I. Introduction
Transportation infrastructure, including railways, roads, airlines, and waterways, is widely considered a key contributor to economic development. This has motivated many countries to devote large amounts of public investment to transportation. For example, in 2011, China invested more than RMB 747 billion on transportation (6.9% of government expenditure), a 36% increase compared to 2010 and eight times the amount invested in 2005 (Ministry of Railways 2011). Similarly, the United States spent USD 136 billion (2.5% of government expenditure) on transportation in 2010. The emphasis on public investment in transportation is even more prominent during recessions. For instance, 38% of China’s RMB 4 trillion stimulus package is devoted to transportation (Global Times 2010). Improving transportation infrastructure is also considered essential for reducing poverty (Ali and Pernia 2003; Gibson and Rozelle 2003). Around 20% of World Bank loans, which is larger than the total share allocated to education, health, and social services, were allocated to transportation projects in 2007 (World Bank 2007).

The importance of transportation appears to be justified by the observation that the more developed economies usually have better transportation systems, and large transportation projects often occur during periods of rapid economic growth. However, correlation does not mean causality. Whether transportation investment leads to or simply follows economic development is still unclear. Given the large amount invested in transportation, estimating the causal effect of transportation on economic development is crucial.

Our article exploits a natural experiment created by a newly constructed railway in China, the Qingzang railway, to study the effect of railway infrastructure on the local economy. The findings of this study are relevant to trans-
Transportation investment in developing economies, as the railway is located in two of the least developed provinces (Qinghai and Tibet) in China. In addition, due to the underdevelopment of the local economy, Qingzang railway was mainly initiated by the central government. As a result, the project does not appear to have been driven by local demand or local policies, which provides a better context for addressing the reverse-causality problem than that used in many of the existing studies.

We apply the difference-in-difference (DD) method to estimate the railway effect on the local economy. Specifically, we distinguish railway counties (counties along the railway) from off-railway counties and compare the economic development of these two groups before and after the arrival of the railway. This method eliminates time-invariant differences between the two groups. We also allow the time trends of a county’s economic growth to vary with its initial economic status such as initial level of gross domestic product (GDP) per capita and population.

The baseline results show that the Qingzang railway stimulated a 33% or so increase in annual GDP per capita in the railway counties. The estimate provides an aggregate effect of the railway on GDP. It includes the direct gains resulting from the decrease in transportation costs of output, intermediate goods, labor, and technology, as illustrated in our theoretical model. It also includes the indirect gains through the railway effect on urbanization, market integration, economies of scale, economies of agglomeration, and so on. The railway effect on total GDP was similar to the railway effect on GDP per capita because the railway effect on the number of permanent residents was nonsignificant. This result is not surprising given the presence of the strict Hukou system in China.

The estimate is robust to different ways of constructing the treatment and control groups, including defining a county’s treatment status on the basis of the distance between the county seat to the nearest railway station and excluding provincial or prefectural capitals. The results from the counterfactual test, the matching DD method, and the weighted regression method all increase our confidence in the baseline estimate. However, we do not have a natural experiment that randomly places the railway, and we cannot observe all characteristics that affect both economic growth and railway placement. As a result, we cannot completely rule out the possibility of selection bias.

We also find that the railway effect on GDP worked mainly through the positive impact on manufacturing, whereas the agriculture and service industries were largely unaffected. This result is consistent with the prediction of our theoretical model, given that the manufacturing industry is more likely
to compete in the national market and benefits from the reduction in the transportation cost of inputs and technology. In addition, the result shows that the railway’s effect did not vary significantly with a county’s initial economic status. The construction before the operation of the railway had a positive impact on the local GDP, but this effect was much smaller than that of the railway after it opened.

The railway could have spillover effects on the off-railway counties that are within the same region as the railway counties, as the economic impact of a new railway is prone to leak outside of the immediate economic area (Rephann and Isserman 1994). However, we do not find evidence for such a spillover effect for GDP per capita, population, agriculture, and the manufacturing industry. Nevertheless, there was a significant decline in the value added of service industries in the off-railway counties adjacent to the railway counties, meaning that there could be a displacement effect on the service sector of these untreated adjacent counties.

