Industrialization from Scratch: The "Construction of Third Front" and Local Economic Development in China's Hinterland

Jingting Fan and Ben Zou*

Jan 2021

Abstract

This paper investigates whether temporary subsidies to the manufacturing sector at an early stage of structural transformation stimulate economic development. We study the "Construction of Third Front" (TF), a massive yet short-lived industrialization campaign in China's under-developed hinterland. Motivated by defense considerations, location choices of TF projects followed a peculiar set of criteria, which generates plausibly exogenous variation in manufacturing capacity before market reforms started in the mid-1980s. We find initial advantages in manufacturing have long-run positive effects on the structural transformation of the local economy. The effects are driven by new entrants in the private sector, consistent with the existence of local agglomeration economics. However, there is no evidence that agglomeration forces are stronger in initially less-developed regions. While the TF reduced regional inequality, it likely hurt the aggregate efficiency.

JEL code: O14, O18, R1

Keywords: structural transformation, regional economic development, agglomeration economics, industrial policy

---

*Fan: Pennsylvania State University, jxf524@psu.edu. Zou: Michigan State University, benzou@msu.edu. We are grateful to Chris Ahlin, David Atkin, Loren Brandt, Sebastian Galiani, Michael Gechter, Steven Haider, Judy Hellerstein, Ginger Jin, Ethan Kaplan, Yi Lu, Nancy Qian, Heiwai Tang, Yuan Tian, Haishan Yuan, and participants at the AREUEA, Econometric Society Winter Meeting, Urban Economics Association Conference, Northeast Universities Development Consortium, Development Day at Northwestern University, Fudan University, Michigan State University, Penn State University, Lehigh University, University of Calgary, University of Maryland, University of Michigan, and University of Oklahoma for helpful comments. All errors are our own.

1 Introduction

Governments around the world frequently adopt policies that direct manufacturing investment to underdeveloped regions of a country, in the hope that such investment could kick-start industrialization and accelerate local economic growth. From a neoclassical perspective, such policies make little sense—by directing investment to less productive regions, such interventions only introduce distortions and reduce efficiency. However, dating back to as early as Marshall (1890), there have also been many accounts of external economies of scale in the manufacturing sector (for a recent review, see Rosenthal and Strange, 2004). If these externalities are stronger in underdeveloped regions, policies that encourage investment in these regions, though often at expense of investment in other regions, could potentially lead to an increase in the aggregate efficiency.

The key to understanding the effect of such policies thus lies in whether investment in underdeveloped regions could indeed generate long-term spillovers, and whether such spillovers are stronger than in more productive areas.[1] Answering these two questions is challenging, however, due to the potential endogeneity problem—governments take growth prospects into account when deciding where to direct investment, rendering it difficult to estimate causal effects. In cases where such policies could be taken as exogenous, the variations are usually small and limited to local areas, making it difficult to examine the channels at work, let alone to explore heterogeneous effects across regions with different initial conditions.

This paper answers the two questions above by exploiting variations from a particular location-based industrialization policy in China, known as the "Construction of the Third Front" (henceforth "the TF"). Launched in 1964 and lasted for over a decade, the TF built hundreds of large manufacturing plants in the country's remote and mountainous hinterland, an area later known as the "Third Front Region." The TF significantly altered China's economic geography by creating patches of manufacturing clusters in this otherwise underdeveloped region, right before the market reforms started in the mid 1980s. We investigate whether the manufacturing capacity built by the TF stimulated structural transformation of the local economy in the two subsequent decades. With variation in industrial compositions of manufacturing capacity across regions with different levels of initial development, we are able to shed more light on the nature of local agglomeration

[1] In closed-economy settings, this rationale has spurred a literature on the "big push" strategy to stimulate development (Rosenstein-Rodan, 1943; Hirschman, 1958; Murphy et al., 1989). In the regional development context, because the opportunity cost of investment in one region is investment in other regions (and the corresponding spillover effects), the spillover effect being stronger in under developed regions is important for development programs to have a positive aggregate effect. See, for example, Kline and Moretti (2014a).

economics and how it is related to local economic conditions.

Our main finding is that, through within- and between-industry spillovers and other aspects of local agglomeration economics, the manufacturing capacity built up by the TF promoted the development of the local manufacturing sector *after* preferential policies had ended. However, the strength of the spillover effects neither depended on the investment intensity, nor varied across regions with different levels of initial economic development. The lack of state-dependent or nonlinear spillover effects in our setting casts doubt on the existence of poverty trap at the regional level. These findings lead us to conclude that the TF benefited targeted local economies and reduced regional inequality, but did not increase the aggregate efficiency.

Our identification strategy exploits the sudden turns in geopolitical situations and the peculiar site-selection criteria for TF plants. In 1964, when China was embroiled in disputes with the USSR, the Vietnam War escalated. To be prepared for a possible direct military confrontation with either of the two superpowers, China launched the TF, the goal of which was to build self-sustaining industrial clusters in the hinterlands so that the country could have more "strategic depth" during a war. To avoid destruction from mass bombing and to ensure access to transportation, TF plants were instructed to be scattered about rugged areas along the railway lines. Projects followed the gradually expanding railway network. However, after the rapprochement between the US and China in the early 1970s, the strategic priority was abruptly shifted away from the TF. As a result, among the vast rugged areas in the TF region, only places that were connected to earlier segments of the railway network received substantial investment, while places that were connected to the railway slightly later did not.

This prompts us to use the *timing* of the railway construction as an instrumental variable for the size of the manufacturing capacity built by the TF. The identification assumption is that, conditional on the access to the railway network completed by the end of the TF, which we always control for, access to *earlier* segments of the railway network affects local development only through its impact on how much TF investment a region received. We conduct a host of tests to check the validity of this identification assumption.

We compile a panel data of prefectures spanning more than 70 years. The key measure of TF-built manufacturing capacity was from a newly digitized industrial census. Conducted in 1985, the census provides a unique snapshot of the initial economic conditions at the eve of China's market reforms.[2] We find that a one percentage point (p.p.) increase

[2] While 1979 is widely accepted as the starting year of the market reforms, initially reforms were concentrated in rural areas and in the agricultural sector (Naughton, 2007). Reforms in the urban sector, especially those on state-owned enterprises (SOEs) did not start until the mid-1980s.

in the share of employment from TF plants raised the local manufacturing employment share by 2.7 p.p. twenty years after the start of the market reforms. The results are robust to alternative measures of TF-built manufacturing capacity and identification strategies.

About 78% of the effect is accounted for by the private sector, which was essentially non-existent at the time the TF ended. This result is unlikely to be driven by continued state subsidies or the privatization of state-owned plants—we show that the result is quantitatively similar when we focus on sub-samples of firms that were less likely to have been privatized, when we directly control for the growth of the local state-owned sector, and when we explicitly account for the differential privatization process across industries. Using individual-level migration data, we further show that the increase in manufacturing employment represents structural transformation of local economies rather than migration of manufacturing workers across regions, so the TF indeed brings benefits to residents in recipient local economies.

These findings are supportive of the existence of local agglomeration forces in the context of economic development: the TF plants generated positive spillovers and made it easier for private firms to enter and operate productively. Consistent with this hypothesis, we show that TF-built manufacturing capacity increased wages and productivity of the local manufacturing sector. With cross-industry variation in the TF investment, we provide additional evidence on the channels through which agglomeration economics operates. Overall, a substantial share of the spillover effect takes place within the same 2-digit industry of the original TF plants. Investment in a given industry also tends to have larger impacts on the employment and productivity of industries that are economically close, as measured by overlaps in labor pool, and the strength of technological and input-output linkages between industries.

We then investigate the aggregate implications of the TF. It is easy to imagine that directing investment from the productive coastal region to the less developed TF region might result in substantial spatial misallocation of manufacturing activity. However, if a region is stuck in a "poverty trap," the marginal benefit of manufacturing investment there could be either particularly large by greasing the wheels of economic development (e.g., Murphy et al., 1989), or particularly small due to the lack of complementary production factors. In the former case, even if reallocating manufacturing plants to this region reduces their productivity, the aggregate efficiency could still increase. Distinguishing empirically which case is more relevant in the regional setting is thus critical.

To formalize this intuition, we build a simple model of structural transformation and derive a set of testable conditions under which the TF *could* increase aggregate output. Chief among these is whether the spillover effects are stronger in initially less-developed regions. Using variation in initial conditions among local economies *within* the TF region, we find that the TF investment has almost identical spillover effects in places with different initial conditions. It follows that manufacturing investment is unlikely to have generated much stronger spillovers in the TF region than had the investment been made to more productive places. The effects also seem to be linear, rejecting the hypothesis that only large-enough investment could jump start structural transformation.

On the other hand, we found that the TF was effective in alleviating regional inequality. Back-of-the-envelope calculations suggest that, had the TF plants been built in a similar-sized prefecture in the coastal region instead of a typical TF prefecture, the gap in GDP per capita between the coastal and the TF prefectures would be 20% larger than it actually was in 2010.

Limited by the availability of historical data on the costs associated with the TF, this paper does not provide a comprehensive cost-and-benefit analysis of the TF. Due to its poor planning and hasty execution, the TF was likely among the most wasteful of all industrial campaigns. Yet this paper's main finding—that once the campaign is carried out and manufacturing capacity is built, it may not increase aggregate efficiency but mainly operates through redistribution—has broader implications for other local industrial policies. After all, for any such program to work, the same economic mechanisms generating spillovers, such as the ones we documented in this paper, must be at play.

1.1 Related Literature

This paper contributes to the empirical studies of place-based policies (Gottlieb and Glaeser, 2008; Wang, 2012; Neumark and Simpson, 2015; Kline and Moretti, 2014b; Alder et al., 2016; Lu et al., 2019). We find that the effects of a favorable place-based policy may persist long after the policy has ceased, but that they are not stronger in poorer regions. This implies that the main effects of the policy are distributional, as opposed to net efficiency improvements.[3] Our identification strategy is related to recent studies that use quasi-experimental variation in estimating agglomeration effects (see, for example, Redding and Sturm, 2008; Greenstone et al., 2010).

The focus on long-run effects in the context of economic development and heterogeneity by initial conditions differentiates this paper from much of the literature, which

[3]Since this paper focuses on the long-run effect of a temporary intervention, it is also related to the literature investigating whether temporary shocks might have persistent impacts on equilibrium outcomes (Bleakley and Lin, 2012; Jedwab et al., 2017). In that literature, population is often the key outcome of interest. In this paper, however, the TF does not significantly affect the population in the treated regions. Our analysis focuses on the industrialization of the local economies.

either is set in developed countries, or focuses on contemporaneous or short-term effects on wage and employment.[4] An exception is Kline and Moretti (2014a), who examine the impact of the Tennessee Valley Authority (TVA) on the long-run local economic development. By comparing counties inside the TVA region with those just outside, the authors find continued growth in manufacturing employment and productivity in TVA counties after subsidies had ceased. Our paper is different in several important dimensions. First, the TVA was a Big Push investment strategy with broad-based investment in infrastructure, energy, and agriculture. In contrast, the TF is unique in directly investing in manufacturing plant. By focusing on manufacturing plants built during the TF, we can speak to the effects of development policies that favor the manufacturing sector, which are popular in many developing countries (such as the "unbalanced" development strategy proposed by Hirschman (1958)).[5] Relatedly, variation in the composition of industries in the TF investment further allows us to study the channels through which local agglomeration economics takes effect. Second, when the TF was initiated, The TF region was much less developed than the TVA region in the 1930s.[6] Moreover, private firms in China were much more likely to suffer from financial constraint and other distortions, which are also prevalent in many developing countries. A priori, it is unclear whether these shortcomings would imply stronger or weaker spillover effects. The TF thus provides a unique opportunity to test whether manufacturing subsidies can have a large impact at early stage of economic development. Last but not least, with well-documented site-selection criteria, we are able to exploit exogenous variation *within* the targeted region, which adds to the credibility of identification.

More broadly, this paper also contributes to the long-standing debate on the role of government interventions in stimulating industrialization in underdeveloped economies. We find that the TF investment benefits local economies through agglomeration economics, even though the site selection decisions were not made to maximize economic return. But we do not find evidence supporting a regional poverty trap. Empirically, different from earlier research that studies industrial policies and poverty traps using cross-industry or cross-country variation (see, e.g., Beason and Weinstein, 1996; Easterly,

[4]Most related, Wang (2012), Alder et al. (2016), and Lu et al. (2019) look at the effect of special economic zones in China. Most of their results should be interpreted as contemporaneous effects as these special economic zones are still in place when the outcomes are measured. Moreover, because economic zones usually include a varieties of policies, it is hard to attribute the outcome to any specific policy. In contrast, the setting studied in this paper allows for a sharper interpretation of the results.

[5]The TF also included infrastructure projects such as railways and hydraulic dams, but they are not the focus of this paper.

[6]In 1964, the year when the TF was initiated, the average urbanization rate of our sample was merely 7 percent. In contrast, in the 1930s, the average urbanization rate in the TVA region was around 40 percent.

2006), this paper joins a new and fast-growing group of studies that exploit variation at the sub-national level generated by natural experiments (e.g., Criscuolo et al., 2019; Juhász, 2018).

The rest of the paper is organized as follows: Section 2 describes a brief history of the TF; Section 3 describes the data, key measures, and the sample; Section 4 documents the site selection criteria for the TF plants, which motivates our identification strategy; Section 5 presents the empirical strategies and the baseline results; Section 6 explores mechanisms; Section 7 discusses the implications on the aggregate efficiency and equity; Section 8 concludes.

2 A Brief History of the TF

In the early 1960s, China was among the poorest countries in the world with GDP per capita about one-fiftieth of that in the United States. The economy was largely agrarian, with a small industrial sector (manufacturing, mining and extractions, and utilities) that accounted for less than 10% of the employment. Existing industrial clusters were concentrated in the northeast and the lower reaches of the Yangtze River. The vast hinterlands of the country was agrarian and rural, with less than 5% employment in the industrial sector. The economy was just recovering from the Great Leap Forward and the ensuing Great Famine between 1959 and 1961. The primary goal for economic development between 1966 and 1970, according to the *initial* draft of the *Third Five-Year Plan*, was to meet the basic needs of everyday life.[7]

The focus of economic planning changed dramatically in 1964 as the geopolitical situation facing China took a sudden downturn. To the north, the relationship with the Soviet Union deteriorated, with the Soviet troops stationing along the China-Mongolia border, just 500 kilometers across open ground from Beijing. To the southeast, the Vietnam War escalated after the Gulf of Tonkin Incident in August. The resulting mass bombing of the North Vietnam menaced China's border. China's leaders worried that should China find itself at war with the Soviet Union or the United States, the country's existing industrial clusters were too close to the potential war fronts and vulnerable to military attacks.

