |
| Population (2010, log) | 13.45 | 0.90 | 13.54 | 13.36 |
| *Panel C: Economic development* | | | | |
| Employment share (ind., 1982) | 0.29 | 0.21 | 0.33 | 0.25 |
| Employment share (ind., 2010) | 0.30 | 0.15 | 0.28 | 0.32 |
| *Panel D: Matching controls* | | | | |
| Travel cost to prov. capital (log) | 13.29 | 2.67 | 13.20 | 13.38 |
| Population (1953, log) | 12.16 | 1.37 | 12.17 | 12.15 |
| County area (log) | 6.76 | 0.99 | 6.88 | 6.64 |
| Travel cost to coal mines (log) | 11.85 | 4.19 | 11.45 | 12.25 |
| Travel cost to coke (log) | 9.96 | 5.73 | 9.94 | 9.98 |
| Travel cost to ore (log) | 14.65 | 1.74 | 14.65 | 14.65 |
| Number of ind. firms (log, 1840-1937) | 0.92 | 1.21 | 0.85 | 0.99 |
| *Panel E: Additional controls* | | | | |
| Distance to military airfields (log) | 10.47 | 1.01 | 10.39 | 10.54 |
| Travel cost to major ports (log) | 13.23 | 2.42 | 13.57 | 12.88 |
| *Panel F: Later economic shocks* | | | | |
| Third Front Movement | 0.17 | 0.38 | 0.20 | 0.13 |
| Cultural Revolution (victims) | 0.03 | 0.04 | 0.04 | 0.03 |
| Industrial parks (1990–2010) | 0.34 | 0.36 | 0.31 | 0.37 |
| Observations | 196 | | 98 | 98 |

Notes: *Penalized distance* is the standardized flying cost ($V_i$, with mean 0 and variance 1 over all counties in China)—see Appendix C.2 for details. The variables in Panels B and C come from Population Censuses. The construction of matching variables (D) and additional controls (E) is explained in Appendix B.3. *Third Front Movement* is an indicator variable equal to 1 if the county is located in a Third-Front province (see Fan and Zou, 2021) and 0 otherwise. *Cultural Revolution (victims)* is the ratio of victims of the Cultural Revolution to the 1953 Census population (Walder and Su, 2003). *Industrial parks* refers to the number of industrial parks created in the county between 1990 and 2010 per 10,000 inhabitants (using population from the 1953 Census). We list and describe the variables used in the matching procedure, the baseline specification, and robustness checks in Appendix B.3 and Appendix C.1.

## 3 The rise-and-fall pattern

This section quantifies the implications of early industrialization in counties hosting MRPs, with an aggregate rise and fall driven by production outside the MRPs.

### 3.1 Baseline results

The influence of the MRPs on their local economies spans two periods: the rise and the fall, which we capture with cross-sectional analyses in 1982 and 2010. In our view, the period from 1953–1960 to 1982 is the treatment period—encompassing the initial investment and the later allocation of resources in a command economy. We use the year 1982 to define the end of this treatment period for two reasons: (i) it is, unfortunately, the earliest year for which county-level output can be observed (about 25 years after the initial investment); and (ii) it coincides with the end of the command economy, thus capturing the possible head start of industrialized counties before reforms of the non-agricultural sector were gradually introduced.

**The rise** We first describe empirical facts about the local treatment effect of MRPs in 1982; the analysis and the choice of outcomes are unfortunately limited by the availability of information at the county level. Table 3 shows OLS and TSLS estimates of the relationship between the presence of am MRP and population, the share of urban residents, output per capita, and the employment share in industry.

We find that industrial investment under the "156" program had a positive impact, albeit moderate, on population at the end of the treatment period. Treated counties are 11.5% more populated than control counties in 1982 (column 1, Panel B, $\exp(0.108) - 1 \approx 0.115$). The treatment effect on urbanization is large; the share of the population with non-agricultural household registration is about 30 percentage points higher in treated counties (column 2, Panel B). The impact of the MRPs thus shows a large reallocation of labor, which can be interpreted as evidence of structural transformation and urbanization. The higher share of urban residents is associated with a much higher output per capita and a higher industry share in the local economy (columns 3 and 4, Panel B). GDP per capita is 80% larger in treated counties, and the employment share in industry is 22 percentage points higher. The magnitude of these differences goes far beyond the mere output of the average MRP (see Appendix A.4), indicating that counties are richer and more developed—the effect on output per capita is slightly lower than the difference between the 25th percentile and 75th percentile of the control-group distribution.

A few remarks are in order. First, the TSLS estimates are larger than the OLS estimates, which could be explained by a standard attenuation bias in the OLS specification due to measurement error (e.g., we nest treatment based on its initial allocation in the 1950s, and a few MRPs were later displaced or closed) or by omitted variation. For instance, counties that were allocated MRPs could have different unobserved fundamentals: they might have had a lower potential for industrialization in the early period, consistent with planners redistributing industrialization across space at the time of the First Five-Year Plan and targeting agrarian regions under the "156" program (Bo, 1991), but they might have better first- and second-nature fundamentals that mattered for later economic growth. Second, the rise of our treated counties may have limited external

**Table 3**. Treatment effect on employment, output, and urbanization in 1982 and 2010.

| | Population | Urban reg. | GDP p.c. | Share industry |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *Panel A: OLS specification* | | | | |
| Treatment effect (1982) | 0.067 | 0.124 | 0.296 | 0.092 |
| | (0.093) | (0.040) | (0.102) | (0.031) |
| | [196] | [196] | [196] | [196] |
| Treatment effect (2010) | 0.125 | 0.156 | 0.100 | -0.013 |
| | (0.110) | (0.036) | (0.100) | (0.026) |
| | [196] | [196] | [196] | [196] |
| *Panel B: TSLS specification* | | | | |
| Treatment effect (1982) | 0.108 | 0.297 | 0.595 | 0.218 |
| | (0.113) | (0.035) | (0.116) | (0.028) |
| | [196] | [196] | [196] | [196] |
| | *77.891* | *77.891* | *77.891* | *77.891* |
| Treatment effect (2010) | 0.140 | 0.354 | -0.195 | -0.057 |
| | (0.102) | (0.035) | (0.094) | (0.032) |
| | [196] | [196] | [196] | [196] |
| | *77.891* | *77.891* | *77.891* | *77.891* |

Notes: Each cell is the outcome of a separate regression. Standard errors are clustered at the level of 400 × 400-kilometer cells (reported between parentheses). The unit of observation is a county (Administrative level 3); the number of observations is reported between square brackets. The instrument is the minimum distance to the U.S. and Taiwanese military airfields penalized by the proximity to U.S.S.R. and North Korean airfields; we report the first-stage F-statistics in italics. All specifications include (i) matching-pair fixed effects, (ii) matching controls, i.e., (log) travel cost to the provincial capital, (log) population in 1953, (log) county area, (log) travel cost to resources (coal, coke, ore), (log) number of industrial firms over the period 1840–1937, and (iii) the additional controls, i.e., (log) travel cost to major ports through the river network, (log) distance to military airfields, and penalized flying cost to enemy airfields in 1964. *Population* is the logarithm of total population in the county, *Urban reg.* is the share of the population that has a non-agricultural household registration (*hukou*), *GDP p.c.* is the logarithm of GDP per capita within the county, and *Share industry* is the industrial employment share.

validity. Before the advent of the reforms, the government instructed workers where to live and where to work in order to accommodate rising demand for labor and control the growth of the plants and of the local economy.[19] Changes in labor allocation mostly reflected government intervention, which may temper agglomeration effects. The population increase, while larger than the expected labor force of the MRP itself, probably lagged behind labor demand in treated counties.