The remainder of this article is organized as follows: Section II introduces the background and reviews the literature. Section III presents a model to illustrate the direct effects of the railway. Section IV then describes the data and the empirical model. Section V presents the results of the baseline models, as well as the spillover effects. Section VI then tests the robustness of the baseline results. Section VII analyzes the potential heterogeneity and dynamics of the railway effect. Finally, Section VIII concludes the article.

II. Background and Literature

A. Qingzang Railway

The Qingzang railway, or the Qinghai-Tibet railway, is one of China’s four major projects in the twenty-first century, as stipulated by the Tenth Five-Year Plan of China (Zhu 2001). This high-altitude railway crosses the Qinghai-Tibet Plateau, known as “the Roof of the World,” running from Xining, the capital of Qinghai province, to Lhasa, the capital of the Tibet Autonomous Region. It is the first railway to connect Tibet with other provinces in China. This railway enables trains to run from metropolitan areas like Beijing, Shanghai, Guangzhou, Chongqing, and Lanzhou to Tibet (fig. 1). The railway was built in two parts. The first section connected Xining to Golmud (the second largest city in Qinghai province) and was completed in 1984. Construction of

1 The other main projects proposed in the Tenth Five-Year plan include the transmission of natural gas and electricity from western to eastern regions, the Beijing-Shanghai high-speed railway, and a project to divert river water from south to north.
the second section, connecting Golmud and Lhasa, started in July 2001 and began operating in July 2006. The total length of Qingzang railway is 1,956 kilometers, and the length of the second section alone is 1,142 kilometers.

The proposal to build a railway on the Qinghai-Tibet plateau can be traced back to 1919 (Sun 1928). The project was not started until 1958. The first section opened to traffic in 1984, but the second section was delayed due to the technical difficulties arising from the construction of a railway in a permafrost region. In 2001, the second section was demonstrated to be technically feasible, prompting the government to start the construction of this section. The second section cost RMB 33 billion to construct, of which 75% came from the Ministry of Finance, and the remaining 25% from the Railway Construction Fund. Local governments did not directly contribute to the construction of the railway.

It is within reason to claim that the construction of the second section provides us a natural experiment to examine how the introduction of a railway can stimulate the local economy of an underdeveloped region. There are two reasons for this claim. First, Qinghai and Tibet are much less developed than other provinces in China, meaning that the Qingzang railway did not result from local economic development. To be exact, the GDP per capita of Tibet and Qinghai in 2000 was at only 58% and 65% of the national average in China.
Second, although local residents and local governments welcomed the railway, they neither proposed nor drove the project. Rather, the project was mainly promoted and financed by the central government, with the aim of stimulating the local economy and strengthening the ties between Tibet and the rest of China. In addition, the choice of the railway route was mainly determined by technical and cost concerns, rather than the economic development of the counties along the railway. The Qinghai-Tibet route was adopted over three alternative proposals because of its lower cost, shorter length, and superior geographic condition (*Economic Daily* 2001).

Between 2006 and 2011, the Qinghai-Tibet railway was estimated to transport over 41 million passengers and 180 million tons of cargo, constituting about 75% of the cargo in this region. The railway boasted that it had a one-way capacity nearly 40 times that of the existing transportation infrastructure and that it reduced transportation costs by more than half (*People’s Daily* 2006, 2011).

The contribution of the Qingzang railway to economic growth appears to be evident in the aggregate provincial data, particularly for Qinghai. Before its construction, Qinghai and Tibet showed modest growth in GDP per capita, at 7.2% and 9.9% respectively (1997–2000). After the Qingzang railway was put into operation, the average growth rate of GDP per capita rose to 11.1% in Qinghai and 11% in Tibet (2007–9).²