The "Construction of Third Front" was launched later that year with the goal of establishing self-sustaining industrial clusters in China's southwest and northwest, an area later known as the "Third Front Region." The purpose of the construction was to build "strategic depth," so that China could still have the industrial capacity to support a war

[7] "*jiejue chichuanyong*," which translates into "to solve the problems that involve eating, wearing, and everyday use."

even if it were to lose its factories in the east. As Mao Zedong put it, "it's a way of buying time against the imperialists, against the revisionists....So that if war breaks out we have nothing to fear" (Naughton, 1988).

Figure 1 shows China's distinct "three-rung-ladder" topography with rising elevations from east to west. The TF region is encircled by the red line, which largely corresponds to the second geographic rung. The TF region is a vast area that overlaps with 9 provinces and accounted for about 20 percent of China's population in 1964.[8] Except for a few flat patches, the region is mountainous and the terrain is rugged.

While the TF was planned as a long-term strategy, the harsh geopolitical tensions started to ease sooner than expected. Nixon's surprise visit to China in January, 1972 started the US-China rapprochement. Increasingly, the TF was seemed unnecessary and wasteful. New projects were halted. By the mid-1970s, only unfinished projects received continued investment until they became operational. There was no official end date of the TF. However, the Communist Party's Assembly in 1978, which marked the start of China's gradual transition into a market-based economy, was the *de facto* end of the TF.

Figure 2 documents the scale of the TF investment. The thick blue line shows capital formation per capita (relative to the national average) in provinces that were included in or intersected with the TF region. The thin red line shows the same measure for the rest of the country. In line with the historical account, capital formation per capita in the TF region was noticeably higher between 1964 and the mid-1970s, during which 40 percent of the nation's industrial investment was estimated to have been dedicated to the TF region (Naughton, 1988; Central Documentary Office, 1992). There are two waves of investment in the TF region. The first wave peaked around 1965, when massive projects were initiated in response to the escalation of the Vietnam War. The second wave peaked around 1970, after the Sino-Soviet relation further deteriorated, culminating with the skirmish on the Zhenbao Island in the spring of 1969. Investment in the TF region declined precipitously after 1972. In the Online Appendix, we document the same shift in industrial activities using output-based measures.

Most of the TF investment was made in the manufacturing sector, either to construct new factories or to expand smaller existing ones. Many factories were spinoffs of existing factories in the Northeast or the coastal areas. Some were entirely relocated to the TF region. Investment prioritized plants that produce basic industrial materials, such as steel and chemicals, as well as sophisticated products like machineries and semiconduc-

[8] The TF region include Sichuan (including Chongqing, which became a provincial-level municipality in 1997), Guizhou, Yunan, Gansu, Ningxia, southern Shaanxi, northwestern Guangxi, and the western parts of Hubei and Hunan.

tor instruments. A relatively small number of light industry plants were established as supporting facilities. Some investment was made to infrastructure projects such as roads and dams. Investment in agriculture was minimal.

The geographical and sectoral shifts in the focus of industrial investment during this period are also evident from firm-level data. The *Complete List of Large and Medium Sized Industrial Plants* in the *1985 Industrial Census* provides a detailed snapshot of China's major industrial presence by the end of the country's command economy era. Table 1 documents the changes in characteristics of plants that were established before the TF (before 1964, Column 1) and those during the TF (between 1964 and 1978, Column 2). Panel A shows the characteristics of counties in which the plants were located. Only 13% of the plants established before the TF period were located in the TF region, compared with 32% among those established during the TF period. Plants established during the TF period were half as likely to be located in a provincial capital. They were more likely to be located in counties that were farther away from provincial capitals, in higher elevations, and had more rugged terrains. Column 3 reports the differences in mean values of each characteristic and the associated standard deviations. It shows that these differences are large and are all statistically significant. Considering that some of the TF investment was made to expand existing plants in the hinterlands, the differences shown here likely understate the actual significance of the TF. Panel B shows that plants established during the TF period were less likely to be in mining, utilities, and light industries. They were significantly more likely to be in industries that produce chemical products, machinery, electric equipments, and electronic devices.

The TF substantially altered the landscape of China's economic geography. In 1979, provinces in the core TF region (Sichuan, Guizhou, Shaanxi and Gansu) accounted for 9 percent of the national industrial output, compared to less than 6 percent in 1964.

Within the TF region, some locations received large plants and developed substantial manufacturing capacity by time market-oriented reforms started in China's urban sector. For example, Mianyang in Sichuan province developed into a cluster of electronic device makers; Shiyan in Hubei province was home to China's second largest automobile factory. At the same time, many places in the region, similar in initial economic and geographical conditions, remained largely agrarian and undeveloped.

In a series of reforms, the TF plants, all of which state-owned, were gradually granted more autonomy and were weaned from government subsidies. Some TF plants, situated in extremely difficult locations, were relocated to nearby cities. Private firms were allowed to enter and grew rapidly. By 2004, over 60% of manufacturing employment nationwide was in the private sector. We investigate whether the manufacturing capacity built by the

TF contributed to the continued structural transformation of local economies and the rise of the private sector,[9] and if so, whether the effect was stronger in initially poorer regions.

3 Data, Measurement, and Sample

3.1 Data Sources

From various sources, we assemble a dataset of Chinese prefectures covering a period over 70 years. Some data are digitized from historical archives and are, to the best of our knowledge, new to the literature. The data cover a wide range of topics, including local economic conditions, demographic and geographic characteristics. Most of these data provide information at the county or prefecture level, while others are at the person or firm level. This section briefly describes the main sources of data. More details on variable construction are provided in Online Appendix A.

Our data can be broadly classified into three groups by the time of measurement: *before* the TF, *right after* the TF but before the market reforms, and the *long run* after the market reform. Data in the "before" period are measured in or before 1964. They provide information on local economic conditions prior to the TF, which took place between 1964 and 1978. Information during the TF is unfortunately scarce and of low-quality, despite our best effort in data collection. While we provide some evidence using contemporaneous measures during the TF period, baseline analyses use a treatment variable constructed from censuses and national surveys conducted in the early 1980s, which provides a snapshot of the Chinese economy right before the start of the urban-sector market reforms. Finally, we measure the "long run" outcomes on local economic development using data from the early 2000s, two decades after the start of market reforms.

Data sources for the "before" period mainly include the *1936 National Industrial Survey*, *1964 Population Census*, and GIS maps for geographic and infrastructure features. The *1936 National Industrial Survey* provides the first ever snapshot of the industrial sector in China. We digitize the *Survey* and measure manufacturing employment in each prefecture. Statistics from the *1964 Population Census* is available at the county level and is used to construct two measures of a prefecture's pre-TF economic conditions: urban

9 Existing views differ with regard to this question. On the one hand, it is possible that large plants stifled entrepreneurship in the local economy (Glaeser et al., 2014). Indeed, looking back on China's transition into a market economy, a popular view is to see the inefficient state sector as a drag on local economic development. On the other hand, equipped with arguably the most advanced technology and skilled workers in China at that time, the TF plants could facilitate the entry and growth of local private firms by passing on know-how and skills (e.g., Glaeser and Kerr, 2009).

population share, defined as the fraction of local workers holding an urban *hukou*,[10] and population density. We extract geographic characteristics from GIS maps obtained from the *China Historical GIS Project*, with which we construct each location's average elevation, average slope, and distance to provincial capitals.

Data sources for the "right after" period include the *1985 Industrial Census*, which we digitize from various volumes of statistical publications, and the *1982 Population Census*. Both datasets include county level statistics; in addition, the *1985 Industrial Census* reports a complete list of large and medium sized (LMS) manufacturing plants in the country. For each firm, the list reports its name, detailed address, industry (divided into 39 industries, roughly corresponding to 2-digit SIC codes), employment, capital, output, year of opening, and names of its key products. These firms accounted for over 70 percent of total manufacturing output in the TF region in 1985.

Data sources for the "long run" period include firm-level data from the *Annual Manufacturing Firm Survey in 2004*, the *2004 Economic Census* and individual-level data from the *Intra-decennial Population Census of 2005*. We use firm-level information to measure industrial development by firm ownership, size, and industry; we use individual-level information to construct variables on income, migration, and additional demographic characteristics.

Digital maps of railway lines are from Baum-Snow et al. (2017). We use two snapshots of the railway networks. The first snapshot includes railway lines that *had existed* or *were under-construction* in 1962, right before the TF. We call this the *existing* railway network. The second snapshot includes railway lines in 1980, right after the TF, which we call the *complete* railway network. We measure distances between each location to these two networks.

The unit of analysis is a prefecture, the administrative unit between a county and a province. In 2000, China is divided into 345 prefectures. A prefecture is a suitable definition of a local economy. With about 25 thousand square kilometers and 4 million residents in a typical prefecture, commuting across prefecture borders is rare so spatial spillovers are less of a concern. Furthermore, some TF plants were relocated from their original sites to nearby cities after the reforms, but they typically did not move across prefecture boundaries. So the measure of TF investment is less likely to be confounded by later, potentially endogenous, plant relocations. Prefecture boundaries have changed little over time. We aggregate variables from the county level to the prefecture level using

[10] *Hukou* refers the residence-registration system in China. Before the market reforms, *hukou* status was tightly linked to occupation. Only those who worked in the urban sector—manufacturing, services, government, and military, and their children—had urban *hukou*.

consistently-defined county boundaries. Details of this process are provided in Online Appendix A.

3.2 Measurement

The treatment of our empirical analysis is the size of manufacturing sector in a prefecture by the end of the command economy, which we measure as the share of prefecture *employment* from the LMS manufacturing plants listed in the *1985 Industrial Census*. This measure captures the variation in the size of manufacturing sector due to the TF because major manufacturing plants located in the TF region in the early 1980s were primarily a result of the TF. We call this measure *TF-built manufacturing capacity*. Since some TF investment was made to expand existing firms in the region, in the benchmark analysis we include all LMS plants on the list regardless of years of opening.

Our baseline measure focuses on employment, rather than capital stock or revenue, because prices in a planned economy might be distorted and therefore fail to capture the true size of a manufacturing sector. There are two concerns on whether this measure truly represents the manufacturing capacity the TF had built. First, this measure might miss investment in smaller plants. Because the TF plants were generally large enough to qualify for the list of the large and medium plants, we believe the first concern is unlikely to be important. In a robustness check, we control for employment from smaller plants and show that the result is unchanged.

Second, given this variable is measured a few years after the end of the TF, even if the distribution of the *initial* TF investment is exogenous, what we measure in 1985 might be subject to survival bias, i.e., plants located in places with better economic development potentials would have grown larger by the time data were collected. The timing of the *1985 Industrial Census* and the reform in China's industrial sector help alleviate the second concern. For the survival biases to be significant, factories need to operate with autonomy in the market environment for sometime. In the setting here, although much of the TF investment was made in the late 1960s and the 1970s, it took many TF plants several years before they finally became operational. Moreover, before the market reforms in the industrial sector started in the mid-1980s, state-owned firms had little autonomy in hiring and firing decisions. For these reasons, the size of these plants in 1985 remains a good proxy for capacity built during the TF. The instrumental variable strategy introduced in the next section addresses the survival bias directly. To further alleviate this concern, we show that results are robust to alternative variables that provide direct and contemporaneous measures of the TF investment.

The outcomes of interest are measures of local manufacturing sector performance. The baseline outcome variable is manufacturing employment share in 2004, constructed from the *2004 Economic Census* and the *2000 Popular Census*. To highlight local spillovers and agglomeration economics as the potential mechanism, we also examine the employment share of the *private* manufacturing sector. The private sector was non-existent before the market reforms took place, so its share of prefecture employment could be interpreted as the *level growth* since the market reforms. To rule out privatization as a contributor to our finding, we follow a narrow definition of the private sector, excluding all firms that are solely or partly owned by all levels of governments, and all joint-venture firms between any entities and the state sector. Additional robustness checks are conducted to rule out privatization as a confounding factor

3.3 Sample

We focus on the TF region, which contains 89 prefectures. Several restrictions are made to refine the sample. First, in China's hierarchical system of cities, provincial capitals receive preferential treatments and usually follow different development paths compared with the rest of the province, so we exclude the 8 provincial capitals in the region. Second, some places in the TF region are extremely remote and sparsely populated, but had a high urban rate prior to the TF because they were home to special industries. For example, Jiuquan of Gansu province, which is in the Alxa desert, has been the site of China's rocket-launching facility since 1957. The combination of low population density and the large number of on-site workers makes urban rate in 1964 a bad proxy for pre-TF economic conditions or development potentials. We follow a simple cutoff rule and drop prefectures with a share of urban population of 15% more higher in 1964. The baseline sample includes 73 prefectures. We show in the Online Appendix that our results are robust to both the inclusion of provincial capitals and using alternative cutoff values of the 1964 urban rate.

Table 2 reports the summary statistics of key variables for the baseline sample. The average urban rate in 1964 was 7.14%. Employment share from the LMS manufacturing plants accounted for about 1.11% of the total employment in the average sample prefecture. By 2004, the average manufacturing employment share was 4.29%, among which over 60% was from the private sector.

4 Site Selection of the TF Plants

This section describes the site-selection criteria for TF plants, which motivate our identification strategy. The TF was planned and carried out in a top-down fashion. The *Central Committee of the Third Front Constructions* (hereafter *Central Committee*) was in charge of distributing hundreds of projects across the region. Defense from potential military attacks, especially mass bombing and nuclear attacks, was the top consideration. The *Committee* issued site-selection criteria which stated that the TF plants should be "*dispersed, hidden, close to mountains*" (Central Documentary Office, 1992), so damage from weapons of mass-destruction could be minimized. According to these criteria, existing population centers and places with relatively flat terrain were to be avoided.

Consistently with these criteria, Table 3 Panel A shows that among the LMS plants located *within* the TF region, those built during the TF period were much less likely to be in a provincial capital compared to those built before the TF period. The site chosen for TF plants also tended to have more a rugged terrain, a higher elevation, a lower population density and urban rate in 1964, and a smaller industrial sector in 1936. These patterns are similar to the nationwide shift in plant location choices documented in Table 1. Also similar with the nationwide shift, Panel B shows that plants built during the TF period are more likely in advanced industries (machinery, electric and electronics) and less likely in light industries.

Hiding factories in mountains far away from major population centers made the shipment of construction materials, capital equipment, and output a challenge, so in implementation, access to railway became a necessity. Although multiple railway lines in interior China had long been planned, in 1964, when the TF was launched, only a subset of them had actually been built. Pushing forward the TF program in a great sense of urgency, decision markers therefore had to spread a large number of plants along existing railway lines. With railway network in the region gradually expanding, additional plants were then scattered along the segments of the *new* lines that cut through mountainous areas.[11] The planned multi-decade program was cut short,[12] however, because of decreasing strategic value due to the easing tension between China and the U.S. in the early 1970s, combined with the realization by decision makers of the egregious costs of building large manufacturing plants in the mountains. The TF investment dwindled quickly since the

[11]Many attribute railway construction in the region during the period to the TF program. However, many of these lines were already planed before the TF was launched. One legitimate concern is that the TF program might have affected which of the "planned" railway lines ended up being built first, leading to potential endogeneity. We discuss this possibility extensively in Section 5.2.