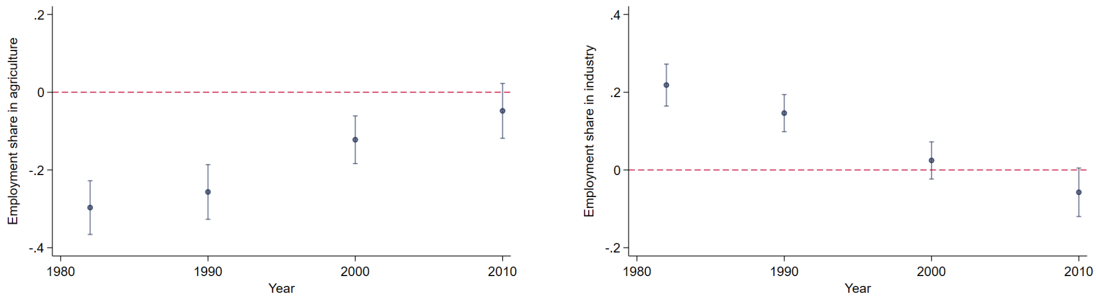
To summarize the impact of MRPs between 1953 and 1982, we find a moderate effect on urban population, and a *very* large effect on output. The substantial productivity gap between treated and control counties indicates that treated areas entered the sub-

[19]While some free movement of labor still occurred after the advent of "New China" in 1949, mobility was subject to authorization from the late 1950s onward. The government tightened its grip on labor movement in the wake of the Great Leap Forward, when famines threatened the sustainability of urban food provision systems.

sequent period with a substantial head start. The liberalization of the economy should have allowed agglomeration economies to operate, and one could have expected treated counties to drift further apart from the rest of the economy. As we see next, we find the opposite.

**The fall** There is (more than) a full catch-up between 1982 and 2010—see Table 3. We find that population remains higher in treated counties (column 1, Panel B); treated locations also continue to have a higher share of registered urban population (column 2, Panel B). In stark contrast to the treatment effect in 1982, however, output per capita and the industry share are now lower in treated counties (columns 3 and 4, Panel B). The significant gap in industrialization before the transition has thus fully eroded: treated counties are about 20% less productive than control counties, and the employment share in industry is 6 percentage points smaller.

**Figure 3**. The dynamics of counties hosting MRPs.




A. Employment share in agriculture B. Employment share in industry

Notes: Panels A and B display the treatment effect for the employment share in agriculture and industry (1982, 1990, 2000, 2010).

We provide further insights into the dynamics of treated counties from 1982 onward in Figure 3 and in Appendix E.1. More specifically, we estimate Equation (1) separately for different years and report the treatment effect on employment shares in agriculture and industry (1982–2010).[20] We find that there is a gradual decrease in economic activity in treated counties, with a clearer inflection during the 1990s.

This fast reversion to the mean is puzzling for three reasons. First, it does not result from a swift decline of MRPs themselves; they remain very large and productive.[21]

[20]The choice of outcomes in Figure 3 stems from the lack of coherent data that can be consistently harmonized over time and to cover the dynamics of *counties* between 1982 and 2010. In Appendix E.1, we also provide some insight into: more recent production dynamics, around 1998–2010; nighttime luminosity (1992–2012); and urban "extent" as identified from satellite imagery (1992–2012).

[21]The dynamics of the local economy could be affected by the dynamics of the MRPs themselves, as they are large employers within a county—but by no means monopsonists. In Appendix A.4, we document

Second, their own, *direct* influence on aggregate productivity is non-negligible and positive, and we will see that numerous connected firms benefit from the presence of MRPs (as separately documented in Giorcelli and Li, 2022). What the previous results indicate is that the other, unconnected firms—still constituting the lion's share of a county's output—must be very unproductive. Third, we document limited treatment effect heterogeneity across MRP characteristics in Appendix E.3, despite significant differences in the nature of their production or in their "initial completeness" (Giorcelli and Li, 2022). We find that upstream MRP(s) exert more negative effects on innovation and productivity than downstream MRP(s), and counties associated with low-investment or closed MRP(s) tend to perform, if anything, slightly better in the long run.

Before turning to the characterization of these other establishments in treated counties, we provide a series of robustness checks.

### 3.2 Robustness checks and sensitivity analysis

Our empirical strategy exploits temporary, exogenous variation in the probability of hosting an MRP among suitable counties. The geographical variation induced by vulnerability to bombing between 1950 and 1960 may, however, coincidentally correlate with other geographic determinants of later economic growth. We provide a comprehensive sensitivity analysis to reduce concerns that the observed rise and fall are driven by factors other than hosting MRPs. This section summarizes these robustness checks; a detailed discussion of the results is provided in Appendix D.

We interpret the previous estimates as the effect of MRPs on the local economy. One concern is that the instrument, which relies on the distribution of air bases across space, may correlate with other geographic factors that have independent effects on the distribution of economic activity across China over time. We specifically address this issue in Appendix D.2. First, we condition the baseline specification on measures capturing an environment generally (un)favorable to economic take-off, e.g., land supply constraints, pollution, connectedness, travel cost to the coast, elevation, land-form characteristics, and potential and actual crop yields, to reduce concerns about bias from unfavorable county characteristics (Table D3). Second, the geographic nature of our instrument im-

the size of MRPs and describe the dynamics of production in the average MRP. More specifically, we show that (i) they experience an *increase* in patenting activity and stable employment throughout the recent period; and (ii) their production share in local economies is non-negligible (around 3% of the wage bill and 2% of value added within their county) but insufficient to explain our findings without spillovers to the rest of local production. Further, we provide a sensitivity analysis in Appendix D.2 where we exclude the few counties hosting closed and displaced factories to demonstrate that the fall is not related to the fall of MRPs themselves. Finally, we examine treatment heterogeneity in Appendix E.3 where we show that the fall of treated counties is quite homogeneous across the different MRP types. Interestingly, while we find little heterogeneity, counties that received lower investments or whose MRP closed down prematurely experienced a smaller rise, but also a smaller fall.

plies that the residual TSLS variation partly exploits counties located in a "second front," stretching from the Northeast region to the interior provinces of Gansu/Shaanxi (see Appendix C.2 for an explicit illustration). For this reason, we probe the robustness of our findings to the possible confounding roles of other, secular dynamics affecting (i) the Chinese Rust Belt (mostly in the Northeast, closer to our treated counties) or, conversely, (ii) the Southern coastal provinces (possibly closer to our control counties): (i) we further account for the pre-treatment industrialization of Northern provinces, with alternative matching procedures that include and condition on initial industrial output (Table D5), and we condition the analysis on the more recent demise of State-Owned Enterprises (Table D3); and (ii) we show that our results are not driven by the unequal geography of the economic take-off in China by matching or conditioning on access to the main trading ports or to the coast, controlling for exposure to the reduction in trade uncertainty following China's accession to the World Trade Organization in 2002, and excluding all coastal provinces, the Pearl River Delta, or the South of China (Tables D3 and D5). Third, we control in Table D4: for later place-based policies that could reduce the gap between treated and control counties (e.g., Third Front Movement, Special Economic Zones); for factors underlying these later decisions (e.g., vulnerability to U.S.S.R. strikes after the rapprochement between the U.S. and China in 1972); for other policies driving the economic evolution of China (e.g., the Great Leap Forward, the Great Famine, or the Cultural Revolution); and for possible direct negative impacts of the few liquidated MRPs on local employment.

Finally, we consider a more general sensitivity analysis in Appendix D.3/D.4 and variations in: the matching strategy, e.g., by changing the set of variables used for matching or the process to select control counties; the treatment of spatial correlation; the construction of vulnerability; and the measurement of economic development.

### 3.3 Production in *other* establishments

Treated counties experience a swift reversion to the mean in *aggregate* output, which includes the production of MRP(s). Production in *other* establishments must therefore be limited enough to account for this county-level reversion pattern. We now rely on micro-data to better characterize the structure of production in those non-MRP establishments.