However, this concurrence between economic growth and the construction of Qingzang railway does not immediately imply causality. There could be contemporary policies or changes that also affected the local economy. We discuss in further detail this possible contamination of our estimates of the railway’s effect in Section IV. However, a review of the policies in the Tenth and Eleventh Five-Year Plans of Qinghai and Tibet (2000–2005, 2006–10) shows that the counties affected by other contemporary policies were unlikely to largely overlap with those traversed by the railway. Specifically, most of the contemporary provincial-level policies—the policies on tax, health care, and education—had been implemented in these two provinces before 2006.³ Policies that targeted specific counties, such as the establishment of the Lhasa Economic and Technological Development Zone in 2003 or the poverty reduction programs

² The national average was 7.6% and 10.4% in these two periods, respectively.
³ For example, one of the most important policies in rural areas is the introduction of the New Cooperative Medical Scheme. The program was piloted in 2003 and launched in all counties in Qinghai and 83% of counties in Tibet in 2005. Also, both provinces abolished agricultural taxes in 2005. The exemption of tuition and fees for compulsory education in rural areas was implemented nationally in 2006.
for some counties, were unlikely to impose significant effects on all the railway counties. Finally, the locations of other infrastructure projects, such as the construction of roads, dams, airports, and electricity networks, did not coincide with the placement of railways.

B. Literature

Beginning with the influential work of Aschauer (1989) and Barro (1990), several empirical studies have tried to examine the stimulation effect of infrastructure on the local economy. The earlier studies focused on the effect of infrastructure on productivity and estimated the output elasticity of infrastructure through the macroeconomic method of production or cost function. Most of these studies argue that increases in infrastructure raise output and productivity (e.g., Keeler and Ying 1988; Lynde and Richmond 1992; Munnell 1992; Nadiri and Mamuneas 1994; Fernald 1999).

However, the estimates vary substantially, depending on the form of the production or cost function, data, and econometric methods (Munnell 1992). Particularly, several papers find insignificant effects of infrastructure investments on aggregate economic performance (e.g., Evans and Karras 1994; Holtz-Eakin and Schwartz 1995; Garcia-Milá, McGuire, and Porter 1996; Holtz-Eakin and Lovely 1996). In addition, as highlighted by Gramlich (1994) and Haughwout (2002), the production or cost function method shows correlation but not causality. Studies based on panel data controlling for fixed effects do not solve the problem of reverse causality or simultaneity of infrastructure provision and economic growth either.

More recent studies examine the issue by relying on plausible exogenous changes in the transportation system. These studies exploit variations across subnational units such as cities or states instead of cross-country variations (Haughwout 2002). In addition, because aggregate-output data mask interindustry differences, the focus of these studies shifts away from aggregate productivity or output to the composition of economic activity and the population distribution (Chandra and Thompson 2000; Duranton and Turner 2012).

These studies find that transportation systems can have different impacts across industries. Not surprisingly, market-oriented industries and traffic-related industries are more likely to be stimulated by these systems. For example, Chandra and Thompson (2000) find that the introduction of a new highway raised economic growth, as measured by total earnings, in nonmetropolitan counties that the highway directly passed through. This highway effect was mostly concentrated on the manufacturing industry, service industry, retail trade, transportation, and public utilities. Rephann and Isserman (1994) find similar highway effects only for the highway counties that are in close proximity to large cities or
have some degree of prior urbanization. In contrast, Duranton and Turner (2012) find that the changes in the highway system between 1983 and 2003 did not have a significant effect on the composition of industrial activity in a city. For the railway, Haines and Margo (2006) find that the introduction of a railway in the 1850s raised the likelihood of participation in the service sector, increased urbanization, and decreased agricultural yield.

The literature also highlights transportation as a force of changing the spatial allocation of economic activity. For example, Chandra and Thompson (2000) find that highways increase the level of economic activity only in the highway counties, drawing activity away from adjacent counties, and have no significant effect on total regional growth. Rephann and Isserman (1994) find that highways do not significantly affect off-highway counties.