[12]The TF was planned as a long-term program, as Mao famously claimed that although the exact time is impossible to predict, a war will break out eventually and China should always be ready for it.

mid-1970s. As a result, only places near railway lines that were completed earlier received much TF investment, while places near railway lines that were completed slightly later did not.

Figure 3 provides graphic support for the site-selection criteria of the TF plants. Each red dot represents 25 workers from the LMS plants in 1985. The colored image of the topography shows elevation and terrain. Red lines show the railway lines that had existed or were under construction in 1962, which we refer to as the *existing* railway network. We refer the railway lines finished by 1980 as the *complete* network, which is plotted in blue lines.[13]

Two patterns are apparent from the map. First, except for the provincial capitals, an overwhelming share of the TF investment was in mountainous areas. It is striking that, much of the Sichuan Basin (the round green area in the middle of the map) outside of Chengdu and Chongqing, being the only large plot of flat land in the region, had little manufacturing capacity by 1985. Second, while TF plants are generally scattered along the railway, only places along the existing railway lines had substantial manufacturing employment.

Panel A of Table 4 documents statistical evidence for the location choice criteria. The dependent variable is the employment share from LMS plants and the independent variables are various proxies for the criteria. *Dispersion* is measured as distance to the provincial capital, which usually is the most important population and economic center in a province; *hiddenness* and *closeness to mountains* are measured with two geographic characteristics: average ruggedness (slope) of the terrain and average elevation; *access to the railway* is measured as the distance to railway networks. We control for province fixed effects in all regressions to account for cross-province differences.

Column 1 shows that places with better access to the complete railway network had larger TF-built manufacturing capacity. Column 2 shows that TF investment is particularly concentrated in prefectures closer to the existing railway network before the TF. Conditioning on the distance to the complete railway, if the distance to the existing railway decreases by 1 log point, TF investment increases by 1 p.p., which is about 90% of the sample average. In Column 2, the log distance to the complete railway network is no longer statistically significant and has the "wrong" sign. This suggests that places connected to stretches of railway lines that were built between 1962 and 1980 did not receive much investment, as is evident in Figure 3.

[13] The minor misalignments between the blue and red lines in early segments of the railway network are due to measurement errors—the GIS railway maps are digitized from maps in hardcopies (Baum-Snow et al., 2017).

Column 3 adds other geographic proxies for site selections. Prefectures with more rugged terrain had larger manufacturing capacity, although conditional on the distance to railways and ruggedness, TF investment did not go to places with the highest elevation. Distance to the provincial capital is not a predictor of manufacturing capacity, suggesting that instead of concentrating around the population and economic centers, TF investment is dispersed across the province.

The remaining columns suggest that our proxies for the selection criteria are not masking some fundamental economic advantages of prefectures that receive large amounts of TF investment. Columns 4 and 5 show that, in general, places had larger manufacturing capacity built up by the TF also had better initial economic conditions prior to the TF. However, Columns 6 and 7 show that, once we control for the selection criteria, initial economic conditions no longer play a role—the test for joint significance of these conditions has a p-value of 0.5. These results confirm that the site selection criteria were followed and that conditional on these criteria being met, local economic potentials were not important considerations for site selection of the TF plants.

As a placebo test, Panel B of Table 4 repeats the same exercises on non-provincial-capital prefectures *outside* the TF region. The dependent variable is the share of employment from the LMS manufacturing plants in 1985. Columns 1 and 2 show that manufacturing employment is clustered near railway lines, but importantly, the clustering is not particularly pronounced along the existing lines. Controlling for distance to the existing railway line on top of distance to the complete railway lines adds little explanatory power. Column 3 shows that places closer to provincial capitals, not those with a more rugged terrain or higher elevation, had larger manufacturing capacity in 1985. Perhaps most convincingly, Columns 4 to 7 show that economic conditions before the TF period are strong predictors for the spatial distribution of manufacturing capacity and remain so after proxies for the location-choice criteria are controlled for (the p-value of the joint test for initial conditions remains close to 0 in Column 7). These results confirm that the test in Column (7) of Panel A was a result of the site-selection criteria for the TF, rather than some systematic measurement issues.

5 Empirical Strategy and Baseline Results

The facts documented in the previous section about the TF indicate that we should control for the site selection criteria. Moreover, we can use the *timing* of railway construction as an instrument for TF investment, which would exploit the interaction between the unexpected cutback in the aggregate scale of the TF and the gradual expansion of the railway network.

5.1 Regression Specification

The baseline empirical model is as follows:

$$y_{i,04} = \beta ManuCap_{i,85} + LocChoice_i \cdot \theta + PreTFCond_{i,64} \cdot \gamma + s_p + \varepsilon_i. \qquad (1)$$

Each observation is a prefecture. $y_{i,04}$ is some measure of the manufacturing sector in prefecture i in the post-reform period. In the baseline specification, it is the prefecture's share of manufacturing employment in 2004. $ManuCap_{i,85}$ is the TF-built manufacturing capacity, measured in 1985 as the share of prefecture's employment in the LMS manufacturing plants. $LocChoice_i$ is a set of proxies for location-choice criteria, which include distance to the provincial capital, average elevation and slope, and distance to the complete railway network in 1980. Distance to the complete railway network serves two purposes: First, it is a proxy for the planned network in the information set of decision makers when laying out the TF program; Second, it also captures any direct effect of access to railway by the end of the TF on subsequent economic development.

$PreTFCond_{i,64}$ is a vector of prefecture economic conditions prior to the TF, which includes urban rate and log population density in 1964, and share of industrial employment in 1936. We further control for the endowment of natural resources, proxied by share of employment in the mining sector. s_p is a set of province fixed effects. ε_i is the error term, which is estimated using the heteroskedasticity-robust formula. β is the parameter of interest. Its interpretation is the percentage point (p.p.) increase in the share of manufacturing employment in 2004 as a result of a 1 p.p. increase in the employment share from TF manufacturing plants. β can be loosely interpreted as the job multiplier. We are interested in whether β is greater than 0, indicating whether initial manufacturing capacity has any positive long-run spillover impact, and also whether β is greater than 1, indicating whether initial investment in manufacturing has a multiplying effect.

Given TF's site selection criteria documented above, one may argue the selection-on-observable assumption of the TF investment is likely satisfied, so in principle Equation 1 can be estimated using the OLS. However, there might still be concerns that the control variables do not capture all relevant information used in decision making. Moreover, the time gap between the end of the TF and when the manufacturing capacity was measured could potentially introduce survival bias. We return to both issues in the robustness checks presented below. In the baseline, we address both concerns using an instrumental variable for TF investment.

Specifically, we use the distance to the existing railway in 1962 as an instrument for the TF-built manufacturing capacity. Since we always control for the distance to the complete railway in 1980, which captures any direct effect of railway on the local economy, the instrument should be interpreted as the effect of being connected to the railway network *earlier*.[14] It effectively exploits variations from the interaction between the sharp decline in the TF investment due to changes in geopolitical situations and the gradual expansion of the railway network. If valid, the instrument also solves the survival bias problem.

5.2 Identification Strategy and Potential Threats

The relevance of the instrument is evident from our earlier discussion of Figure 3 and Table 4 Panel A. First stage regressions are shown in baseline results reported in Table 6. The validity of the instrument requires that conditional on the access to the complete railway network, a prefecture's access to the existing railway before the TF is uncorrelated with unobserved factors that determine local economic development. This assumption seems reasonable in our context, given the window of time difference exploited here is relatively narrow—most of the *existing* railway lines in our definition were not completed until in the 1950s, and a large fraction were still under construction when the TF was initiated; the *complete* railway lines, on the other hand, were all completed by 1980. With such a small difference in the timing of railway completion, the relevance of our instrument comes primarily from the fact that the TF was launched unexpectedly in 1964 and then halted prematurely in the early 1970s. This short window also alleviates any concern that railway lines built earlier might have a different *direct* effect on economic development after the market reforms.

One may be skeptical that because distances to the complete and existing railway networks are likely highly correlated, the instrumental may not be credibly identified. The correlation coefficient between the two log distance variables is 0.77, but the strong raw correlation does not impose a threat to identification by itself. Panel A of Table 4 shows that the coefficient associated with the log distance to existing railways is statistically significant and remain stable in magnitude across various specifications with additional controls. Importantly, from Column 1 to Column 2, when the log distance to existing railways is added alongside with the log distance to complete railways, the model fits

[14] The identification in our specification is different from a growing body of literature on the effect of transportation infrastructure that exploits variation in *route* of railway or highway lines (e.g., Banerjee et al., 2020; Faber, 2014). Our specification conditions on the access to the complete network in 1980 and uses variation from the *timing* of construction. It is similar in spirit to Sequeira et al. (2020), who use the interaction between the timing of immigrants arrival in the United States and the timing of railway construction to instrument for the spatial distribution of immigrants across U.S. counties.

the data much better — the R^2 rises from 0.317 to 0.531, a substantial increase with just one additional control variable. This is in stark contrast to prefectures outside the TF region. Panel B of Table 4 shows that, although distance to railway is highly correlated with manufacturing capacity (Column 1), when both distance variables are included in the regression, neither is statistically significant (Column 2, although they remain jointly significant at the 5% level). The R^2 rises only slightly, from 0.247 to 0.252. These results confirm a unique feature of the TF – investment follows the *timing* of railway construction.

One may also be rightfully concerned that places with better economic potentials were connected to the railway earlier. The nature in the design of China's railway network helps alleviate this concern. China's railway lines were planned to connect major cities, mostly provincial capitals. Other places were linked along the way simply because they happened to be *en route* between two nodes (Banerjee et al., 2020). Because provincial capitals are excluded from the sample, the instrument exploits, among the remaining prefectures, the variation in which of them happened to be sited between pairs of provincial capitals that were connected *earlier or later*. Since the existing railways were all built or planned before the TF, when most of these prefectures were rural areas with little industrial presence, it is unlikely that they had an influence on which provincial pairs got connected first.

Instead, one major factor determining which pairs of provincial capitals were connected earlier is the engineering constraint. This was especially true in the TF region due to its rugged terrain and complicated geological features. Often, engineering difficulties in a small segment could drastically delay the completion of the whole railway line, affecting railway access in places hundreds of kilometers away. Due to those engineering constraints, the timing of access to a railway is farther removed from a location's economic potential. For example, the idea of a *chuan-han* line, which connects Sichuan with Wuhan, dated back as early as the 1910s, and was partially constructed in the 1950s. It was not completed until 2009, almost a century behind schedule, due to engineering difficulties near the border between Sichuan and Hubei, where the Yangtze River forms a deep gorge at the fall line.

While the identification assumption is not directly testable, under the presumption that local economic potentials are positively correlated with pre-TF economic conditions, we can test whether prefectures with better economic potentials were more likely on the earlier segments of the railway network. Specifically, we regress initial economic conditions prior to the TF period on the distance to the existing railway, while controlling for the distance to the complete railway and a set other pre-determined characteristics. Columns 1 through 3 in Table 5 show that there is no evidence that the distance to earlier segments is statistically correlated with other economic conditions prior to the TF period.

Our identification requires the *timing* of railway construction to be exogenous to the sample prefectures. A perhaps more demanding version of the identification assumption is that the *route* of the railway network is exogenous to the sample prefectures—if which of the prefectures would *eventually* be connected is exogenous, it would be even less likely that which of them were connected *earlier* was endogenous. We provide direct evidence that the route is itself uncorrelated with proxies for economic prosperities. Columns 4 through 6 of Table 5 show that overall, distance to the complete railway network is uncorrelated with pre-TF economic conditions. For example, the coefficient in Column 4 shows that a 1 log point decline in the distance to railway (about two-thirds of a standard deviation) is associated with a 0.4 p.p. increase in urban rate in 1964, or one-sixth of a standard deviation. The correlation is also statistically insignificant at conventional levels. Therefore, there is no statistical support that the railway lines endogenously zig-zag to pass by prefectures with better economic potentials.

To further mitigate the concern that endogenous railway routes could implicate our identification strategy, we construct hypothetical railway networks in which the routes are determined only by geographical and engineering parameters. Given a pair of terminals (nodes) of a railway line, we construct two sets of hypothetical routes that minimize either the length or the construction cost. We then recover the railway networks in 1962 and 1980 by inspecting which pairs of nodes were directly connected in each year.[15] Appendix Figure E.1 illustrates the hypothetical railway networks (with actual networks in the background as a comparison). Appendix E.1 describes the details on how the hypothetical networks are constructed. To insure against distance to nodes in itself having an impact, we further control a prefecture's distance to the nearest node when we use distances to the hypothetical networks.[16]

[15] In their pioneering work, Banerjee et al. (2020) uses straight lines between terminal cities in order to purge out endogenous route choices; Faber (2014) constructs a minimum-cost network of China's highway system. Our approach is slightly different from both. In these studies, the counterfactual optimization re-designs the whole network, which may involve connecting pairs of nodes that are different from the actual network. In our case, we take railway lines connecting two nodes as given, and only optimize the routes at the railway line level.

[16] A related concern is the choice of nodes would affect the routes of the hypothetical railway lines. This concern turns out to be unimportant. Most of the nodes are provincial capitals or highly-industrialized cities by 1964, which are already excluded from our sample. Excluding all node cities from our sample would not affect the result either.

5.3 Baseline Results

Table 6 reports our baseline results. All models except the one in the last column are estimated using the Two-Stage Least Squares (2SLS) estimator, with distance to the existing railway lines as an instrument for the TF-built manufacturing capacity.

Column 1 presents the result from the most parsimonious specification, where the only control variables except for manufacturing capacity are the log distance to the complete railway network and a set of province fixed effects. The instrumental variable in the first-stage regression is statistically significant and has expected signs. The Kleibergen-Paap F-statistic is about 9, indicating a reasonably strong first stage. Column 2 adds proxies of other site-selection criteria: average slope, average elevation, and the log distance to the provincial capital. Column 3 further adds pre-TF economic conditions: urban rate and population density in 1964, industrial employment share in 1936, and a proxy for natural resources. The power of the first stage becomes a bit stronger while the coefficient associated with pre-TF manufacturing capacity barely changes.

The models in Columns 4 and 5 replace distances to the actual railway networks with those to the hypothetical ones. Using either the minimum-distance network or the minimum-cost network, we have similarly strong first stages. The key coefficients in the first stages have expected signs, and the coefficients of interest are positive and statistically significant, though somewhat smaller than that in Column 3. Column 6 reports results from estimating Equation 1 using the OLS. Reassuringly, the coefficient is comparable with the 2SLS estimates.