Our analysis relies on: measures of factor productivity at the establishment level, with a production function identified using an exogenous labor supply shifter (see Imbert et al., 2022, and Appendix B.5); patents linked to establishments (He et al., 2018); and markups (computed following De Loecker and Warzynski, 2012, see Appendix B.6).[22] In

[22]In this section, we describe the treatment effect on the structure of production using a selection of outcomes, and we leave the detailed analysis of factor use, factor productivity, firm characteristics, invest-

Table 4, we estimate Equation (1) at the establishment level, pool all establishment × year observations between 1998 and 2007, and regress an outcome on the treatment $T_i$, instrumented by $V_i$. We exclude the MRPs and their descendants from the sample and we cluster standard errors at the level of 400 × 400-kilometer cells. Differences in firm outcomes across counties could theoretically be tied to composition differences induced by local industrial structure, products, the presence of public and subsidized enterprises (Harrison et al., 2019), and the age of establishments. We thus add year indicators interacted with: 4-digit industry fixed effects; 6-digit product fixed effects; firm ownership type (e.g., "State-owned enterprise", "Collective enterprise", "Private limited liability company", "Foreign-Chinese joint venture", "Wholly foreign enterprise"); and age.[23]

**Table 4**. Structure of firm production in the average other establishment.

| | Labor<br>(1) | Capital<br>(2) | Wage<br>(3) | TFP<br>(4) | Patents<br>(5) | Markup<br>(6) |
|---|---|---|---|---|---|---|
| Treatment | 0.168<br>(0.049) | 0.205<br>(0.054) | -0.256<br>(0.033) | -0.257<br>(0.063) | -0.051<br>(0.013) | 0.111<br>(0.034) |
| Observations | 348,632 | 348,632 | 348,632 | 348,632 | 348,632 | 223,341 |

Notes: Standard errors are clustered at the level of 400 × 400-kilometer cells. The unit of observation is an establishment × year. We exclude the MRPs and their descendants from the sample. All specifications include: the extended controls of Table 3 (including matching-pair fixed effects interacted with year fixed effects); 4-digit industry × year fixed effects; 6-digit product × year fixed effects; firm type × year fixed effects; and age × year fixed effects. *Labor* is the logarithm of the number of workers; *Capital* is the logarithm of real capital; *Labor cost* is the logarithm of total compensation per employee; *TFP* is the logarithm of firm-specific total factor productivity as computed in Imbert et al. (2022); *Patents* is the number of patent applications registered by the firm; *Markup* is a dummy equal to one if the markup is above the median within a 4-digit industry × year cell, and markups are computed following De Loecker and Warzynski (2012)—see Appendix B.5 and Appendix B.6.

We find that establishments in treated counties are larger and slightly more capital-intensive than those in control counties (columns 1 and 2). Labor cost and total factor productivity are much lower in treated counties (around 25%, columns 3 and 4). This could indicate either that the treatment generates differences in technology or that it generates differences in price setting between control and treated counties. We investigate these two aspects next. We find that establishments in treated counties produce far fewer patents. The treatment effect on patents amounts to 30% of the yearly number of

ment and subsidies, patenting behavior, and price setting (and their evolution over time) to Appendix E.1.

[23]We study the role of entry, exit, and firm composition in a dedicated Appendix E.2. In particular, we analyze how differences in production structure may reflect differences in the ownership structure, or the presence of establishments at different stages of their life cycles, and we provide a sensitivity analysis for Table 4 with more limited sets of "composition" controls. Controlling for the local industry structure is only quantitatively important for patenting behavior. Indeed, the presence of MRP(s) tilts the local industrial fabric toward innovative and competitive sectors; these innovative sectors are, however, far less innovative and competitive in treated counties. In general, we find that controlling for the life cycles of local firms and for the demise of public establishments is quite innocuous. Public establishments in treated counties are not less innovative or less productive than in control counties, relative to other firms.

patents produced in the average establishment: very few patents are thus registered in treated counties. Finally, the productivity effect cannot be explained by markups; they are on average higher in treated counties (column 6): the probability for a firm to charge a markup above the median is 11 percentage points higher in treated counties.

The next section identifies the spillovers within the local economy that drive this marked decline in innovation, productivity, and competition.

## 4 Mechanisms behind the fall of treated counties

This section analyzes the negative spillovers exerted by MRP(s) on local production. In the first step, we zoom in and out of the production chain of MRPs, characterize the structure of local production, and examine the treatment effect of MRP(s) on closely-linked establishments versus average establishments within the same county. In a second step, we identify the effect of hosting an MRP on *missing industries* through the local supply of entrepreneurs. In a final step, we discuss external validity and the role of other externalities than production and technological spillovers, e.g., those operating through the provision of factors or political favoritism.

### 4.1 Production structure around MRPs

We first identify linkages between establishments to better understand concentration *around the MRP(s)* and the characteristics of linked plants.

**Product concentration around MRPs** Establishment-level data allow us to observe potential links between establishments and local MRP(s) and thus to identify differences in the local structure of production. For instance, one may benchmark the activity of downstream establishments in a treated county against similarly defined establishments in control counties. A difference-in-difference specification, however, cannot be implemented as such, due to treatment heterogeneity, and we rationalize the use of a slightly more involved empirical strategy in Appendix C.3, where we (counterfactually) allocate MRP(s) from their matched treated county to each control county.

In order to capture potential linkages between establishments and local MRPs, we combine U.S. input-output accounts (Stewart et al., 2007) with our classification of manufacturing establishments by product codes (Imbert et al., 2022). The idea is to measure proximity between an establishment and the local MRP(s) through their products. Let $m_{q,p}$ denote the input share of product $q$ in product $p$, and let $\Theta$ be the set of products produced by MRP(s) in a county. For each establishment producing $p$ in a treated-control pair $\{t_i, c_i\}$, we compute $\max_{P\in\Theta(t_i)}\{m_{P,p}\}$, i.e., the largest input share that could be supplied by the MRPs of county $t_i$. We then define an indicator, *Downstream*, as equal to 1 if this value lies above the 95th percentile across establishments, identifying those most likely to be downstream of local MRPs. By analogy, *Upstream* equals 1 if $\max_{P \in \Theta(t_i)}\{m_{p,P}\}$ lies above the 95th percentile. We further define *Same industry*, as equal to 1 if the establishment operates in the same 4-digit industry as one of the MRPs. In addition, we construct two broader measures: (i) *Technological closeness*, based on the 95th percentile of patent citation intensities across industries (Bloom et al., 2013); and (ii) competition for factor markets, based on revealed factor intensities predicted by trade patterns in 2000 (see Shirotori et al., 2010).[24]

**Table 5**. Production linkages with the MRPs.

| | Downstream | Upstream | Same industry |
|---|---|---|---|
| *Panel A: Production linkages* | (1) | (2) | (3) |
| Treatment | 0.021 | 0.023 | 0.021 |
| | (0.017) | (0.010) | (0.008) |
| | [348,506] | [348,506] | [348,506] |
| | H-intensity | K-intensity | T-intensity |
| *Panel B: Factor demand* | (1) | (2) | (3) |
| Treatment | 0.003 | 0.011 | -0.017 |
| | (0.011) | (0.013) | (0.011) |
| | [257,063] | [257,063] | [257,063] |
| | Tech. clos. | Tech. clos. (Mah.) | |
| *Panel C: Technology closeness* | (1) | (2) | |
| Treatment | 0.019 | 0.013 | |
| | (0.007) | (0.006) | |
| | [348,506] | [348,506] | |

Notes: The unit of observation is a firm × year, we exclude the MRPs from the sample, standard errors are clustered at the level of 400 × 400-kilometer cells, and all specifications are weighted by output. All specifications include the baseline controls of Table 4. *Downstream* (resp. *Upstream*) is a dummy equal to one if the firm is down (resp. up) the supply chain with respect to one of the MRPs; *Same industry* is a dummy equal to one if the firm is in the same industry as one of the MRPs. *F-intensity* is a dummy equal to 1 if the revealed factor intensity of factor *F* (using product codes) is higher than that of the average associated MRP. *Tech. clos.* (*Tech. clos. (Mah.)*) is a dummy equal to one if sectors in which the establishment and the MRP(s) operate are linked through patent applications (based on Mahalanobis distance).