Transportation can also change the spatial allocation of population or labor, as well as accelerate urbanization (Burchfield et al. 2006; Haines and Margo 2006; Baum-Snow 2007; Michaels 2008; Atack et al. 2010; Duranton and Turner 2012). For example, Atack et al. (2010) find that railroads have a strong effect on urbanization but only a small effect on total population growth. Duranton and Turner (2012) find that a 10% increase in a city’s initial stock of highways caused about a 1.1% increase in its employment between 1983 and 2003, while Baum-Snow (2007) shows that one new highway passing through a central city reduced its population by about 18% between 1950 and 1990. Aside from this, literature also shows that transportation can increase market size, reduce the price gap, and increase trade-related activities (Keller and Shiue 2008; Michaels 2008; Donaldson 2010).

We note that most of the above-mentioned papers discuss the historical expansion of the US interstate road system. Few studies have examined the causal relationship between China’s investment in transportation and its remarkable economic growth. Previous studies on China mostly use aggregate time series data, or provincial panel data to estimate the output elasticity of infrastructure (Fleisher and Chen 1997; Mody and Wang 1997; Démurger 2001). However, these studies do not adequately address the endogenous problem of infrastructure. There are three recent papers that make genuine progress on this issue. Specifically, Banerjee, Duflo, and Qian (2012) solve the reverse causality problem by exploiting an exogenous source of variation in access to transportation networks, that is, the distances from the counties to the straight line that connects historical cities or treaty ports. They find that the proximity to transportation networks has a moderate positive effect on a county’s long-term

---

4 The exceptions include Keller and Shiue (2008), who consider the introduction of steam trains in Germany, and Donaldson (2010), who considers railroad construction in nineteenth-century India.
per capita GDP. By applying instrumental variable (IV) estimation based on the construction of least cost paths, Faber (2012) finds that China’s National Trunk Highway System had led to a reduction in GDP growth among nontargeted peripheral counties due to the increase in localized concentrations of industrial production. Finally, Baum-Snow et al. (2012) construct the IV based on historical transportation infrastructure and find that in the last 20 years, the expansion of Chinese railroad networks displaced industrial GDP, while road networks displaced the population from central cities to surrounding regions.

As a whole, the studies on China’s transportation infrastructure offer mixed conclusions regarding the effect of transportation infrastructure. Moreover, these papers do not look into the importance of the railway for the more underdeveloped areas, particularly where the transportation network is far behind that of the other regions in China and the local economy has been isolated from the main economic actors.

III. Theoretical Framework

In this section, we develop a spatial competition model to illustrate how the railway can directly affect the local economic growth by reducing transportation costs. We study not only the railway counties but also the off-railway counties located in the same region. This model borrows from Chandra and Thompson (2000), and we extend their model to incorporate the production process and consider the influence of the railway on the choices of input.

Suppose firms have a standard production function with three inputs: \( Y = F(L, K, I) \), where \( Y \) is the output, \( F(\cdot) \) is the production function, and \( L, K, \) and \( I \) represent labor, capital, and intermediate goods, respectively. The production function is “constant returns to scale,” which implies that the marginal cost (MC) of production is constant and equal to the unit cost. Assume there is no entry barrier, firms are price takers, and \( P^S \) denotes the price of the output. The constant-returns-to-scale production function alongside free entry implies that firms earn zero profit. Therefore, in an equilibrium, we must have \( P^S = MC \).

Our model focuses on three channels of the railway effect: reducing the cost of transporting the output, reducing the cost of transporting input, and reducing the cost of technology diffusion or adoption. The qualitative theoretical predictions are derived from the first channel. The other two channels quantitatively magnify these predictions.

A. Channel 1: Reducing the Cost of Transporting the Output to Consumers

To simplify the analysis, we assume for the moment that the railway does not affect the MC and \( P^S \) of the firms, as well as the input prices. The delivery cost
of output creates a price wedge between the price for a firm \((P^S)\) and the price for consumers \((P^D)\). Specifically, \(P^S + t \times D = P^D\), where \(t\) represents the travel cost and \(D\) represents the distance between the firm and consumers. Reasonably, \(t\) decreases with the transportation infrastructure \((G)\) and the effect of \(G\) on the economy equates to the negative of the effect of \(t\) on the economy.