Our preferred specification is the one in Column 3. The coefficient suggests that if a prefecture's share of employment in TF plants increases by 1 p.p. (about 90% of the sample mean), share of manufacturing employment in the prefecture two decades later increases by 2.7 p.p. Considering that the average value of the outcome variable in the sample is 4.3%, this effect is substantial. We can reject the null hypotheses that the coefficient is not statistically different from zero or from one at the 95% confidence level.

5.4 Robustness Analyses

5.4.1 Alternative Measures of the Pre-reform Manufacturing Capacity

In the baseline, TF-built manufacturing capacity is proxied by the share of the prefecture's workers from the large- and medium-size manufacturing plants that were listed on the *1985 Industrial Census*. Since LMS plants were the backbones of China's industrial presence, and TF projects were predominately large, the measure using LMS plants well represents the industrial legacy from the TF. However, because we do not directly measure the TF investment and there is a time gap between the end of the campaign and our measurement, there may be concerns about this measurement. In this section we spell out several possible concerns, construct alternative measures, and conduct robustness checks. The results from those robustness checks are summarized in Table 7.

One obvious omission from using the LMS plants as the measure of manufacturing capacity is that it misses out smaller plants. It is not clear whether smaller plants should be counted as legacies of the TF. Some of those smaller plants may be part of the TF investment, producing parts and supportive equipments for larger plants; others were local township and village enterprises (TVEs), which started to sprout in the early 1980s. The *1985 Industrial Census* reports the total industrial employment at the county level. We measure the employment in smaller plants by deducting the employment in LMS plants from total industrial employment. Column 1 includes employment share from those smaller plants. Employment from smaller plants have an independent effect on the development of the manufacturing sector. One additional worker from those smaller plants can be roughly translated into 1.3 additional manufacturing workers in 2004. Important for us, the coefficient associated with employment share from LMS plants remains similar in magnitude, which suggests that our baseline result is unlikely to be biased due to omission of employment from smaller firms.

The TF (unofficially) ended in 1978 while our measure is from 1985. Because we are interested in the effect of the initial manufacturing capacity before the market-oriented reforms, which did not start until the mid-1980s, measures from 1985 is well-suited for our purposes. Clearly, there were changes in manufacturing employment and output between 1978 and 1985 that could be endogenous. However, to the extent that the growth (or decline) is determined by time-invariant local fundamentals, such as geography and natural resources, our instrumental variable addresses this endogeneity issue.

Yet the endogeneity in changes of the manufacturing sector between 1978 and 1985 could also come from some *time-varying* unobservable factors, *and* these time-varying factors could also affect local economic development after 1985. For example, it is possible that in the late 1970s, the prefecture was led by someone who was open to economic reforms, whose reforms not only promoted the growth of the local manufacturing sector between 1978 and 1985, but also sowed seeds for subsequent economic development. If this is the case, our instrumental variable, which exploits variation in industrial investment during the TF period and we argue is uncorrelated with local fundamentals, does not solve the new endogeneity issue that took place after the TF.

To provide a robustness check immune to this concern, we collect additional data on prefectures' manufacturing capacity in 1978. Statistical records for the period before the market reforms were sporadic and typically not of high-quality. We scrape volumes of statistical yearbooks and local gazetteers to compile a panel dataset on *industrial output* which is relatively consistently defined across different data sources.[17] We measure manufacturing capacity as the prefecture's industrial output (in 1978 *yuan*) divided by the prefecture's total employment (in hundreds, from the 1982 population census). Column 2 shows that a higher industrial output per worker in 1978 leads to a higher manufacturing employment share in 2004. A 100-yuan increase in the industrial output in 1978 leads to 2.1 more manufacturing workers the prefecture in 2004. The data sources we have do not provide information on the number of manufacturing workers before 1985, so we cannot directly compare this coefficient with the baseline estimate. But back-of-the-envelope calculation suggests that the magnitudes are similar.[18] The similarity of the magnitudes from using measures in 1978 and 1985 suggests that time-varying endogenous factors between the end of the TF and the time of our main measure is not a serious concern.

The panel data on prefecture-level industrial output also allows us to construct the change (in levels) in the per capita industrial output between 1964 and 1978, which provides a *contemperaneous* measure of the magnitude of the TF investment. This measure captures the degree to which the industrial sector had developed during the TF period, although it does not speak to the cost involved. Column 3 shows that larger TF investment leads to a larger manufacturing sector in 2004.

Alternatively, while manufacturing employment measured in 1985 may be confounded by endogenous growth since the end of the TF, the restructuring and bankruptcy of state-owned enterprises took place much later.[19] So using the number of plants as a measure is less susceptible to the endogeneity issue between 1978 and 1985. As another robustness check, we replace employment share with the number of LMS plants (per 10,000 workers). Column 5 shows that the first stage is weaker with an F-statistic of 5.7, suggesting that the number of plant by itself is a crude measure of TF investment—plant size also matters. The coefficient suggests that an additional plant leads to 8,400 additional manufacturing

[17] In addition to manufacturing, the industrial sector might also include utilities, mining and extraction, and industry-related services. We are able to find records for 67 prefectures in our sample. See Appendix A.2 for details on definitions and data sources.

[18] The *1985 Industrial Census* shows that the average industrial output for a prefecture in the sample is 7.1 million *yuan* (in 1985 current). With an annual inflation of 3.6% and an annual real GDP growth of about 8% between 1978 and 1985 (those numbers are from the World Bank World Development Indicators), we can convert the industrial output to be 3.2 million in 1978 *yuan*. The number of workers from the LMS plants was 21,997 for an average prefecture in the sample. The coefficient in Column 2, 2.11, could be translated into a coefficient of 3.1 if the independent variable is the share of workers from the LMS plants. This is very close to the coefficient in the baseline, which is 2.7.

[19] The first documented corporate exit took place in Liaoning Province in 1986.

workers in 2004. Using the average size of an LMS plant (about 1,600 workers), the effect can be converted into one that is comparable with the baseline specification (effect of LMS employment share in 1985 on manufacturing employment share in 2004). The coefficient in Column 4 implies an effect of 5.1, which is almost twice as large as the baseline result. The large coefficient may be a result of a weaker first stage, which could exacerbate the small-sample bias of the 2SLS estimate.

Finally, we construct a measure of *planned* industrial centers based on the *The Third Five-year Plan*, published *before* the TF constructions.[20] We identify 10 designated industrial centers out of 73 prefectures in the baseline sample. Probably due to the small variation in the binary variable indicating an industrial center, the first stage is rather weak. Access to the existing railway does predict the location of TF industrial centers, but the coefficient is barely significant at the 5% level. The coefficient in Column 5 shows that being a designated industrial center results in a manufacturing employment share in 2004 that is 21 p.p. higher.[21]

5.4.2 Conditional Independence Assumption

Consistently estimating β in Equation 1 using the OLS requires the usual conditional independence assumption:

$$E[ManuCap_{i,85} \cdot \varepsilon_i | LocChoice_i, PreTFCond_{i,64}, s_p] = 0. \tag{2}$$

The common critique of such an assumption is the potential omitted variable bias: econometricians are unable to exhaust all the possible factors that are correlated with both the treatment and the outcome variables. In our specific setting, however, given the focus of the TF on military considerations and the resulting explicit criteria for site selection, The assumption represented in Equation 2 seems more plausible. Indeed, findings in Table 4 suggest that conditioning on the location choice criteria, proxies for economic development are not correlated with investment decisions. To the extent some unobserved economic fundamentals might matter for long-run development, they are unlikely to matter beyond the information readily available to the decision maker of the TF. In Online Appendix D, we report analysis based on the OLS, which finds very similar results.

[20] Although the *Plan* did not disclose exact locations of TF projects, in some cases it indicated whether a prefecture was designated to be a center of certain industries.

[21] To compare this coefficient with the baseline result, we note that in 1985, LMS plants account for 3.416% of employment in those TF industrial centers, compared with 0.7488% in non-centers. This would translate to an effect of 7.858 (20.959/(3.416-0.7488)). This is much larger than the baseline estimate. Again, the weak first stage may have exacerbated the small sample bias of the 2SLS estimate.

In Online Appendix D, we also present a case study of a large TF project, whose site selection process resembles a tournament. Following Greenstone et al. (2010), we show that comparing with the runner-up, the winning prefecture had a larger manufacturing sector in 2004. Based on the identification assumption for consistent estimates using tournaments, we devise a novel approach to *directly* test the conditional independence assumption. We find that the omitted variable bias, if any, is negligible.

6 Evidence of Local Agglomeration Economics

The baseline results show that the allocation of manufacturing capacity by the TF has a long-lasting multiplying effect on the structural transformation of local economies. The results so far are silent on the mechanisms through which persistent effects work. In this section, we show evidence that local agglomeration economics plays an important role: the proximity to TF plants made it easier for other plants to enter, grow, and become productive. To highlight the spillover effects, we focus on *private* manufacturing firms. Because the private sector was virtually non-existent at the time when the TF ended, the effects of pre-reform manufacturing presence on the size of the private manufacturing sector can be interpreted as effects on the level growth since the start of the market reforms. We rule out several alternative interpretations and then test specific channels through which agglomeration economics works. Finally, we ask whether the TF, as a *regional* industrial policy, truly benefited local residents.

6.1 Spillover to the Private Manufacturing Sector

Table 8 reports a set of regressions using the benchmark specification, with the dependent variable being various performance measures of the private manufacturing sector. Column 1 suggests that a 1 p.p. increase in the LMS employment share in 1985 increases the private manufacturing employment share in 2004 by 2.1 p.p. Considering the effect on the manufacturing sector as a whole is 2.7 p.p. (Table 6 Column 3), private sector accounts for the bulk of the baseline effect. The sample prefectures had an average private manufacturing employment share of 2.7 percent and a standard deviation of 3.1, so the effect is 78% of the mean or two-thirds of a standard deviation.

Besides local agglomeration economics, there are several alternative explanations for the effect on private manufacturing sector growth. During the market reforms, many state-owned enterprises (SOEs) were privatized. One important concern is that our estimate may be simply picking up changes in ownership of the TF plants. It is worth noting that the way we define the private sector already provides assurance against this concern. Specifically, we exclude all firms with any state shares, and all joint ventures between state and non-state sectors. Since the restructuring and privatization typically took the form of creating joint ventures and mixed-ownership firms (Naughton, 2007), our measure minimizes the risk of picking up the privatized SOEs.

We provide additional evidence to rule out privatization as an alternative explanation. In Column 2 of Table 8, the outcome variable is the employment share from private firms that were established between 1985 and 1998. With this restriction, we include only firms that entered after the market reforms had started but before the large wave of privatization.[22] This measure is therefore less likely to be affected by privatization. As shown in the bottom of the table, about a third (0.875/2.69) of the private manufacturing employment in 2004 are in these firms. The estimated coefficient indicates that if the TF employment share increases by 1 p.p., employment share from these firms increases by 0.73 p.p., which is about 83% of the sample mean. This relative magnitude is similar to that from the benchmark specification in Column 1, suggesting that privatization is unlikely an important concern.

The second piece of evidence relies on the fact that TF plants tended to be large. The smallest 1% of TF plants had more than 200 workers in 1985, and the median had over 1,000 workers, so small private firms in 2004 were unlikely to be privatized TF plants. We use employment share of small private firms (with fewer than 25 workers in 2004) as an alternative outcome measure. Column 3 shows that the effect remains economically and statistically significant. The estimated effect is about 62% (0.018/0.029) of the sample average of the dependent variable.[23]

The third robustness test is to control for changes in employment share of the state-owned manufacturing plants between 1985 and 2004. If the result is driven by privatization, places with more rapid state-sector growth (indicating less privatization) should have a smaller private sector. Column 4 shows that the opposite is true. This additional control variable is endogenous: places more suitable for the development of the manufacturing sector would have a faster-growing manufacturing sector, both state-owned and private. If some of the SOEs were then privatized, it could also explain the growth in the private manufacturing sector. If this is the case, then controlling for the growth in the

[22] Zhu Rongji became China's premier in 1998 and carried out a large wave of privatization of the SOEs.

[23] The 25-worker cutoff is of course arbitrary. The results are robust to alternative cutoffs. It is also possible that smaller TF plants were established and later privatized. Yet we still find a positive and significant effect when we focus on private firms that were established after 1985. Results are also robust to directly controlling for employment from smaller manufacturing plants in 1985, as is shown in Table 7 Column 1.

state-sector employment, which implicitly (and partly) controls for unobserved locational characteristics, should reduce the coefficient associated with the TF investment. Instead, we find that the coefficient becomes larger.

Finally, we exploit cross-industry variation in the state-sector reforms. Since the start of the market reforms, some industries gradually opened up to private entrants, while others, especially those related to infrastructure, natural resources and national defense, are still largely under the state control. Private entry in these industries was limited. If privatization is an important driver for our result, we should expect prefectures receiving more investment in industries under stricter state control have a smaller private sector. To implement this test, we construct a "state-share index" for each prefecture. Specifically, using data from the *2004 Economic Census*, we first calculate the state sector's share in each 2-digit manufacturing industry j at the *national* level, denoted as τ_j. The shares of state-sector employment vary greatly across industries, ranging from 97% in the tobacco industry to 13% in the furniture industry. We then calculate the "state-share index" of the manufacturing capacity in prefecture i as $S_i = \sum_j (\tau_j - \bar{\tau}) s_{ij} / \sigma_\tau$, in which $s_{ij} = ManuCap_{ij,85} / \sum_j ManuCap_{ij,85}$ is the share of industry j in prefecture i's TF-built manufacturing capacity, $\bar{\tau}$ is the average share of state-sector employment across all industries and σ_τ is the standard deviation of τ_j's. S_i is zero if the TF-built manufacturing capacity in prefecture i is evenly distributed across all industries, and is higher if more capacity in prefecture i is in industries dominated by the state sector. We include S_i and its interactive term with $ManuCap_{i,85}$ in the regression, instrumenting the interactive term with the interaction between the log distance to existing railways and the state share index. If privatization drives the result, the coefficient associated with the interactive term would be negative and the coefficient of $ManuCap_{i,85}$ would become much smaller. Column 5 shows that this hypothesis is not supported by the data. The interactive term does have a negative sign but is not statistically different from zero; the coefficient associated with $ManuCap_{i,85}$ is quantitatively similar to that in Column 1. One concern for this specification is the weak first stage. In Online Appendix Table D.1, we perform the same exercise using OLS. The main messages are the same.