Table 5 (Panel A) reports the relative presence of establishments operating downstream from, upstream from, or in the same industry as the local MRP(s). We report

[24] For factor markets, we define a dummy, *More F-intensive*, as equal to 1 for factor *F* (human capital, physical capital, or land) if the average factor intensity of an establishment's goods exceeds that of local MRPs. The rationale is that MRPs, with lower search frictions and privileged resource access, may exert stronger bargaining power, particularly affecting firms with high demand for the same factor.

a measure capturing the probability that a unit of *output* is tied to MRPs and discuss the probability that a *worker* is tied to MRPs in Appendix E.4. In column (1), we report the result of a specification in which the measure of downstream linkages at the establishment level is regressed on the treatment, instrumented by vulnerability to air strikes (with the same controls as in Table 4). We find that treatment increases the probability that an establishment operates downstream of the MRP by about 2.1 percentage points. Columns (2) and (3) of Table 5 report the relative incidence of upstream linkages and horizontal linkages in treated counties. The treatment effect on the probability for an establishment to operate upstream of the MRP is non-negligible (2.3 percentage points, to be compared with an average of 5% across treated and control counties), even though MRPs tend to operate early in the production chain. The treatment also affects the probability to operate in the same 4-digit industry, which increases by 2.1 percentage points—an effect that might be attributed to economies of scale (Ciccone, 2002). Overall, while the hypothetical production chain of MRP(s), excluding the MRP(s), would represent about 6.2% of output in control counties (accounting for the fact that the previous categories are not exclusive), the actual production chain accounts for almost 10% of output in treated counties.

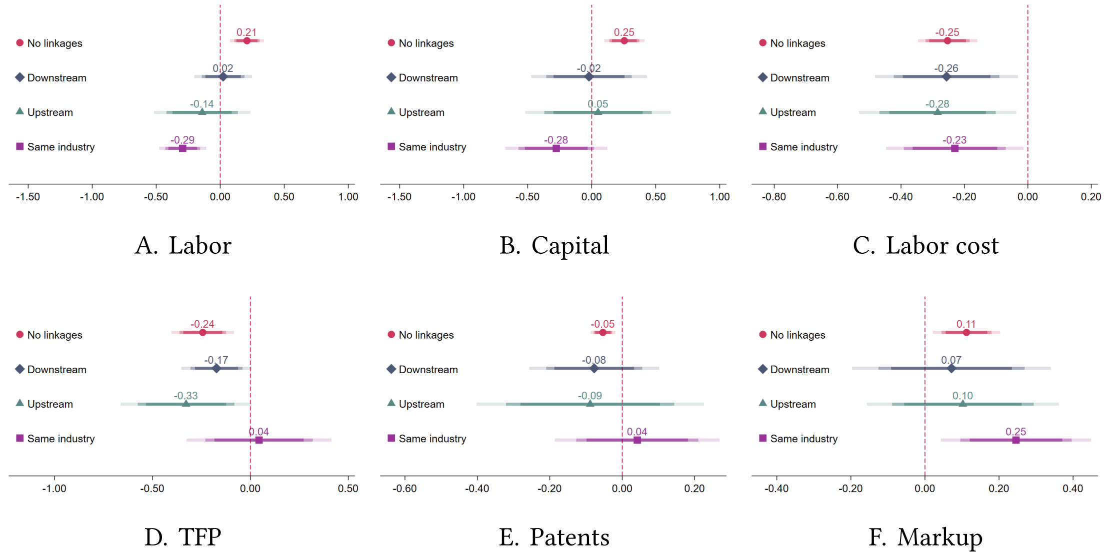
Panel B of Table 5 reports the relative presence of establishments with more acute demand for human capital (column 1), physical capital (column 2), and land (column 3) than the local MRP(s). The differences between treated and control counties are not very large—a few percentage points, to be compared with averages at around 50%. These findings provide little support for the existence of spillovers in factor markets. Finally, Panel C reports the relative presence of establishments with a technology closeness measure above the 95the percentile. The difference between treated and control counties is positive, albeit smaller than for production linkages.

**Production in linked and non-linked establishments** The previous evidence has identified the change in the structure of production induced by the presence of the MRP(s): there are many establishments operating along their production chains. We now characterize these establishments by looking at treatment heterogeneity on the following selected outcomes: factor use (labor and capital), labor cost, total factor productivity, the number of registered patents, and markups.

Figure 4 reports these estimates for non-linked (red circle), downstream (blue diamond), upstream establishments (green triangle), and establishments within the same 2-digit industry as MRPs (purple square).[25] Treatment heterogeneity is limited, yet not

[25] We use a dichotomy based on 2-digit industries in this exercise, as there are too few establishments in control counties within the same 4-digit industries as hypothetical MRPs.

**Figure 4**. Productivity, innovation, and pricing in establishments along MRPs' production chains.




A. Labor B. Capital C. Labor cost

D. TFP E. Patents F. Markup

Note: This Figure displays the treatment estimates for the following outcomes: *Labor* is the logarithm of the number of workers; *Capital* is the logarithm of real capital; *Labor cost* is the logarithm of total compensation per employee; *TFP* is the logarithm of firm-specific total factor productivity as computed in Imbert et al. (2022); *Patents* is the number of patent applications registered by the firm; *Markup* is a dummy equal to one if the markup is above-median within a 4-digit industry × year cell, computed following De Loecker and Warzynski (2012). The unit of observation is a firm × year, we exclude the MRPs from the sample, and standard errors are clustered at the level of 400 × 400-kilometer cells. All specifications include the extended controls of Table 4. In each panel, we report four estimates: one obtained using the sample of non-linked firms (*No linkages*); one obtained using the sample of downstream firms (*Downstream*); one obtained using the sample of upstream firms (*Upstream*); and one obtained using the sample of firms within the same 2-digit industry as MRPs (*Same industry*).

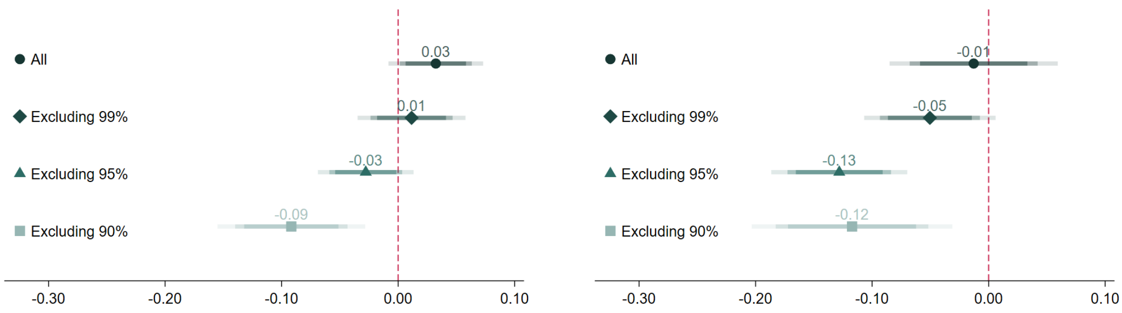
negligible. First, the average non-linked establishment is larger, less productive, less innovative, charges higher markups, and faces lower labor costs than in control counties. The spillovers exerted by MRPs thus seem to extend beyond their production chains—an effect that we further investigate in Section 4.2, where we look at (missing) entrepreneurship. Second, the treatment effects on linked establishments differ slightly from the average effect: for instance, the average downstream or upstream establishment is not larger than in control counties, while same-industry establishments are smaller. Downstream and upstream firms tend to be less productive and innovative in treated counties—an effect that is absent for same-industry establishments.[26] By contrast, all establishments charge high markups.[27]

[26]The less negative productivity and patenting effects for same-industry establishments could be due to *technological* linkages shared with the local MRP(s). While the present exercise focuses on "production" linkages, i.e., on establishments that operate through and close to the production chain of local MRPs, we expand this exercise to technological linkages and other firm characteristics (public, subsidized, young) in Appendix E.2. We find that establishments with direct technological linkages to the MRPs appear to benefit from the presence of the large plant(s), although the effects are estimated with noise: for instance, they are not unproductive and produce many patents. We do not see, however, qualitative differences in treatment effects along the other firm characteristics, i.e., when focusing on public firms, subsidized firms, or young establishments.