Assume firms, or in this case the counties, produce the same good and compete for consumers. Two routes, route 1 and route 2, connect firms and consumers, as indicated in figure 2. Route 1 connects to route 2 at point \(O\). Their unit travel cost is \(t_1\) and \(t_2\), respectively. Route 1 is the route that introduces the railway, and the new railway reduces \(t_i\). We focus on how the railway affects the local firm \(A\) that the railway passes through.

We first consider the case in which the local firm \(A\) competes with another local firm \(B\) that is also located at route 1. As described in Chandra and Thompson (2000), this case can be labeled as a case of competition within the market of the railway region. For this case, it is easy to show that (see the proof in app. A) the railway increases firm \(A\)'s output \(\left(\frac{d D_A}{dt_1} \leq 0\right)\) if and only if firm \(A\) has a lower MC (and thus \(P^S\)) than firm \(B\).

The local firm \(A\) can also compete with an off-railway firm \(C\) that is located at route 2. Firm \(C\) may be located in a county adjacent to the railway county. It could also be located in a county farther away from the railway, such as a county outside of the provinces served by the railway. The marginal consumer who is indifferent to purchasing from \(A\) or \(C\) can be located at route 1 or route 2. The second case considers the scenario in which the marginal consumer is located on route 1, say, at point \(j\). Since the marginal consumer is within the railway region, the competition in this case remains primarily in the railway regional market. Appendix A proves that the railway increases local firm \(A\)'s

![Figure 2. Effect of reducing transportation cost of route 1](http://www.journals.uchicago.edu/t-and-c).
output if and only if $P_A^s \leq P_C^s + t_2 \times D_{CO}$. This means that firm C contracts (i.e., loses consumers) unless its MC is low enough to offset its higher cost of traveling to the railway.

In the third case, firm A still competes with an off-railway firm C, with the difference that now the marginal consumer is located on route 2, say, at point $k$. This case corresponds to the case of nationally traded goods as looked into by Chandra and Thompson (2000), where firm A trades with consumers outside of the railway region, or in the national market. Appendix A shows that in this case, firm A always expands after the arrival of the railway, while firm C always contracts. This is intuitive because firm C does not use the railway to reach the marginal consumer and does not enjoy the drop in travel cost.

In order to take into account a firm’s cost of traveling to the railway, we define the effective MC as the MC of production plus the firm’s travel cost to the railway. Then, the following conclusion holds for both along-railway firms and off-railway firms: a firm expands after the arrival of the railway as long as it has a lower effective MC than its competitors. Off-railway firms are left at a disadvantage by the railway because of the higher cost of traveling to the railway.

The three cases also imply three testable predictions. First, the railway effect on the output of a local firm (or county) is ambiguous if this firm competes for marginal consumers within the railway regional market. Firms with a lower effective MC expand, while firms with a higher effective MC contract. As a result, the aggregate railway effect on local firms is ambiguous. Second, the railway definitely stimulates the output of a local firm if the firm competes for marginal consumers in the national market (i.e., out of the railway region). Finally, the off-railway counties in the same region as the railway counties are likely to contract in the regional market competition unless they have a much lower production MC to offset the higher transportation costs in comparison to the railway counties. However, they may gain consumers in the national market.

It is difficult to ascertain whether firms compete within the regional market. However, it is reasonable to argue that the competition of manufacturing products is likely to reach beyond the railway region; that is, manufacturing goods are likely to be nationally traded. In contrast, agricultural products and services are more likely to be traded within the regional market. Therefore, the railway effect on the local manufacturing industry is likely to be positive, while the railway effect on the agricultural and service industries remains ambiguous. Some subindustries of the service industry, such as the tourist industry, may be traded nationally. Hence, the railway effect on the service industry can be slightly
stronger than that on agriculture, depending on how important the subindustry is in the county.

The difference across industries also matters for off-railway counties that are adjacent to the railway counties. Particularly, the displacement spillover effect is more likely to occur in the agricultural and service industries than in the manufacturing industry.