Another alternative explanation for our results is continued government subsidies. Specifically, even after series of market reforms, many TF plants remained under the state control. They might also continue to receive direct and indirect subsidies from the government, which might spillover into the private sector and drive our results. These subsidies can take the form of direct transfer from higher-level government, barriers to the entry of private competitors, contracts from government, and infrastructure investment in the prefecture. Columns 4 and 5 allow us to alleviate this concern. If continued government subsidies matter, once we control for the state-sector employment growth, which likely correlates strongly with subsidies, the estimated effect of manufacturing capacity would be much smaller. Column 4 suggests otherwise. In addition, since continued government subsidies are likely larger in industries that are still under the state control, to the extent that those subsidies are driving our results, controlling for the "state-share index" would substantially reduce the coefficient associated with $ManuCap_{i,85}$. Results in Column 5 reject this hypothesis.

To further demonstrate the existence of local agglomeration economics, we look directly at the entry and the productivity of private firms. In Column 6, we look at the entry margin using firm counts per 100 counts of employment as the dependent variable. According to this estimation, as the employment share from LMS plants increases by 1 p.p., there is 16 additional private manufacturing firms for every 100,000 workers in the prefecture in 2004. This is roughly a 42 percent increase from the sample mean, so a significant share of the effect on private manufacturing employment could be accounted for by firm entries.

The remaining two columns report the effects on efficiency and wage in the manufacturing sector. In Column 7, we use firm-level data from the *2004 Annual Manufacturing Firm Survey*, which provides detailed information on firms' balance sheet. We estimate the same specification, using firm-level TFP (constructed as the Solow residual) as the dependent variable and weighting each firm by its share of the total private manufacturing employment in the prefecture. The standard error is clustered at the prefecture level. Column 7 shows that a 1 p.p. increase in the employment share of the LMS plants increases the TFP of local private manufacturing plants by about 10 log points.

In Column 8, we use individual monthly wage from the *2005 Inter-Decennial Population Census* as the dependent variable. All workers in the manufacturing sector are included in the regression sample—the data does not allow us to distinguish workers in the private sector from those in the state sector. Each observation is weighted by the inverse of the number of manufacturing workers in the prefecture, and the standard error is clustered at the prefecture level. The estimate shows that 1 p.p. increase in the employment share of the LMS plants increases the local manufacturing wage by about 3 log points.

We interpret the findings in Table 8 as evidence that TF plants generated positive spillovers to local private firms through agglomeration economics— by passing on knowledge about production, market, and management, TF plants made it easier for new firms to enter and grow. Indeed, many of China's post-reform entrepreneurs spent years in a state firm before starting their own shop. Agglomeration economics may also work through input and information sharing: proximity to TF plants could provide new firms access to suppliers and buyers, as well as a pool of workers with relevant skills. We explore these mechanisms in detail below.

6.2 Inter-Industry Linkages

To directly assess agglomeration economics through inter-industry linkages, we estimate the effect of TF-built manufacturing capacity at the prefecture-industry level using the following specification:

$$EmpShare_{ij,04} = \rho \sum_{k} IndLink_{jk} \cdot ManuCap_{ik,85} + \lambda_i + I_j + \varepsilon_{ij}, \tag{3}$$

In Equation 3, i indicates a prefecture, j indicates one of the nineteen 2-digit manufacturing industries. $EmpShare_{ij,04}$ is the employment of private manufacturing firms in 2004 in industry j as a share of prefecture i's total employment. $IndLink_{jk}$ is some directional measure of linkages between industries j and k (influence from k to j). λ_i is the prefecture fixed effect. I_j is the industry fixed effect. The prefecture fixed effects control for the total LMS manufacturing employment in each prefecture, so the specification exploits only variation in industrial compositions of TF-built manufacturing capacity across prefectures. Our instrumental variable is at the prefecture level and is thus no longer relevant. We therefore estimate Equation 3 using OLS. The identification assumption is that the unobserved prefecture heterogeneity relevant for long-run growth do not affect industrial compositions of the TF-built manufacturing capacity.

To investigate the effect on the efficiency of private firms, we run the following specification:

$$TFP_{fij,04} = \rho \sum_{k} IndLink_{jk} \cdot ManuCap_{ik,85} + \lambda_i + I_j + \varepsilon_{fij}, \tag{4}$$

where f indicates a firm. Each firm is weighted by its number of workers as a share of prefecture total employment in the private manufacturing sector. Standard errors are clustered at the prefecture level.

Recent studies in urban economics have identified a few channels through which agglomeration economics works among local firms. These channels include knowledge spillovers, labor sharing, and input sharing (Rosenthal and Strange, 2004). All these channels are likely to be stronger for firms within the same industry. Column 1 in Table 9 Panel A shows that a 1 p.p. higher share of LMS employment share increases the private employment share in the same 2-digit industry by 0.3 p.p. This number is around 14% of the baseline result reported in the first column of Table 8. With a total of 19 industries, the average effect on one industry would account for about 5% of the total effect. Within-industry spillovers account for a disproportionately large share of the overall spillover effect.

More generally, following the literature, we use three broad sets of measures that capture different channels of local agglomeration economics: labor market pooling, technological spillover, and intermediate input linkages. To circumvent data limitations and to avoid endogeneity issues, we use measures based on the U.S. data from Greenstone et al. (2010), which capture aspects of industrial relations that are exogenous to the specialties of the Chinese economy. Labor market pooling is measured as the intensity of worker flows between industries (worker transitions from the *Current Population Survey*), technological spillovers are measured using patent citation patterns and the input-output table for R&D expenditure (technology input/output), and intermediate input linkages are captured by the input-output table. Detailed definitions of these measures are listed in Online Appendix Table A.2. The measures are standardized so the coefficients are comparable across columns. Columns 2 to 7 of Table 9 show that all these measures are positively correlated with employment size and many of them are statistically significant, which suggests that industry linkages play a role.

Panel B repeats the same exercises on firm efficiency. We find that all coefficients have the expected sign, though they are not statistically significant at the conventional levels. For example, according to the point estimate in Column 1, a 1 p.p. increase in the employment share from the LMS plants raises the productivity of private firms in the same industry by 3.8 log points. This is an economically meaningful effect despite not precisely estimated.

6.3 Local Structural Transformation versus Inter-Regional Migration

The results so far indicate that the manufacturing capacity built during the TF increased local manufacturing employment in the long run, and it exhibits a multiplying effect. This increase could come from two sources: structural transformation of workers within the local economy, or reallocation of manufacturing workers across regions. These two channels have very different implications for the aggregate effects of the TF (Gottlieb and Glaeser, 2008; Kline and Moretti, 2014a). Using individual-level data from the *2005 Intra-decennial Population Census*, we assess the importance of these two channels.

We first look at whether places that had a larger pre-reform manufacturing capacity attracted more migrants after the reforms. We define migration based on where one's *hukou* is registered. China's *hukou* system by-and-large ties each person to one's birthplace. We classify an inward migrant if one works in the current prefecture but one's *hukou* is registered in a different prefecture. Defined this way, migrant workers accounted for about 1.1% of the workforce in the average sample prefecture in 2005. Column 1 of Table 10 shows that TF investment has essentially no effect on inward migration. One possible reason for that is places in the TF region remained much less developed than other parts of the country. For workers from a prefecture in the TF region that did not receive much investment, once they pay the fixed cost to migrate, they are most likely to migrate to more prosperous regions in the coast, where wages are much higher, rather than to migrate to another city within the TF region.[24]

After ruling out recent migration as the source of increased manufacturing employment share, we provide direct evidence for structural transformation of the local workforce by looking at the impact of TF-built manufacturing capacity on the share of employment in manufacturing among workers with a local rural *hukou*— workers who are registered with a rural community in the same prefecture.[25] Column 3 shows that increasing the LMS employment share by 1 p.p. raises the manufacturing share among those workers by 0.23 p.p. The effect is significant at the 10% level and amounts to 11% of the sample mean. Together, these results suggest that the growth in the local manufacturing sector since the market reforms was driven by structural transformation of the local economy rather than migration of manufacturing workers across prefectures.

Indeed, TF-built manufacturing capacity transformed the local economy in a wide range of dimensions beyond the manufacturing sector. Table 11 shows a 1 p.p. increase in the LMS employment share in 1985 raises the prefecture's urban rate by 3 p.p. (10% from the sample mean), share of workers with a college degree by 0.85 p.p. (34% of the mean), and per capita GDP by 11%.

7 Aggregate Implications

The TF can be thought as a transfer of manufacturing capacity to the hinterlands from other parts of the country. In this section, we investigate its implications on the

[24]Thousands of people migrated, some forced to, to the TF region during the 1960s and 1970s. Those migrants are mostly factory workers that helped staffing the new TF plants. Had they remained in the region, they would have obtained a local *hukou* long before 2005. Those with *hukou* registered outside of the place-of-residence are mostly new migrants.

[25]There are two types of *hukou*: rural and urban. A rural *hukou* indicates that a person was born in a rural community and was originally tied to agriculture. When rural workers find jobs in the urban sector, some are able to change their *hukou* registration into urban. However, urban *hukou* is often rationed and many manufacturing workers have to keep the rural *hukou*. Our estimate is based on these workers. Because of the possibility of switching to urban *hukou* (this was made easier after the late 1990s), our estimate likely understates TF's impact on rural-to-urban transformation.

aggregate economy. A comprehensive cost-and-benefit analysis is beyond the scope of our analysis here, mostly because costs associated with the TF are unknown to us. We also take as given the amount of manufacturing capacities to be allocated by the government, thus abstracting away from sectoral allocation and distortive taxations in order to fund government industrialization initiatives.

We investigate TF's impacts on the aggregate efficiency of the manufacturing sector as well as regional equity. To this end, we develop a simple model to identify necessary conditions for the TF to be efficiency-improving in the aggregate. In the model, the TF benefits the local economy because it pulls agricultural workers into manufacturing, which has a higher labor productivity. But local gains come at potential aggregate losses because the opportunity cost of the TF is the return to investment in other, more productive, regions.[26] The key condition for the TF to improve the aggregate efficiency is one of heterogeneous effects by initial development levels. It is likely to improve the aggregate efficiency if the spillover effects are especially strong in less developed regions. To preview the results, our finding does not support this condition, and we conclude that the TF reduces the aggregate efficiency. On the other hand, the TF did significantly reduced regional inequality. Had the manufacturing capacity in the TF moved to the productive coastal area, the gap in per capita GDP would have been 20% larger than it actually was in 2010.

7.1 The Conceptual Framework

We delegate the full version to Online Appendix C and summarize only the key elements of the model here. The model is at the prefecture level. Each prefecture is endowed with a number of workers, denoted L. To capture the constraint on worker mobility in China and to be consistent with our empirical evidence (Table 10 Column 1), we assume that workers are immobile across prefectures. In order to speak to the gains from structural transformation, we assume that there are two production sectors. All workers originally work in the rural (agricultural) sector, and each needs to pay a $\frac{\tau-1}{\tau}$ ($\tau > 1$) fraction of urban income to switch to the urban (manufacturing) sector. Agricultural production in each region is subject to decreasing return to scale in labor input with land

[26] This is related to, but different from, a point about the aggregate effect of place-based policies in the context with free mobility (Kline and Moretti, 2014b; Gottlieb and Glaeser, 2008). The main message in that context is that agglomeration effect in one region from a place-based policy comes at the expense of reduced agglomeration in other regions. In our context, because of migration frictions, the reallocation is best thought of as between agricultural and manufacturing sectors *within* a local economy. Yet a similar logic applies—without the TF, the investment could have gone to other potentially more productive regions, so to investigate the aggregate effect of TF investment, we need to take into account this "opportunity cost."

being the fixed input. Manufacturing production is modeled in a monopolistic competitive framework. Private manufacturing firms pay a fixed entry cost and then produce using a linear technology with labor as the only input; on the other hand, state-owned firms are modeled as firms existing in the economy with the fixed setup cost already paid for, before private firms decide whether to enter. To capture the existence of local agglomeration economics, we assume that local manufacturing productivity increases in the number of state-owned firms in the economy before private entry.

We can write the per-capita output in prefecture i as

$$y_i = \underbrace{\frac{w_i^A l_i^A}{\alpha}}_{\text{Agricultural income}} + \underbrace{w_i^M l_i^M}_{\text{Non-agricultural labor income}} + \underbrace{\frac{w_i^M N^{ini} f}{L_i}}_{\text{TF firm profits}}, \tag{5}$$

where w_i^s, l_i^s, $s \in \{A, M\}$ are wage and the employment *share* in sector s, respectively. α is the labor share in agricultural output; f denotes the fixed entry cost for private firms. Equation 5 simply states that the average income is the weighted sum of agricultural and manufacturing income, and per-capita profit from the state firms.[27]

Given Equation 5, the effect on per-capita income of a marginal increase of the investment in a prefecture, δ, can be decomposed into three channels—wage, reallocation, and profit—as follows:

$$\begin{aligned}\frac{\partial \log(y_i)}{\partial \delta} &= \underbrace{l_i^A \frac{w_i^A}{\alpha y_i}\frac{\partial \log(w_i^A)}{\partial \delta} + l_i^M \frac{w_i^M}{y_i}\frac{\partial \log(w_i^M)}{\partial \delta} + \frac{f w_i^M}{L_i y_i} N_i^{ini} \frac{\partial log(w_i^M)}{\partial \delta}}_{\text{wage channel}} \\ &+ \underbrace{\frac{(w_i^M - \frac{w_i^A}{\alpha})}{y_i}\frac{\partial l_i^M}{\partial \delta}}_{\text{reallocation channel}} + \underbrace{\frac{f w_i^M}{L_i y_i}}_{\text{profit channel}}\end{aligned} \tag{6}$$

The wage channel consists of three components. The first two components capture the effect of the investment through raising agricultural and manufacturing wages directly. The third component captures the indirect effect of a wage increase on profits of the state firms—because profits are proportional to the manufacturing wage, as the wage increases, profits increase, too. The reallocation channel captures the effect on income by reallocating workers from agriculture to manufacturing. The strength of this component depends on two factors: the income gap between the two sectors, which is equal to

[27] The free entry condition implies that private firms make zero profit. The migration costs rural workers have to pay can be seen as a transfer to the local government, so they do not show up in the output accounting.

$(w_i^M - \frac{w_i^A}{\alpha})$, and the effect of the investment on reallocation, $\frac{\partial l_i^M}{\partial \delta}$. Finally, the profit channel simply captures the profit of the marginal firm being reallocated. This channel is more important when manufacturing wage w_i^M is higher relative to the average income y_i.

We use the model to evaluate the aggregate effect of the TF. Consider two prefectures, one inside and the other outside the TF region (denoted as TF and NTF, respectively). The two prefectures have the same population ($L_{TF} = L_{NTF}$), but due to its better geographic conditions, NTF has a higher initial manufacturing productivity. Consistent with the historical background, we assume that all state firms are initially in NTF, that is, $N_{TF}^{ini} < N_{NTF}^{ini}$. We view the TF as reallocating a small number of plants, δ, from NTF to TF.