[27]The effect of a large, innovative establishment on the markup set by intermediaries could be ambigu-

In general, establishments along the production chain of MRP(s) seem to enjoy a rent from their proximity to a highly productive and innovative factory; these linked establishments extract part of the final value added when operating at one point of the production chain, whether upstream or downstream, and do not need to incur innovation efforts. This rent might provide incentives for establishments and entrants to tie their production to MRP technology, thereby explaining the cluster of specialized production units around the MRP.

**Figure 5**. Production clustering in treated/control counties.




A. Herfindahl index B. IO-weighted Herfindahl index

Note: This Figure displays the treatment estimates for the following outcomes: *Herfindahl index* is a Herfindahl index based on output shares within 6-digit product categories computed from the firms of our sample in 2004 and nested at the county-level (panel A); *IO-weighted Herfindahl index* is a Herfindahl index based on output shares within 6-digit product categories, mediated by input shares from the U.S. IO matrix (panel B). In each specification, the unit of observation is a county, standard errors are clustered at the level of 400 × 400-kilometer cells, and we include the extended controls of Table 3 and measures of the (log) number of firms at the county-level and (log) total output by these firms—to control for a possible mechanical effect of market size on production clustering.

**Production concentration in treated/control counties** The previous evidence documents a grouping of production units through the production chain of MRPs. This expected result, however, masks a less intuitive picture of production concentration across counties. We now shed light on this general production structure by extracting county-level Herfindahl indices of production concentration (see Appendix B.4 for a detailed description of these indices).

More specifically, we construct a county-specific vector $\mathbf{S}_i = \left(s_{pi}\right)_p$ of output shares across 6-digit product categories indexed by $p$. An common measure of product concentration is the Herfindahl-Hirschman Index, $h_i = \sum_p s_{pi}^2$ (hereafter "Herfindahl index").

ous: the production chain probably generates high rents and there may be a hold-up problem if the final good requires key intermediary inputs to be produced; however, the large MRPs may benefit from a more advantageous bargaining position when negotiating with intermediaries.

This standard index can be written as:

$$h_i(\mathbf{I}) = \mathbf{S}_i'\mathbf{I}\mathbf{S}_i,$$

where $\mathbf{I}$ is the identity matrix. However, this index does not capture product concentration through production chains: A pair of establishments would be linked—with intensity 1—if and only if they operate in the exact same 6-digit product market. To better account for production linkages, we further rely on the input share of product $q$ into product $p$, $m_{q,p}$, and the associated square matrix $\mathbf{M} = \left(m_{q,p}\right)_{q,p}$. An "IO-weighted Herfindahl index" can be written as:

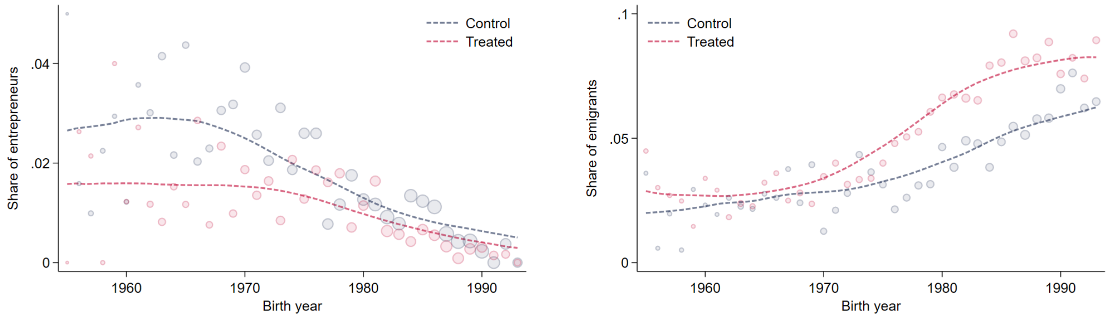
$$h_i(\mathbf{M}) = \mathbf{S}_i'\mathbf{M}\mathbf{S}_i.$$

We estimate the treatment effect on these product concentration indices by considering the county-level specification (1), and we report the estimates in Figure 5 for two outcomes: the standard "Herfindahl index", $h_i(\mathbf{I})$ (panel A); and the "IO-weighted Herfindahl index", $h_i(\mathbf{M})$ (panel B). The first estimate relies on all product codes in each county and shows a slightly positive difference between treated and control counties—a difference that fades away for the IO-weighted Herfindahl index, despite the previously-documented inflated production chain of MRPs in treated locations. The second estimate drops the top-1% county/product-code output shares from the calculation of the "Herfindahl index"; the third drops the top 5% and the last drops the top decile of output shares. These latter estimates show that product concentration is (much) lower in treated counties once we exclude the largest products, and the differential is even more pronounced when accounting for production linkages. For instance, the differential in product concentration between treated and control counties of -0.12/-0.13 in product concentration indices (as shown for the last two estimates of panel B) is equivalent to more than half of a standard deviation: outside the production chain of MRPs,—which is typically dropped from those selected product samples,—there are fewer product clusters in treated counties and production is largely scattered across smaller production chains.[28]

With production mostly concentrated around MRPs and a non-innovative nucleus of firms, treated counties do not benefit much from externalities in local technological progress, whether within or across industries (Glaeser et al., 1992). One question remains: why are there so few clusters of establishments outside the production chain of MRPs and why are they less productive, innovative, and competitive?

[28]We provide further robustness checks in Appendix E.4, where we use other matrices based on (i) technological linkages or (ii) product similarity based on language similarity.

**Figure 6**. The missing entrepreneurs.




A. Top managers B. Emigrants

Note: Panel A displays the share of individuals that are in top managerial positions among working-age individuals—between 22 and 60 years old—with a college education in control counties (blue dashed line) and in treated counties (red dashed line). Treatment is defined at the location of residence. The lines are local polynomial approximations as a function of the year of birth. Panel B displays the share of emigrants among working-age individuals—between 22 and 60 years old—with a college education in control counties (blue dashed line) and in treated counties (red dashed line). Treatment is then defined at the location of registration, and emigrants need to live outside of their *province* of origin.

### 4.2 The missing entrepreneurs

So far, we have shown that treated counties have a nucleus of establishments possibly tied to MRPs, but (i) their own contribution to local production and innovation is limited; (ii) other productive production chains appear to be missing; and (iii) the average other establishment outside the MRPs' production chains is not productive, not innovative, and not competitive. One explanation is that treatment directly affects the local production of entrepreneurs or their retention. We rely on the 1% Population Survey of 2015 to assess the effect of treatment on the production of prospective entrepreneurs and their selection into emigration. As a preliminary illustration, Figure 6 shows that the cohorts of college graduates born between 1955 and 1975—thus at the age to occupy top entrepreneurial positions in 2015—are much less likely to do so in treated counties compared to control counties (Panel A).[29] While potential entrepreneurs might be absorbed into MRPs, prospective managers also emigrate and the retention of college graduates is much lower in treated counties (Panel B).