B. **Channel 2: Reducing the Cost of Transporting the Input**

The railway reduces the transportation cost of intermediate goods and, possibly, labor. We now illustrate the effect with intermediate goods. Similar to the output, the travel cost creates a price gap between the price of input ($\theta^i$) for the local firm we are interested in and the supplier’s price ($\theta^s$). Specifically, $\theta^i + t_I \times D = \theta$, where $t_I$ is the unit travel cost and $D$ is the distance between the local firm and suppliers. For simplicity, we assume that the railway does not affect the supplier’s price. Then for the local firm, the railway simply reduces $t_I$ and thus $\theta$. Therefore, the MC of producing output declines for this firm. The discussion regarding the first channel implies that this change helps the firm attract more consumers of the output.

The changes in input costs magnify the difference of the railway effect on output between industries. Specifically, the fall in MC related to intermediate goods is important for the manufacturing industry but does not seem to be essential for the agriculture and service industries. The railway effect on labor can be substantial for both manufacturing and service industries but not for agriculture because labor surplus remains common in rural areas.

C. **Channel 3: Reducing the Cost of Technology Diffusion or Adoption**

The adoption or diffusion of technology can be costly, particularly when the adoption is accompanied with transporting equipment, human capital, and knowledge. Therefore, by reducing the cost of technology adoption, the improvement in transportation can reduce the MC of producing output and increase the competitiveness of local firms in the market. This channel is also likely to magnify the difference in the railway effect between the manufacturing industry and other industries, as the competitiveness of the manufacturing industry depends heavily on technology, more so than the other two industries.

Overall, the model predicts that the railway stimulates the GDP of railway counties through its positive effect on the manufacturing industry. The railway effect on agriculture and service industries tends to be significantly smaller than that on the manufacturing industry. Finally, the railway effect on
off-railway counties that are located in the same regional market as the railway counties is likely to be much weaker than that on the railway counties. In addition, for these off-railway counties, the negative spillover effect is more likely to occur in agriculture and service industries. Nevertheless, the existence of a labor surplus and the immobility of the crucial input (land) can baffle the displacement of agriculture in China.

IV. Data and Empirical Specification

A. Data

Due to a lack of village-scale-level data and the existence of the spillover effect within a county, we use county-level data to analyze the railway effect. Notice that counties along the railway tend to experience a stronger influence than those farther away from the railway. Hence, we distinguish two types of counties in the empirical study: railway counties, which we regard as the treatment group, and off-railway counties, which we consider as the control group. There are 20 railway counties, five in Tibet and 15 in the Qinghai province. Of the 20 counties, three have no passenger stations (of which, two have stations for cargos), and we will exclude them in the robustness test. There are 96 off-railway counties, 68 in Tibet and 28 in Qinghai. We also use geographical information systems (GIS) data to calculate the Euclid distance from each county seat (usually the hub of the county) to the nearest station. We use this distance to measure the actual traveling distance and conduct the robustness analysis.

The construction of the second section of the Qingzang railway started in mid-2001, was completed at the end of 2005, and was opened to traffic in mid-2006. We thus distinguish three periods based on the potential influence of the railway: pre-2000 (pretreatment period), 2002–5 (construction period), and post-2007 (posttreatment period). We exclude 2001 and 2006 because the status of the railway changed in the middle of these years.

We collect information on counties in Tibet and Qinghai between 1997 and 2009. All data come from the statistical yearbooks of Tibet and Qinghai, as well as other related sources.5 Due to a lack of information on GDP and other important variables in Tibet in 1997, 1998, and 2005, we focus on the periods 1999–2000, 2002–4, and 2007–9. The descriptive statistics are shown in table B1. In panel A, we summarize the statistics of all counties in Tibet and Qinghai. We notice that on average the railway counties have a higher GDP

5 More specifically, the data sources include the “Tibet Statistical Yearbook,” “Tibet Yearbook,” “Economic and Social Statistical Book of Tibet,” “Qinghai Statistical Yearbook,” “Qinghai Yearbook,” and “China Statistical Yearbook for Regional Economy.”
per capita than other counties. The main reason for this is that the railway tends
to pass through the provincial or prefectural capitals in these two provinces.
These cities are more developed and more populated. After excluding these
cities, the GDP per capita gets much more similar between the railway counties
and other counties.