From Equation 6, the *aggregate* effect across the two prefectures of reallocating δ from NTF to TF is given by:

$$
\begin{aligned}
\frac{\partial y_{sum}}{\partial \delta} &= y_{TF}\frac{\partial \log(y_{TF})}{\partial \delta} - y_{NTF}\frac{\partial \log(y_{NTF})}{\partial \delta} \\
&= \underbrace{[\frac{f w_{TF}^M}{L_{TF}} - \frac{f w_{NTF}^M}{L_{NTF}}]}_{\text{Profit shifting}} + \underbrace{[(w_{TF}^M - \frac{w_{TF}^A}{\alpha})\frac{\partial l_{TF}^M}{\partial \delta} - (w_{NTF}^M - \frac{w_{NTF}^A}{\alpha})\frac{\partial l_{NTF}^M}{\partial \delta}]}_{\text{Net reallocation effect}} \\
&\quad + \underbrace{[(w_{TF}^M l_{TF}^M + w_{TF}^M \frac{f N_{TF}^{ini}}{L_{TF}} + w_{TF}^A \frac{l_{TF}^A}{\alpha})\frac{\partial \log(w_{TF}^M)}{\partial \delta} - (w_{NTF}^M l_{NTF}^M + w_{NTF}^M \frac{f N_{NTF}^{ini}}{L_{NTF}} + w_{NTF}^A \frac{l_{NTF}^A}{\alpha}))\frac{\partial \log(w_{NTF}^M)}{\partial \delta}]}_{\text{Wage effect}}.
\end{aligned}
\tag{7}
$$

The second equality in the expression uses the migration condition for workers ($\tau w_i^A = w_i^M$, which implies $\frac{\partial \log(w_i^A)}{\partial \delta} = \frac{\partial \log(w_i^M)}{\partial \delta}$). As Equation 7 makes clear, whether the TF leads to aggregate efficiency gains boils down to whether the sum of the three terms are positive. In the following, we empirically determine the signs of each of these three terms.

7.2 Quantifying Aggregate Efficiency

7.2.1 The Profit-Shifting Channel

The profit-shifting effect depends only on which prefecture has a higher wage. Intuitively, the profit of a plant depends on local productivity. Plants in regions with higher labor productivity (hence a higher wage) will make a larger profit. As is shown in Table 12, manufacturing productivity and wages are both higher in the NTF region. The difference in average efficiency ranges between 15% to 30%—by investing in a less productive region, the TF reduces profits of the state firms, resulting in a first-order efficiency loss.

7.2.2 The Reallocation Channel

We now turn to the reallocation channel. According to Equation 7, two terms determine this channel. First, the agricultural-manufacturing income gap, $(w_i^M - \frac{w_i^A}{\alpha})$. This term captures the increase in average income by transforming an additional worker from the agricultural sector to the manufacturing sector. Table 12 shows that this difference is 395 *yuan* per month in the TF region, and 479 *yuan* in the non-TF region—the gains from reallocating one worker from the agricultural to the manufacturing sector does not generate larger gains in the TF region.

The second term important for the reallocation effect is the marginal effect of investment on the reallocation of workers across sectors, $\frac{\partial l_i^M}{\partial \delta}$. One rationale for transferring resources to less developed regions is that these regions might be stuck in a "poverty trap." For example, workers at the subsistence level may not be able to save enough to invest in the manufacturing sector. They may also lack the necessary skills and knowledge to run a manufacturing plant. In such scenarios, transfers to these regions may help overcome these constraints and jump start a virtuous cycle of sustained industrialization, and $\frac{\partial l_i^M}{\partial \delta}$ could be larger in these regions.

Our empirical setting does not allow us to estimate this term in the NTF region. Instead, we exploit variation in levels of initial economic development *within* the TF region and test whether the reallocation channel is stronger in less-developed regions. We split our sample into halves according to a prefecture's urban rate in 1964. Although prefectures in our sample were all relatively undeveloped in 1964, there is still meaningful variation in initial economic conditions. The median urban rate among sample prefectures is 6.5%. The average is 9.1% for prefectures above the median, and 5.1% for prefectures below it. Panel A of Figure 4 shows the relationship between TF-built manufacturing capacity and the 2004 private manufacturing employment share (both residuals from regressions with the same covariates as in the baseline specification) for the two groups. The blue dots represent sample prefectures with a below-median urban rate in 1964 while the red crosses represent the those above the median. First notice that dots and crosses are similarly distributed along the horizontal axis—TF investment is generally uncorrelated with the initial economic development. In addition, there are almost identical marginal effects of TF-built manufacturing on private manufacturing employment between the two groups.[28]

[28] The slope for the linear fit is 1.68 (0.27) for the above-median sample and 1.58 (0.42) for the below-median sample. We use OLS here because the IV is weak in split samples. One may worry that the survival bias might affect our conclusion. To the extent that the survival bias inflates the measure of investment for more prosperous regions, it will lead to an underestimate in the spillover effect for the above-median

One may suspect that the effect of investment could be non-linear: only large-enough investment could pull an economy out of the poverty trap. In the figure, we also fit a flexible functional form for both groups and find no evidence for nonlinear effects. Taken together, the evidence suggests that the spillover effects are neither stronger in initially poorer regions nor dependent on the intensity of the investment, that is, $\frac{\partial l_{TF}^{M}}{\partial \delta} = \frac{\partial l_{NTF}^{M}}{\partial \delta}$.

7.2.3 The Wage Channel

Finally, we turn to the wage channel, which captures a general equilibrium effect: by increasing labor demand from the manufacturing sector, the government investment increases local wages. If this channel is larger in the TF region, then the TF could still have a positive aggregate impact. We would like to compare $\frac{\partial \log w_{TF}^{M}}{\partial \delta}$ and $\frac{\partial \log w_{NTF}^{M}}{\partial \delta}$, but again, we do not have a causal estimate for the latter. Following the same approach as earlier, we look within the TF region to see if the wage effect differs by the prefecture's initial level of development.

We split the TF prefectures according to the 1964 urban rate and separately estimate the effect of TF investment on local manufacturing wages for each group. Panel B of Figure 4 shows that the two fitted lines almost perfectly overlap with each other, indicating that the wage effect is not stronger in initially less developed regions.[29] Panel C shows that the same conclusion holds for firm productivity.

We now turn to the difference between TF and NTF in $(w_i^M l_i^M + w_i^M \frac{fN_i^{ini}}{L_i} + w_i^A \frac{l_i^A}{\alpha})$. Since the TF region had no manufacturing plants in the beginning and still had fewer plants than the NTF region by the end of the TF period, we have $w_{NTF}^{M} \frac{fN_{NTF}^{ini}}{L_{NTF}} > w_{TF}^{M} \frac{fN_{TF}^{ini}}{L_{TF}}$. Using $l_i^A + l_i^M = 1$, the remaining terms could be arranged into $l_{TF}^{M}(w_{TF}^{M} - \frac{w_{TF}^{A}}{\alpha}) - l_{NTF}^{M}(w_{NTF}^{M} - \frac{w_{NTF}^{A}}{\alpha}) + \frac{w_{TF}^{A} - w_{NTF}^{A}}{\alpha}$. Because $w_{NTF}^{A} > w_{TF}^{A}$, $l_{NTF}^{M} > l_{TF}^{M}$, and $w_{NTF}^{M} - \frac{w_{NTF}^{A}}{\alpha} > w_{TF}^{M} - \frac{w_{TF}^{A}}{\alpha}$ (Table 12), the net wage effect is also negative.

sample. Our conclusion that the spillover effects are not stronger in the underdeveloped regions will thus be reinforced once the bias is accounted for.

[29] One valid concern is that we might not be able to find a larger effect in poorer regions because our wage estimates lack accuracy. An alternative way of getting at the wage effect is to rely on the structure of the model and exploit the relationship between the wage effect and the employment effect. Notice $log(w_i^A) = log(w_i^M) - log(\tau)$ and $\frac{\partial \log w_i^A}{\partial \delta} = \frac{\partial \log w_i^A}{\partial \log L_i^A} \frac{\partial \log L_i^A}{\partial \delta} = (\alpha - 1) \frac{\partial l_i^A}{\partial \delta} \frac{1}{l_i^A}$, in which the second equality follows from Equation C.3 in the Appendix. Our employment effect estimate suggests $\frac{\partial l_{TF}^{A}}{\partial \delta} = \frac{\partial l_{NTF}^{A}}{\partial \delta}$. Therefore if any difference, the wage effect should be stronger in prefectures with a lower share of employment in agriculture, l_i^A, namely the NTF region.

7.3 Implications on Regional Inequality

Like most location-based policies, an important aspect of the TF is its role in reducing regional inequality. Although the TF was motivated mostly by national defense considerations, developing the economically backward region was also an implicit goal.[30] Economic development across Chinese regions is highly unequal. Being the only major source of investment in the hinterlands until the 2000s, the TF played an important role in mitigating the large regional inequality.

To calculate the extent the TF has reduced regional inequality, we use GDP per capita as the measure of regional income and compare prefectures in our sample with prefectures in the coastal provinces.[31] In 2010, GDP per capita was about 16,630 *yuan* in the TF provinces and 29,124 *yuan* in coastal provinces. The regional income gap was 175%.

The starting point of our calculations is Table 11 Column 3, which shows that a one p.p. increase in the employment share from the LMS plants leads to an 11% increase in the prefecture's per capita GDP in 2010. We assume this effect is uniform across locations (Figure 4 shows that there is no evidence of heterogeneous effects based on initial development levels).

The sample average of the employment share from the LMS plants was 1.11. According to estimates in Table 11 , GDP per capita in TF provinces would be 12.2% ($1.11 \times 11\%$) lower, to 14,601 *yuan*. Had the same manufacturing capacity moved to the coast area in a prefecture with the same population, its GDP per capita would growth to 32,677 *yuan* ($29,124$ *yuan* $\times\ 112.2\%$). The regional income gap would be 224%. Therefore, the TF reduced regional gap in per capita GDP by 22% ($(224\% - 175\%)/224\%$).

The above back-of-the-envelope calculations maintain several assumptions. First, population is assumed to be fixed. In 2005, as much as 5% of the workforce in the TF provinces work outside of the region. If regional income gaps were to become larger, more people in the TF region would migrate to the coastal area, which in turn tends to reduce regional inequality. However, migration is highly costly in China (Fan, 2019). As the existing large regional disparity in GDP per capita suggests, migration alone would not be enough to substantially close the income gap across regions. The TF thus played an indispensable role in reducing regional inequality. Second, construction of the same factories would

[30] As a fervid advocate of the TF, Mao was susceptible to the regional equality argument. Chen (2003) documents an interesting incident. In a visit to Tianjin, Mao intended to test the popular support of the TF by asking the audience "if the war does not break out, do you think the TF will be a waste?" He was apparently content with the crowd answering that the TF would be a worthy investment to boost the economic development of the country's hinterlands.

[31] The coastal provinces include Beijing, Tianjin, Hebei, Shandong, Jiangsu, Shanghai, Zhejiang, Fujian, and Guangdong. We exclude provincial capitals to be consistent with the baseline sample.

likely cost less in the coastal area and the factories will be more efficient. The same amount of investment, if made in the coastal area, is likely to generate larger output, and the regional inequality may be even larger.

7.4 Discussion of Findings on the Aggregate Implications

The analyses in this section suggest that although the TF benefited recipient local economies, these benefits primarily represent regional redistribution, and that the TF likely reduced the efficiency of the aggregate economy. Policymakers face an efficiency-equity tradeoff. Of course, whether it is worthwhile in giving up at least 15-30% of efficiency in exchange for a regional inequality that is about 20% smaller depends on how much weight the society puts on the two competing goals.

An important finding leading to this conclusion is the lack of heterogeneous spillover effects either by the level of local economic development or by the size of investment. This result is related to a large literature aiming at detecting poverty traps. Most of those studies use country-level data (see Easterly, 2006, for a recent example). By exploiting spatial variation within a country, we do not find evidence for a key implication of models with a poverty trap. Our results are in line with the evidence from the TVA. Kline and Moretti (2014a) find that agglomeration forces are linear and that gains in the TVA region were offset by losses in the rest of the country. Our paper further shows that, even in a much earlier stage of economic development, when the lack of capital stock and know-how is more likely to impose a barrier for industrialization, the agglomeration, though positive, is not particularly larger.

There are two caveats in interpreting our results. First, to reach the above conclusions we use within-TF variation in initial development levels. To the extent that the sample prefectures were all largely undeveloped before the TF constructions, it is possible that the range of variation in our sample does not cover the range that could exhibit heterogeneous effects. Second, we have ignored the potential congestion effects in more developed areas. Additional investment could generate two competing effects: a positive agglomeration effect and a negative congestion effect. One would expect that the agglomeration economics dominates at early stages of economic development, while the congestion effect dominates at more developed stages. If the congestion effect dominates the agglomeration effect in China's coastal area, there is an additional efficiency argument to reallocate resources to less developed regions. In Appendix E.4, we discuss this possibility and show evidence that at least by the 2000s, the congestion effect does not dominate the agglomeration effect in China's most developed cities.

8 Conclusions

The Construction of the Third Front provides a unique setting for estimating the long-run effects of government policies that direct investment to the manufacturing sector in less-developed regions. Our results show that such policies are effective in stimulating structural transformation of the local economy. More than twenty years after the start of the market reforms, places that had larger initial manufacturing capacity due to the TF had a larger and more productive manufacturing sector. By focusing on the private sector and exploiting cross-industry relations, we show that local agglomeration economics played an important role in generating these long-run effects. Despite its substantial positive impacts on the local economy, and its role in reducing regional inequality, the TF did not increase the aggregate efficiency. Central to this conclusion is the finding that the spillover effects of the investment were not stronger in less developed regions. Our finding highlights tradeoff between equity and efficiency in location-based development policies in settings where workers face mobility constraints.

The TF's unique historical background and peculiar site selection criteria provide a rare opportunity to credibly answer an important research question. Yet this context also raises valid concerns over external validity. Indeed, poorly designed and hastily implemented, the true costs associated with generating the initial manufacturing capacity must be much higher than other place based policies, and this paper does not intend to conduct a comprehensive cost-and-benefit analysis of the program. Nevertheless, the TF allows us to estimate the long-run spillover, a key parameter for many place-based policies. Our finding that temporary policies promoting manufacturing sector in poor rural regions can generate long-lasing multiplier effects, but that such spillover effects are not stronger in less developed regions are likely relevant for regional development programs elsewhere.

This paper focuses on the long-run outcomes and overlooks the dynamics of regional development after China's economic reforms. By combining more evidence on the transition path with a structural model, one will be able to explore interesting patterns of regional dynamics. Through the lens of a model, one can also simulate the aggregate and distributional effects of potentially better-designed policies, such as those that try to reduce mobility costs. We leave these explorations to future research.