Table 6 quantifies the extent of missing entrepreneurs within our preferred TSLS specification. We find that the probability that college-educated individuals are unem-

[29]In this exercise, we consider working-age individuals—between 22 and 60 years old—with a college education, and we define "entrepreneurs" as individuals in top managerial positions (i.e., the chief or deputy manager of an enterprise) based on the textual description of occupations. We provide some evidence in Appendix E.5 that our findings are not explained by the local production of elites or by the capture of those elites by the public sector: the incidence of college education is higher in treated counties, and these college-educated workers are not significantly more likely to be employed in the public sector (i.e., education, local government, and health). Appendix E.5 also sheds light on the treatment effect on local entrepreneurial values using survey data (e.g., the expected return to "talent").

**Table 6**. The missing entrepreneurs.

| | Unemployed (1) | Entrepreneur (2) | Manager (3) | Emigrant (4) | E-manager (5) |
|---|---|---|---|---|---|
| Treatment | 0.056 | -0.006 | -0.029 | 0.022 | 0.089 |
| | (0.007) | (0.003) | (0.008) | (0.009) | (0.021) |
| Observations | 59,831 | 43,052 | 42,553 | 55,015 | 6,142 |
| Mean | *0.0539* | *0.0140* | *0.1561* | *0.0519* | *0.0602* |

Notes: Standard errors, clustered at the level of 400 × 400-kilometer cells, are reported between parentheses. The unit of observation is a working-age individual (22–60 years old) with a college degree in the 2015 1% Population Survey; and all regressions are weighted such that each county contributes the same as in Table 3. In columns (2)-(4), we further restrict the sample to employed individuals; and we restrict the sample to individuals in a managerial occupation in column (5). All specifications are TSLS specifications and include the extended controls of Table 3 as well as "cohort" dummies for the year of birth. *Unemployed* is a dummy equal to 1 if the individual is not employed and actively looking for employment; *Entrepreneur* is a dummy equal to 1 if the individual reports being the chief or deputy manager of an establishment/enterprise/corporation; *Manager* is a dummy equal to 1 if the individual occupies a (general) managerial position; *Emigrant* is a dummy equal to 1 if the individual resides outside of her province of registration; and *E-manager* is a dummy equal to 1 if the manager resides outside of her province of registration. In columns (1)-(3), the treatment and control variables are nested at the place of residence; and the treatment and control variables are nested at the place of registration in columns (4)-(5).

ployed is 5.6 percentage points higher in treated counties—with treatment defined at the location of residence; the probability of reporting a *top* managerial position (i.e., chief or deputy manager of an enterprise) is 0.6 p.p. lower; and the probability to occupy a managerial position is 2.9 p.p. lower. These effects are sizable: top managers would constitute 1.7% of working-age graduates in control counties versus 1.1% in treated counties. The explanation does not lie in the quality of education: if anything, treated counties produce more highly educated individuals. Instead, labor demand appears to differ between treated and control counties: college graduates in treated locations are 2.2 p.p. more likely to emigrate to another province, with treatment defined at the location of household registration (*hukou*, see column 4),—and college graduates in managerial positions are 8.9 p.p. more likely to hold such positions in another province than their registration location.

The presence of MRPs appears to create a local environment that is not conducive to firm creation. Even when prospective entrepreneurs are produced, they prefer to export their skills and manage firms in other locations.

### 4.3 External validity

The previous sections present findings that are consistent with negative production and technological spillovers. However, China is characterized by a unique set of institutions and policies that may limit the external validity of these findings. In this section, we provide evidence on various features of the local business environment that might be less applicable in other settings: (i) factor market imperfections and a subsequent misallocation of factors (labor, credit, land) across production facilities in treated and control counties; (ii) urban (dis)amenities; (iii) local values, the local political environment, and corruption; and (iv) the role of confounding macroeconomic factors.

**Dutch disease and misallocation of factors** In China, recent research discusses factor market imperfections, e.g., labor (Brandt et al., 2013; Tombe and Zhu, 2019; Mayneris et al., 2018), capital (Hsieh and Klenow, 2009; Song et al., 2011; Brandt et al., forthcoming), and land (Brueckner et al., 2017; Yu, 2024). The presence of large factories may further distort the allocation of resources and factors across production units. For instance, treated counties may experience a form of Dutch disease, whereby production costs become prohibitive for smaller firms entering the market (Duranton, 2011), or the presence of large factories could affect the availability of prime land parcels for other businesses or residents. We provide evidence that: (i) labor costs are lower in treated counties in spite of a more educated workforce—the accumulation of human capital is indeed higher (in stark contrast with Franck and Galor, 2021), and parental aspirations about education do not explain our findings (Appendix D.2); (ii) access to capital and public subsidies does not appear particularly constrained for the average (other) establishment (E.1); and (iii) our main results are robust to controlling for local land prices and characteristics of the local housing stock (D.2). These findings are inconsistent with stories that are purely based on the (distorted) allocation of production resources.[30]

**Urban (dis)amenities** The rapid urbanization of China over the past decades has posed challenges for policymakers (e.g., urban sprawl, pollution). The presence of a large factory might have exacerbated these issues in treated counties. In Appendix D.2, we show that our main findings are robust to controlling for: NO2, PM10, and PM2.5 concentration, and the average commute time.

**Entrepreneurial spirit and local political environment** The influence of MRPs may affect the business environment through less observable channels. For instance, there may be a tight link between the MRPs and local politicians, which could operate in two ways. First, MRPs may use their influence on the local business environment to gain preferential access to resources (Fang et al., 2023; Harrison et al., 2019). Second, the MRPs may be used and favored by local leaders to alleviate social unrest (Wen, 2019). We have very limited data to shed light on these issues, and we look at indirect indicators about the local business environment, its fairness, and its possible impact on entrepreneurial values: the provision of subsidies (Appendix E.1), indirect measures of corruption (E.1), a

[30]In unreported checks, we also investigate the dispersion of factor productivity in treated and control counties and do not find evidence of greater misallocation of capital and labor across and within sectors.

survey about entrepreneurial values and fairness (E.5), and the local presence of pre- and post-Revolution elites, who have been shown to differ markedly in their values (Alesina et al., 2021). While we do not find evidence that subsidies are differently allocated in treated and control counties or that corruption is higher in treated counties, we do find evidence that individuals hold markedly different beliefs about returns to talent. The less dynamic business environment is thus accompanied by an adjustment of priors about social mobility. We show, however, in Appendix D.2 that none of these possible mitigating factors explain the large and swift decrease in output observed in treated counties.

**Public sector, establishment life cycle, and obsolete sectors** To condition the analysis on the role of confounding aggregate factors (e.g., reforms of the public sector or industrial cycles), our baseline specification in Table 4 already controls for the age of manufacturing establishments, their industry, their 6-digit main product, and their "type" (e.g., whether they are State-Owned Enterprises). Indeed, the presence of a large (initially public) factory may affect the involvement of the state in other establishments. The boom and bust of treated counties could reflect the boom and bust of the state sector over the period 1950–2015 in China (Brandt et al., forthcoming). We further show in Appendix D.2 that our findings are robust to controlling for: (i) the percentage of pre- and post-Revolution elites in the population (as defined by Alesina et al., 2021), (ii) the share of subsidized housing, and (iii) the incidence of the state sector within counties. In summary, our main results are not driven by the boom and bust of MRPs themselves, the demise of public firms, the life cycle of local firms, or the marked change in aggregate sectoral activity over the period.

## 5 Concluding remarks

Industrialization and the concentration of large industrial clusters may have long-lasting effects on local economies. This paper relies on a unique experiment in which large factories (MRPs) were quasi-randomly allocated across suitable counties in China, and it follows the evolution of these locations—rather than the plants themselves—in the longer run. As in Kline and Moretti (2014a), we find that the initial investment was effective in spurring transformation from agriculture to manufacturing and in raising local living standards. However, this head start failed to generate positive agglomeration economies in the later period: the large productivity gains observed in the 1980s fully vanished in the 1990–2010 period, and this reversal of fortune occurred even though the MRPs were still productive, innovative, and dynamic. Treated areas did not merely revert to the path followed by control areas; the (other) production units in treated counties became less productive, competitive, and innovative than in control counties.