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.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian, "On the Road: Access to Transportation Infrastructure and Economic Growth in China," *Journal of Development Economics*, 2020, (102442).

Baum-Snow, Nathaniel, Loren Brandt, J Vernon Henderson, Matthew A Turner, and Qinghua Zhang, "Roads, Railroads and Decentralization of Chinese Cities," *Review of Economics and Statistics*, 2017, *99* (3), 435–448.

Beason, Richard and David E Weinstein, "Growth, Economies of Scale, and Targeting in Japan (1955-1990)," *Review of Economics and Statistics*, 1996, pp. 286–295.

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

Central Documentary Office, "Outlines for the Third Five-Year Plan (1965) "guanyu disange wunian jihua anpai qingkuang de huibao tigang"," in "Selected Important Documents since the Founding of the People's Republic of China," Vol. 20, Central Documentary Press, 1992.

Chen, Donglin, *The Third Front: China's Western Development as a Preparation for War. (sanxian jianshe: beizhan shiqi de xibu kaifa) (in Chinese)*, China Central Party School Press, 2003.

Criscuolo, Chiara, Ralf Martin, Henry G Overman, and John Van Reenen, "Some causal effects of an industrial policy," *American Economic Review*, 2019, *109* (1), 48–85.

Easterly, Willian, "Reliving the 1950s: the Big Push, Poverty Traps, and Takeoffs in Economic Development," *Journal of Economic Growth*, 2006, *11* (4), 289–318.

Faber, Benjamin, "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System," *Review of Economic Studies*, 2014, *81* (3), 1046–1070.

Fan, Jingting, "Internal Geography, Labor Mobility, and the Distributional Impacts of Trade," *American Economic Journal: Macroeconomics*, 2019, *11* (3), 252–88.

Glaeser, Edward L and William R Kerr, "Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain?," *Journal of Economics and Management Strategy*, 2009, *18* (3), 623–663.

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*, 2014.

Gottlieb, Joshua D. and Edward L. Glaeser, "The Economics of Place-Making Policies," *Brookings Papers on Economic Activity*, 2008, *1*, 155–239.

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.

Hirschman, Albert O, *The Strategy of Economic Development*, Vol. 10, Yale University Press New Haven, 1958.

Jedwab, Remi, Edward Kerby, and Alexander Moradi, "History, Path Dependence and Development: Evidence from Colonial Railroads, Settlers and Cities in Kenya," *Economic Journal*, 2017, *127* (603), 1467–1494.

Juhász, Réka, "Temporary protection and technology adoption: Evidence from the napoleonic blockade," *American Economic Review*, 2018, *108* (11), 3339–76.

Kline, Patrick and Enrico Moretti, "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *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.

Lu, Yi, Jin Wang, and Lianming Zhu, "Place-Based Policies, Creation, and Agglomeration Economies: Evidence from China's Economic Zone Program," *American Economic Journal: Economic Policy*, 2019, *11* (3), 325–60.

Marshall, Alfred, *Principles of Economics*, London: Macmillan and Co., Ltd., 1890.

Murphy, Kevin M, Andrei Shleifer, and Robert W Vishny, "Industrialization and the Big Push," *The Journal of Political Economy*, 1989, *97* (5), 1003–1026.

Naughton, Barry, "The Third Front: Defense Industrialization in the Chinese Interior," *The China Quarterly*, 1988, *115*, 351–86.

_ , *The Chinese Economy: Transitions and Growth*, MIT Press, 2007.

Neumark, David and Helen Simpson, "Place-Based Policies," in "Handbook of Regional and Urban Economics," Vol. 5, Elsevier, 2015, pp. 1197–1287.

Redding, Stephen J and Daniel M Sturm, "The Costs of Remoteness: Evidence from German Division and Reunification," *American Economic Review*, 2008, *98* (5), 1766–1797.

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

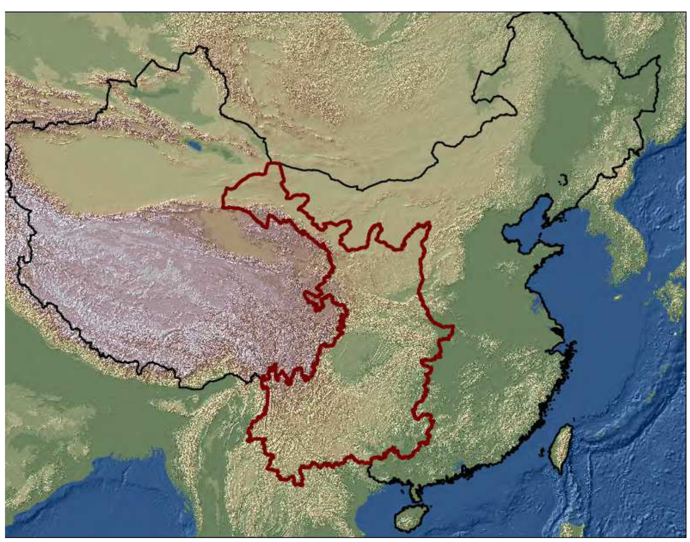
Rosenthal, Stuart S and William C Strange, "Evidence on the nature and sources of agglomeration economies," *Handbook of Regional and Urban Economics*, 2004, *Vol. 4*, 2119–2171.

Sequeira, Sandra, Nathan Nunn, and Nancy Qian, "Immigrants and the Making of America," *Review of Economic Studies*, 2020, *87* (1), 382–419.

Wang, Jin, "The Economic Impact of Special Economic Zones: Evidence from Chinese Municipalities," *Journal of Development Economics*, 2012.

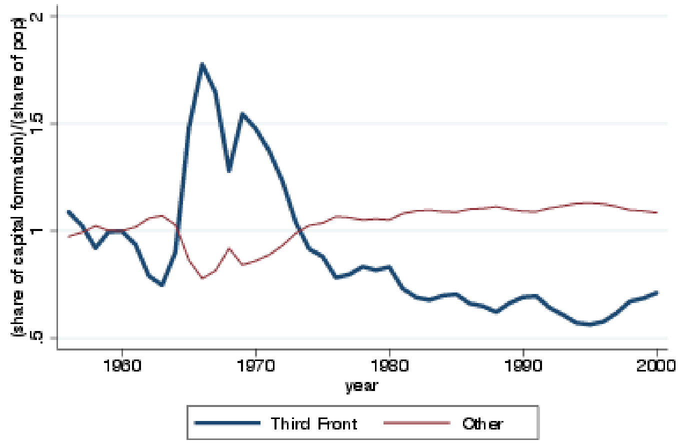
Figures and Tables

Figure 1: Third Front Region



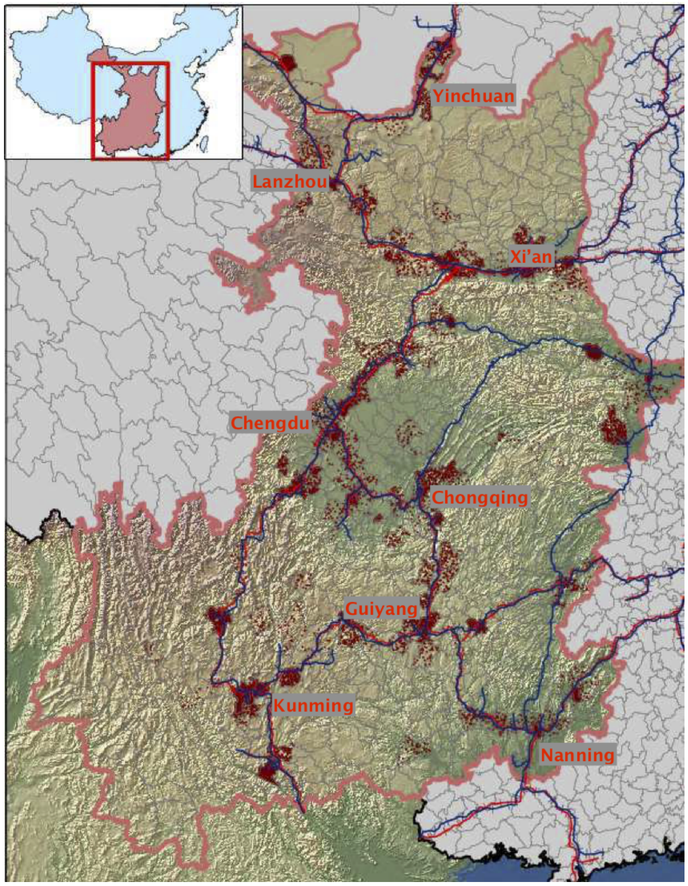
Notes: The Third Front region is shown inside the red line.

Figure 2: Capital Formation by Region



Notes: Relative investment intensity is measured as the region's share of national capital formation divided by its share of national population. Source: *60-year Statistical Summary* constructed by the National Bureau of Statistics of China.

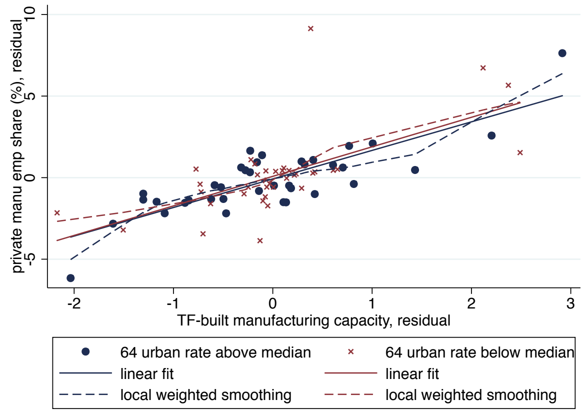
Figure 3: Spatial Distribution of Manufacturing Employment



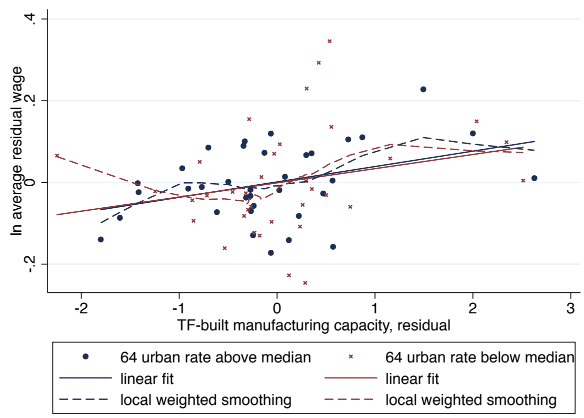
Notes: Each red dot indicates 25 workers from large- and medium-size manufacturing firms in 1985. Red lines show railway lines that existed or were under construction in 1962. Blue lines show existing railway lines in 1980. In segments of existing railway lines, red lines do not always perfectly overlap with blue lines. This is due to measurement errors introduced in digitizing railway lines from hardcopy maps. Provincial capitals are annotated.

Figure 4: Heterogeneous Effects of TF Investment

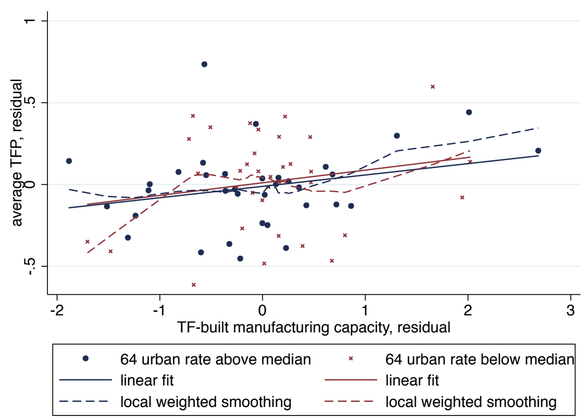
Panel A: Private Manufacturing Employment Share, 2004



Panel B: Average Manufacturing Wage, 2005



Panel C: Average Private Manufacturing Firm TFP, 2004



Notes: The sample prefectures are divided into two equal-sized groups based on whether the prefecture's urban rate in 1964 was above or below the sample median. For each group, the dependent variable and the TF-built manufacturing capacity are first regressed on a set of location-choice variables and initial economic conditions, residuals are plotted. Solid lines are linear fits. Dashed lines are local weighted smoothing.

Table 1: Location Characteristics by Plants' Time of Opening: Nationwide

time of opening	before TF (1)	during TF (2)	(2)-(1) (3)
Panel A: County characteristics			
= 1 if in TF region	0.13	0.32	0.19
	(0.33)	(0.47)	(0.01)
in provincial capital prefecture	0.44	0.21	-0.23
	(0.5)	(0.41)	(0.01)
average slope	1.23	1.92	0.69
	(1.42)	(1.76)	(0.04)
log mean elevation	4.29	5.27	0.97
	(1.88)	(1.72)	(0.05)
log distance to provincial capital	2.71	3.93	1.22
	(2.49)	(2.11)	(0.06)
urban rate, 1964	60.1	35.52	-24.58
	(80.81)	(55.51)	(1.68)
log population density, 1964	6.68	5.76	-0.92
	(1.49)	(1.35)	(0.04)
industrial employment share, 1936	54.55	18.52	-36.03
	(105.77)	(59.01)	(1.99)
Panel B: Sector			
mining	0.06	0.04	-0.02
	(0.23)	(0.2)	(0.01)
light industries	0.3	0.22	-0.08
	(0.46)	(0.42)	(0.01)
power/water	0.07	0.06	-0.01
	(0.25)	(0.24)	(0.01)
chemical	0.14	0.15	0.02
	(0.35)	(0.36)	(0.01)
ferrous and non-ferrous metal	0.1	0.06	-0.03
	(0.3)	(0.24)	(0.01)
machinery	0.25	0.29	0.04
	(0.43)	(0.46)	(0.01)
electric and electronic	0.09	0.16	0.08
	(0.28)	(0.37)	(0.01)
# of plants	4927	2049	-

Notes: Data on manufacturing plants are from the list of LMS manufacturing plants in 1985. Time is divided into two periods: "before TF" indicates years before 1964, "during TF" indicates years between 1964 and 1978. In Columns 1 and 2, statistics are averages across plants. Column 3 reports the mean differences between Column 1 and Column 2. Standard deviations (Columns 1 and 2) or standard errors (Column 3) are in parentheses. County characteristics are from various population censuses, industrial surveys, and GIS maps.