We provide a careful characterization of the structure of local production and innovation to shed light on the nature of these negative externalities. We find that the structure of production is concentrated around the production chain of the MRPs themselves, while the rest of the local economy is scattered across small production clusters. Through the MRPs' production chains, firms appear to extract rents without incentives to innovate; technological spillovers are therefore minimal. Outside these production chains, firms are also unproductive and charge high markups. One plausible explanation is that these counties do not retain highly educated entrepreneurs, who export their skills to other locations. Through these two channels, early industrialization has a persistent, albeit now adverse, influence on local economies.

Our focus on the long-run effects of industrial policies that lead to the emergence of industrial clusters speaks to a wider literature on place-based policies. A growing body of evidence is investigating their short- and medium-run effects on local economic development and finds positive effects on economic activity and possibly welfare gains. However, there is little evidence on whether these (transient) policies induce the creation of self-sustaining economic gains: do such place-based policies change the development path of treated regions, or do regions revert to their previous state in the long run? Our analysis shows that treated regions have a tendency to specialize (consistent with Greenstone et al., 2010), which conditions a positive effect on the local economy in the short run but comes at the expense of economic diversity and productivity spillovers that are necessary for economic performance in the long run (consistent with Bleakley and Lin, 2012). The nature of negative spillovers in our setting is consistent with the findings and postulated mechanisms discussed in Chinitz (1961), Glaeser et al. (1992), or Glaeser et al. (2015) for the United States.

While our empirical assessment accounts for specific circumstances of the Chinese economy after liberalization, a remaining question is whether our findings generalize to other settings. Few contributions analyze local spillovers over a period of 60 years, and those that do (e.g., Garin and Rothbaum, 2025) do not investigate the structure of production as a potential mechanism, making external validity difficult to assess. In our case, the negative economic spillovers from large industrial investments may stem from the nature of the initial technologies and industries and their subsequent *aggregate* dynamics: although locally successful, relatively low aggregate returns to an industrial sector may discourage workers or entrepreneurs from investing in related skills or industries, leading to the "departure" of new ideas or firms. Future research could help clarify the conditions under which local industrial policies generate self-sustaining local economic gains.

## References


**Alder, Simon, Lin Shao, and Fabrizio Zilibotti**, "Economic reforms and industrial policy in a panel of Chinese cities," *Journal of Economic Growth*, 2016, *21* (4), 305–349.

**Alesina, Alberto F, Marlon Seror, David Y Yang, Yang You, and Weihong Zeng**, "Persistence Despite Revolutions," Working Paper 27053, National Bureau of Economic Research 2021.

**Autor, David H, David Dorn, and Gordon H Hanson**, "The China syndrome: Local labor market effects of import competition in the United States," *American Economic Review*, 2013, *103* (6), 2121–68.

**Bleakley, Hoyt and Jeffrey Lin**, "Portage and Path Dependence," *The Quarterly Journal of Economics*, 2012, *127* (2), 587–644.

**Bloom, Nicholas, Mark Schankerman, and John Van Reenen**, "Identifying technology spillovers and product market rivalry," *Econometrica*, 2013, *81* (4), 1347–1393.

**Bo, Yibo**, *Review of Several Major Decisions and Events [Ruogan Zhongda Juece yu Shijian de Huigu]*, Beijing: The Chinese Communist Party School Press [Zhonggong Zhongyang Dangxiao Chuban She], 1991.

**Brandt, Loren, Gueorgui Kambourov, and Kjetil Storesletten**, "Barriers to Entry and Regional Economic Growth in China," *Review of Economic Studies*, forthcoming.

**_ , Johannes Van Biesebroeck, and Yifan Zhang**, "Challenges of working with the Chinese NBS firm-level data," *China Economic Review*, 2014, *30* (C), 339–352.

**_ , Trevor Tombe, and Xiaodong Zhu**, "Factor market distortions across time, space and sectors in China," *Review of Economic Dynamics*, 2013, *16* (1), 39–58.

**Brueckner, Jan K, Shihe Fu, Yizhen Gu, and Junfu Zhang**, "Measuring the stringency of land use regulation: The case of China's building height limits," *Review of Economics and Statistics*, 2017, *99* (4), 663–677.

**Busso, Matias, Jesse Gregory, and Patrick Kline**, "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," *American Economic Review*, 2013, *103* (2), 897–947.

**Chen, Shuai, Paulina Oliva, and Peng Zhang**, "The effect of air pollution on migration: Evidence from China," *Journal of Development Economics*, 2022, *156*, 102833.

**Chen, Ying, J. Vernon Henderson, and Wei Cai**, "Political favoritism in China's capital market and its effect on city sizes," *Journal of Urban Economics*, 2017, *98* (C), 69–87.

**Chinitz, Benjamin**, "Contrasts in Agglomeration: New York and Pittsburgh," *The American Economic Review*, 1961, *51* (2), 279–289.

**Ciccone, Antonio**, "Input chains and industrialization," *The Review of Economic Studies*, 2002, *69* (3), 565–587.

**De Loecker, Jan and Frederic Warzynski**, "Markups and firm-level export status," *American Economic Review*, 2012, *102* (6), 2437–71.

**Desmet, Klaus and Esteban Rossi-Hansberg**, "Spatial development," *American Economic Review*, 2014, *104* (4), 1211–43.

**Dong, Zhikai**, "On the Establishment of the "156 Program" [in Chinese]," *Researches in Chinese Economic History*, 1999, *4*, 93–107.

__ **and Jiang Wu**, *Industry Cornerstone of New China [in Chinese]*, Guangzhou: Guangdong Economy Publishing House, 2004.

**Du, Xuncheng**, *China's National Capitalism (1840–1927) [Zhongguo de Minzu Ziben Zhuyi (1840–1927)]*, Shanghai: Shanghai People's Press [Shanghai Renmin Chuban She], 2014.

__ , *China's National Capitalism (1927–1937) [Zhongguo de Minzu Ziben Zhuyi (1927–1937)]*, Shanghai: Shanghai University of Finance and Economics Publishing House [Shanghai Caijing Daxue Chuban She], 2019.

**Duranton, Gilles**, "California dreamin': The feeble case for cluster policies," *Review of Economic Analysis*, 2011, *3* (1), 3–45.

__ **and Diego Puga**, "Nursery cities: Urban diversity, process innovation, and the life cycle of products," *American Economic Review*, 2001, *91* (5), 1454–1477.

__ **and** __ , "Chapter 5 - The Growth of Cities," in Philippe Aghion and Steven N. Durlauf, eds., *Handbook of Economic Growth*, Vol. 2 of *Handbook of Economic Growth*, Elsevier, 2014, pp. 781 – 853.

**Eckstein, Alexander**, *China's Economic Revolution*, Cambridge: Cambridge University Press, 1977.

**Faggio, Giulia, Olmo Silva, and William C. Strange**, "Heterogeneous agglomeration," *Review of Economics and Statistics*, 2017, *99* (1), 80–94.

**Fajgelbaum, Pablo D and Cecile Gaubert**, "Optimal spatial policies, geography, and sorting," *The Quarterly Journal of Economics*, 2020, *135* (2), 959–1036.

**Fan, Jingting and Ben Zou**, "Industrialization from Scratch: The "Construction of Third Front" and Local Economic Development in China's Hinterland," *Journal of Development Economics*, Sept. 2021, *152.*

**Fang, Lily, Josh Lerner, Chaopeng Wu, and Qi Zhang**, "Anticorruption, Government Subsidies, and Innovation: Evidence from China," *Management Science*, 2023, *69*, 4363–4671.

**Franck, Raphaël and Oded Galor**, "Flowers of Evil? Industrialization and Long Run Development," *Journal of Monetary Economics*, 2021, *117*, 108–128.

**Garin, Andrew and Jonathan Rothbaum**, "The long-run impacts of public industrial investment on local development and economic mobility: Evidence from World War II," *The Quarterly Journal of Economics*, Feb. 2025, *140* (1), 459–520.

**Giorcelli, Michela and Bo Li**, "Technology Transfer and Early Industrial Development: Evidence from the Sino-Soviet Alliance," Working Paper 29455, National Bureau of Economic Research 2022.

**Glaeser, Edward L, Hedi D Kallal, Jose A Scheinkman, and Andrei Shleifer**, "Growth in cities," *Journal of Political Economy*, 1992, *100* (6), 1126–1152.

**Glaeser, Edward L., Sari Pekkala Kerr, and William R. Kerr**, "Entrepreneurship and urban growth: An empirical assessment with historical mines," *Review of Economics and Statistics*, 2015, *97* (2), 498–520.

**Goncharenko, Sergei**, "Beijing and Moscow: From Allies to Enemies [in Chinese]," in Danhui Li, ed., *The Military Causes of the Sino-Soviet Split*, Guangxi Normal University Press, 2002.

**Greenstone, Michael, Richard Hornbeck, and Enrico Moretti**, "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings," *Journal of Political Economy*, 2010, *118* (3), 536–598.

**Harrison, Ann, Marshall Meyer, Peichun Wang, Linda Zhao, and Minyuan Zhao**, "Can a Tiger Change Its Stripes? Reform of Chinese State-Owned Enterprises in the Penumbra of the State," *NBER Working Paper*, 2019, *25475.*

**He, Yimin and Mingchang Zhou**, "The 156 Program and the Development of Industrial Cities in New China (1949–1957) [in Chinese]," *Contemporary China History Studies*, 2007, *14* (2), 70–77.

**He, Zi-Lin, Tony W Tong, Yuchen Zhang, and Wenlong He**, "A database linking Chinese patents to China's census firms," *Scientific data*, 2018, *5*, 180042.

**Henderson, J Vernon, Zmarak Shalizi, and Anthony J Venables**, "Geography and development," *Journal of Economic Geography*, 2001, *1* (1), 81–105.

**Henderson, Vernon, Ari Kuncoro, and Matt Turner**, "Industrial Development in Cities," *Journal of Political Economy*, 1995, *103* (5), 1067–1090.

**Hsieh, Chang-Tai and Peter J. Klenow**, "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

**Imbert, Clement, Marlon Seror, Yifan Zhang, and Yanos Zylberberg**, "Migrants and firms: Evidence from China," *American Economic Review*, 2022, *112* (6), 1885–1914.

**Khanna, Gaurav, Wenquan Liang, Ahmed Mushfiq Mobarak, and Ran Song**, "The Productivity Consequences of Pollution-Induced Migration in China," *American Economic Journal: Applied Economics*, 2025, *17*, 184–224.

**Kim, Sukkoo**, "Expansion of markets and the geographic distribution of economic activities: the trends in US regional manufacturing structure, 1860–1987," *The Quarterly Journal of Economics*, 1995, *110* (4), 881–908.

**Kline, Patrick and Enrico Moretti**, "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *The Quarterly Journal of Economics*, 2014, *129* (1), 275–331.

_ **and** _ , "People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs," *Annual Review of Economics*, 2014, *6* (1), 629–662.

**Lardy, Nicholas R.**, "Economic Recovery and the First Five-Year Plan," in Roderick MacFarquhar and John K. Fairbank, eds., *The Cambridge History of China, Volume 14, The People's Republic, Part I: The Emergence of Revolutionary China 1949–1965*, Cambridge University Press, 1987.

**Lüthi, Lorenz M.**, *The Sino-Soviet split: Cold War in the communist world*, Princeton University Press, 2010.

**Mayneris, Florian, Sandra Poncet, and Tao Zhang**, "Improving or Disappearing: Firm-Level Adjustments to Minimum Wages in China," *Journal of Development Economics*, 2018, *135*, 20–42.

**Mitrunen, Matti**, "War reparations, structural change, and intergenerational mobility," *The Quarterly Journal of Economics*, Feb. 2025, *140* (1), 521–584.

**Naughton, Barry**, *The Chinese Economy: Transformation and Growth*, Cambridge, MA: MIT Press, 2007.

**Neumark, David and Helen D. Simpson**, "Place-Based Policies," in Gilles Duranton, J.V. Henderson, and William Strange, eds., *The Handbook of Regional and Urban Economics*, Elsevier/Academic Press, 2015.

**Ngai, L. Rachel and Christopher A. Pissarides**, "Structural Change in a Multisector Model of Growth," *American Economic Review*, 2007, *97* (1), 429–443.

**Pierce, Justin R and Peter K Schott**, "The surprisingly swift decline of US manufacturing employment," *American Economic Review*, 2016, *106* (7), 1632–62.

**Rawski, Thomas**, *Economic Growth and Employment in China*, Oxford: Oxford University Press, 1979.

**Rosenstein-Rodan, Paul N.**, "Problems of Industrialisation of Eastern and South-Eastern Europe," *The Economic Journal*, 1943, *53* (210/211), 202–211.

**Schweiger, Helena, Alexander Stepanov, and Paolo Zacchia**, "The Long-Run Effects of R&D Place-Based Policies: Evidence from Russian Science Cities," *American Economic Journal: Economic Policy*, August 2022, *14* (3), 322–51.

**Selden, Mark and Patti Eggleston**, *The People's Republic of China. A Documentary History of Revolutionary China*, New York: Monthly Review Press, 1979.

**Shi, Wenjie**, "A Study of the "156 Program" and its Influence on the Early Modernization of New China [in Chinese]," *Financial Information*, 2013, *1*, 42–45.

**Shirotori, Miho, Bolormaa Tumurchudur, and Olivier Cadot**, *Revealed factor intensity indices at the product level*, Vol. 44, UN, 2010.

**Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti**, "Growing Like China," *American Economic Review*, 2011, *101* (1), 196–233.

**Stewart, Ricky L, Jessica Brede Stone, and Mary L Streitwieser**, "US benchmark input-output accounts, 2002," *Survey of Current Business*, 2007, *87* (10), 19–48.

**Tombe, Trevor and Xiaodong Zhu**, "Trade, Migration, and Productivity: A Quantitative Analysis of China," *American Economic Review*, 2019, *109* (5), 1843–72.

**von Ehrlich, Maximilian and Tobias Seidel**, "The Persistent Effects of Place-Based Policy: Evidence from the West-German Zonenrandgebiet," *American Economic Journal: Economic Policy*, 2018, *10* (4), 344–74.

**Walder, Andrew G. and Yang Su**, "The Cultural Revolution in the Countryside: Scope, Timing and Human Impact," *The China Quarterly*, March 2003, *173*, 74–99.

**Wang, Jin**, "The Economic Impact of Special Economic Zones: Evidence from Chinese Municipalities," *Journal of Development Economics*, 2013, *101*, 133–147.

**Wang, Qi**, "An Analysis of the "156 Program" and Sino-Soviet Relations in the 1950s [in Chinese]," *Contemporary China History Studies*, 2003, *10*, 110–116.

**Wen, Jaya**, "The Political Economy of State Employment and Instability in China," Technical Report 2019.

**Yu, Yue**, "The Local and Aggregate Effects of Land-Use Regulation on Farmland Protection," Technical Report 2024.

**Zhang, Jiuchun**, "A Study of the "156 Program" of Industrial Construction in the 1950s [in Chinese]," *Journal of Engineering Studies*, 2009, *1*, 213–222.

**Zhang, Shuguang**, *Economic Cold War: America's Embargo against China and the Sino-Soviet Alliance, 1949–1963*, Stanford, CA: Stanford University Press, 2001.

**Zheng, Siqi, Weizeng Sun, Jianfeng Wu, and Matthew E. Kahn**, "The birth of edge cities in China: Measuring the effects of industrial parks policy," *Journal of Urban Economics*, 2017, *100*, 80–103.