Table 2: Summary Statistics of the Regression Sample

	mean	s.d.	p10	p50	p90
treatment					
TF-built manufacturing capacity	1.11	1.51	0.00	0.54	2.88
site-selection criteria					
log distance to complete railway	3.16	1.39	1.35	3.09	5.00
average slope	4.47	2.39	1.76	4.37	7.25
log mean elevation	7.05	0.58	6.28	7.09	7.73
log distance to provincial capital	5.35	0.45	4.66	5.38	5.89
pre-TF conditions					
urban rate, 1964	7.14	2.70	4.07	6.53	11.68
log population density, 1964	4.42	0.94	3.32	4.50	5.56
industrial employment share, 1936	0.34	1.85	0.00	0.00	0.00
log distance to existing railway	3.68	1.28	2.33	3.76	5.30
2004 manufacturing employment					
manufacturing employment share	4.29	4.05	1.26	2.99	9.43
private	2.69	3.12	0.65	1.69	4.93
private and opened after 1985	1.57	2.02	0.37	0.95	3.09
migration flows in 2005					
in-migration rate	0.96	1.01	0.13	0.76	1.94
workers with rural hukou					
manufacturing employment share	1.96	1.12	0.81	1.77	3.16
non-agricultural employment share	7.91	3.53	4.49	7.37	12.49

Notes: There are 73 prefectures in the baseline sample. Online Appendix Table A.1 shows detailed definitions of the variables. Online Appendix A reports more details on data sources.

Table 3: Location Characteristics by Plants' Time of Opening: within the TF Region

time of opening	before TF	during TF	(2)-(1)
	(1)	(2)	(3)
Panel A: County characteristics			
in provincial capital prefecture	0.39	0.17	-0.22
	(0.49)	(0.37)	(0.02)
average slope	2.58	3.1	0.52
	(2.06)	(2.04)	(0.11)
log mean elevation	6.56	6.7	0.13
	(0.74)	(0.71)	(0.04)
log distance to provincial capital	3.16	4.22	1.06
	(2.56)	(1.99)	(0.13)
urban rate, 1964	36.82	22.28	-14.54
	(21.78)	(21.13)	(1.2)
log population density, 1964	6.05	5.37	-0.68
	(1.12)	(1.01)	(0.06)
industrial employment share, 1936	13.00	3.81	-9.19
	(23.66)	(13.92)	(1.09)
Panel B: Sector			
mining	0.08	0.04	-0.04
	(0.26)	(0.18)	(0.01)
light industries	0.28	0.19	-0.09
	(0.45)	(0.39)	(0.02)
power/water	0.09	0.05	-0.03
	(0.28)	(0.22)	(0.01)
chemical	0.11	0.1	-0.01
	(0.31)	(0.30)	(0.02)
ferrous and non-ferrous metals	0.10	0.07	-0.03
	(0.30)	(0.26)	(0.02)
machinery	0.25	0.35	0.10
	(0.44)	(0.48)	(0.03)
electric and electronic	0.10	0.20	0.10
	(0.30)	(0.40)	(0.02)
# of plants	634	657	-

Notes: The sample includes LMS plants *inside* the TF region. Time is divided into two periods: "before TF" indicates years before 1964, "during TF" indicates years between 1964 and 1978. In Columns 1 and 2, statistics are averages across plants. Column 3 reports the mean differences between Column 1 and Column 2. Standard deviations (Columns 1 and 2) or standard errors (Column 3) are in parentheses. In Panel A, county characteristics are from various population censuses, industrial surveys, and GIS maps. In Panel B, individual LMS plants are grouped into 7 sectors, share among all plants is reported.

Table 4: Determinants of Local Manufacturing Capacity

dep var: share of employment from LMS manufacturing plants							
Panel A: Third Front Region	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log dist to complete railway	-0.543	0.193	0.226			0.243	0.311
	(0.136)	(0.224)	(0.227)			(0.237)	(0.235)
log dist to existing railway		-0.960	-1.042			-1.006	-1.049
		(0.317)	(0.338)			(0.369)	(0.374)
average slope			0.195			0.200	0.191
			(0.091)			(0.094)	(0.104)
average elevation			-0.348			-0.502	-0.161
			(0.259)			(0.245)	(0.408)
log dist to prov capital			0.094			0.119	0.382
			(0.529)			(0.536)	(0.507)
urban rate in 1964				0.165	0.209	0.092	0.110
				(0.065)	(0.076)	(0.061)	(0.072)
log population density in 1964					0.433		0.414
					(0.362)		(0.391)
industrial emp share in 1936					-0.043		-0.026
					(0.053)		(0.040)
R^2	0.317	0.531	0.572	0.215	0.260	0.592	0.604
joint test for initial conditions ($p-$val)					0.041		0.505
Panel B: Non-Third Front Region	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log dist to complete railway	-0.637	-0.457	-0.444			-0.268	-0.319
	(0.169)	(0.330)	(0.327)			(0.249)	(0.252)
log dist to existing railway		-0.228	-0.138			0.025	0.097
		(0.326)	(0.363)			(0.303)	(0.310)
average slope			0.089			0.098	0.072
			(0.134)			(0.135)	(0.133)
average elevation			0.436			0.625	0.915
			(0.661)			(0.592)	(0.598)
log dist to the nearest prov capital			-0.876			-0.921	-0.616
			(0.436)			(0.377)	(0.370)
urban rate in 1964				0.228	0.219	0.208	0.208
				(0.043)	(0.043)	(0.048)	(0.049)
log population density in 1964					0.622		0.600
					(0.191)		(0.232)
industrial emp share in 1936					0.012		0.010
					(0.004)		(0.005)
R^2	0.247	0.252	0.268	0.412	0.451	0.452	0.481
joint test for pre-TF conditions ($p-$val)					0.000		0.000

Notes: The dependent variable is the share of 1985 employment from the LMS manufacturing plants. All models include a set of province fixed effects. Panel A includes 73 sample prefectures. Panel B includes 211 prefectures outside the TF region that are neither provincial capitals, nor in one of the ethnicity-minority autonomous areas: Tibet, Inner Mongolia, and Xinjiang. Robust standard errors are in parentheses. $p-$value is from a joint test for pre-TF conditions: urban rate in 1964, log population density in 1964, and industrial employment share in 1936.

Table 5: Route and Timing of Railway Construction and Pre-TF Economic Conditions

	(1)	(2)	(3)	(4)	(5)	(6)
	timing of railway construction			route of railway		
	urban rate ’64	ln pop density ’64	% ind emp ’36	urban rate ’64	ln pop density ’64	% ind emp ’36
ln dist to complete railway	-0.237	-0.120	0.189	-0.434	-0.080	0.072
	(0.361)	(0.062)	(0.174)	(0.306)	(0.051)	(0.099)
ln dist to existing railway	-0.290	0.059	-0.172			
	(0.527)	(0.085)	(0.208)			
geo chars	X	X	X	X	X	X
province FE	X	X	X	X	X	X
N	73	73	73	73	73	73
mean dep var	7.136	4.418	0.336	7.136	4.418	0.336

Notes: Geographic characteristics include log average elevation, log distance to the provincial capital, and employment share in the mining sector. Robust standard errors are in parentheses.

Table 6: The TF and Manufacturing Employment in 2004

	dep var: share of manu employment in 2004					
	(1)	(2)	(3)	(4)	(5)	(6)
TF-built manu capacity	2.769	2.679	2.730	1.692	1.903	2.218
	(0.233)	(0.370)	(0.394)	(0.401)	(0.396)	(0.311)
geo conditions		X	X	X	X	X
pre-TF economic conditions			X	X	X	X
log dist to nearest node				X	X	
province FE	X	X	X	X	X	X
model	2SLS	2SLS	2SLS	2SLS	2SLS	OLS
railway networks	actual	actual	actual	min. dist	min. cost	actual
first stage						
log dist. to existing railway	-0.960	-0.907	-0.848	-0.866	-0.941	-
	(0.317)	(0.254)	(0.226)	(0.262)	(0.265)	-
log dist. to complete railway	0.193	0.266	0.421	0.635	0.855	-
	(0.224)	(0.185)	(0.199)	(0.260)	(0.310)	-
first-stage F-statistic	9.167	12.754	14.056	10.884	12.622	-

Notes: Each observation is a prefecture and the sample size is 73. The mean dependent variable is 4.287. Columns (1) - (5) use 2SLS specifications, with the instrument being the log distance to the existing railway. Column (6) uses OLS. Log distance to the complete railway is controlled for in all columns. In Columns (1) - (3) and Column (6), distances are calculated based on the actual railway networks. In Column (4), distances are based on the counterfactual railway networks constructed to minimize railway length while still connecting the same pair of terminal cities (nodes). In Column (5), distances are based on the counterfactual railway networks constructed to minimize the minimum construction cost between nodes. Log distances to the nearest node are also controlled for in Columns (4) and (5).

Table 7: Robustness in Measurement of TF-built Manufacturing Capacity

	dep var: manu emp share in 2004				
	(1)	(2)	(3)	(4)	(5)
emp share from LMS plants, '85	2.435				
	(0.333)				
emp share from smaller plants, '85	1.313				
	(0.299)				
ind. output per 100 workers, '78		2.111			
		(0.405)			
Δ in ind. output, '64-'78, per 100 '64 pop			3.187		
			(0.665)		
# of LMS plants per 10k workers, '85				83.850	
				(19.678)	
= 1 if designated TF industrial center					20.959
					(9.836)
N	73	67	67	73	73
first stage					
log dist to existing railway	-0.843	-1.158	-0.767	-0.028	-0.111
	(0.231)	(0.407)	(0.251)	(0.012)	(0.057)
log dist to complete railway	0.423	0.613	0.366	0.016	-0.017
	(0.199)	(0.393)	(0.242)	(0.010)	(0.048)
first-stage F-statistic	13.322	8.101	9.361	5.720	3.703

Notes: Column 1 accounts for employment from smaller plants. Column 2 measures the TF-built manufacturing capacity by the prefecture's industrial output in 1978 (in *yuan*) per 100 workers. Column 3 measures TF investment by changes in the prefecture's industrial output between 1964 and 1978, per 100 residents in 1964. Column 4 measures manufacturing capacity by the number of LMS plants per 10,000 workers in 1985. The main explanatory variable in Column 5 is a binary variable indicating whether the prefecture is designated to be a key industrial center according to the Third Five-year Plan (1966-1970). All models are estimated using the same 2SLS specification, with the same set of controls, as in Column 3 of Table 6. Robust standard errors are in parentheses.

Table 8: The TF and the Local Private Manufacturing Sector

	emp share					# firms	efficiency	
	all	estab. '85-'98	emp ≤ 25	all	all	per 100 emp	ln TFP	ln manu wage
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
TF-built manu capacity	2.124	0.727	0.018	2.588	2.749	0.016	0.096	0.034
	(0.348)	(0.145)	(0.003)	(0.397)	(0.739)	(0.003)	(0.055)	(0.019)
Δ in state manu emp share, '85-'04				1.178				
				(0.289)				
manu capacity$\times$ state-share index					-6.178			
					(5.684)			
state-share index					4.947			
					(3.408)			
geo and pre-TF conditions	X	X	X	X	X	X	X	X
province FE	X	X	X	X	X	X	X	X
mean dep var	2.690	0.875	0.029	2.690	2.690	0.038	0.002	6.383
first-stage F-statistic	14.056	14.056	14.056	17.025	1.307	14.056	18.931	18.787
N	73	73	73	73	73	73	4738	9526

Notes: The table focuses on the local private manufacturing sector in 2004, with variables constructed from the *2004 Economic Census*. All models are estimated using the same 2SLS specification, with the same set of controls, as in Column 3 of Table 6. Robust standard errors are in parentheses.

Table 9: Inter-Industry Linkages

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		inter-industry linkage measures					
	within 2-digit industry	worker transitions	citation pattern	tech input	tech output	manu input	manu output
Panel A	dep var: share of 2004 employment from private manufacturing firms						
same-ind manu capacity	0.297						
	(0.128)						
linkages		0.005	0.022	0.020	0.046	0.017	0.020
		(0.002)	(0.015)	(0.008)	(0.018)	(0.008)	(0.015)
industry FE	X	X	X	X	X	X	X
prefecture FE	X	X	X	X	X	X	X
N	1387	1387	1387	1387	1387	1387	1387
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel B	dep var: private firm TFP						
same-ind manu capacity	0.038						
	(0.063)						
linkages		0.003	0.028	0.028	0.024	0.021	0.031
		(0.002)	(0.046)	(0.023)	(0.025)	(0.025)	(0.027)
industry FE	X	X	X	X	X	X	X
prefecture FE	X	X	X	X	X	X	X
N	4656	4656	4656	4656	4656	4656	4656

Notes: This table focuses on the private sector manufacturing firms in 2004. In Columns 2 to 7, each linkage is indicated in the header. Each observation in Panel A is a prefecture-2 digit industry pair. Each observation in Panel B is a firm. All models include 2-digit industry fixed effects and prefecture fixed effects. All models are estimated using the OLS. Standard errors are clustered at the prefecture level.

Table 10: Post-Reform Inter-Prefecture Migration and Local Structural Transformation

	% of local workers with *hukou* from a diff. pref.	% of manu workers among workers with local rural *hukou*
	(1)	(2)
TF-built manu capacity	-0.082	0.225
	(0.090)	(0.118)
geo and pre-TF conditions	X	X
province FE	X	X
N	70	70
first stage F-stat	14.184	14.184
mean dep var	1.114	1.964

Notes: Data are from the individual-level sample of the intra-decennial population census of 2005. All columns are estimated using the 2SLS estimation. The dependent variable in Column 1 is the percent of workers in a prefecture with *hukou* from another prefecture. The dependent variable in Column 2 is the percent of local manufacturing workers among those with a rural *hukou* registered in the same prefecture. Robust standard errors are in parentheses.

Table 11: Additional Indicators for Regional Economic Development

	(1) % urban 2010	(2) % college 2010	(3) ln p.c. GDP 2010
TF-built manu capacity	3.142	0.850	0.111
	(1.132)	(0.364)	(0.067)
geo and pre-TF conditions	X	X	X
province FE	X	X	X
N	73	73	73
mean dep var	31.409	2.486	9.618
first-stage F-stat	14.056	14.056	14.056

Note: All models are estimated using the 2SLS, where the distance to the existing railway network is used to instrument for TF-built manufacturing capacity. The specification is the same as Table 6 Column 3. The dependent variable is the percent of urban population in 2010 (Column 1), the percent of college graduates among people between 16 and 64 years old in 2010 (Column 2), and log GDP per capita in 2010. Sources for dependent variables are the county-level tabulations of the *2010 census* (Columns 1 and 2) and provincial statistical yearbooks (Column 3). Robust standard errors are in parentheses.

Table 12: Parameters for TF Region and non-TF Region

parameter	**variable**	**TF Region**	**non-TF Region**
W^M	average manufacturing wage	687	900
W^A	average agricultural wage	234	337
$W^M - \frac{W^A}{\alpha}$	manu-ag income gap (absolute value)	395	479
l^A	agriculture employment share	0.80	0.47
TFP^s	average log TFP of state firms	-0.25	-0.04
TFP^{ns}	average log TFP of private firms	-0.10	0.04

Notes: Wage data are from the *2005 intra-decennial census*. Wages reported here are monthly income. TFP is calculated from the *2004 Annual Manufacturing Firm Survey*. The value of α is 0.8.