The information on industries is only available for counties in the Qinghai
province. Panel B of table B1 gives the statistics and reveals that, compared with
off-railway counties, the railway counties in Qinghai are less likely to rely on
the agricultural industry and are more likely to have a bigger service industry.

B. Baseline Model

We apply the DD method to estimate the effects of the Qingzang railway.
In particular, we compare the difference in the outcomes between the treat-
ment group and the control group before and after the railway began oper-
ating. In the baseline model, the treatment group includes all the counties
along the railway, while the control group includes all the off-railway counties
in these two provinces.

The first model we consider (model 1) can be described as follows:

\[ \ln Y_{it} = \beta_0 + \gamma D_i \times T + \beta_1 T + \beta_2 Z_i + \delta X_{it} + \epsilon, \]  

(1)

where \( Y_{it} \) is the outcome of county \( i \) in period \( t \), such as the GDP per capita and
the value added in different industries; \( D_i \) is a dummy variable indicating the
treatment status (1 for railway counties and 0 for off-railway counties); \( T \) is the
indicator of the posttreatment period, with 1 for the posttreatment period and
0 for the pretreatment periods; and \( Z_i \) represents the indicator of county \( i \).
It controls for the fixed effects of time-invariant county characteristics, such as
whether the county is along the railway, county area, and whether the county is
the provincial or prefectural capital. Variable \( X_{it} \) is a list of control variables that
change with time, such as the provincial consumer price index (CPI; year 1997 =
100), and \( \epsilon \) is the residual.

In this model, \( \beta_1 \) represents the common trend of the outcome for all
counties in the sample periods, \( \beta_2 \) represents the time-invariant difference be-
tween counties, and \( \gamma \) represents the effect of the Qingzang railway on the out-
come of the treated counties. The identification assumption is that the change
in the outcome should be parallel between the treatment and the control groups
in the absence of the railway. However, this assumption can be problematic. For
example, the railway counties are on average richer than the off-railway coun-
ties, and this gap may change due to reasons other than the arrival of the rail-
way. To minimize the bias, we allow the trend of the outcome to vary with some
initial status, such as the GDP per capita before the treatment. Model 1 can then be revised to the following model (model 2):

$$\ln Y_{it} = \beta_0 + \gamma D_i \times T + \beta_1 T + \beta_2 Z_i + \delta X_{it} + \theta Z_{i0} \times T + \epsilon,$$  

(2)

where $Z_{i0}$ is the pretreatment variable that affects the potential time trend of the outcome. Unless otherwise specified, we focus on the GDP per capita and population in the earliest pretreatment year with data available. Since model 2 allows for different time trends of the outcome, the identification assumption is weaker in model 1. Thus, the estimates may be more reliable. Although $\theta$ may incorporate some of the effects of the railway and, in this fashion, model 2 may underestimate the railway effect, the inclusion of $Z_{i0} \times T$ is important in order to rule out the possibility that the positive correlation between railway and local economy resulted from unparalleled time trends between treatment and control counties.

One caveat here is that even in model 2, the estimates of $\gamma$ can still be up-biased due to omitted variables that are positively correlated with both GDP and the location of the railway. For instance, the existence of contemporary changes mentioned in Section II can still contaminate the estimate of the railway effect, although we do not expect a substantial overlap between counties affected by the contemporary changes and counties along the railway. The bias is more likely to be upward because major cities such as the provincial or prefectural capitals are more likely to be passed by the railway, and at the same time the contemporary changes, particularly the road improvement and the establishment of Lhasa Economic and Technological Development Zone, are more likely to have stronger effects on these major cities. Given that it is impossible to control for all the contemporary changes, we interpret our estimate as the upper-bound effect of the railway on the local economy.

V. Results of the Baseline Model

We first focus on the effects of the operation of the Qingzang railway. To that end, we now compare the outcomes post-2007 and those pre-2000.

A. Economic Development

We first examine the effect of the Qingzang railway on the GDP per capita. As there is no information on the GDP of Tibet before 1999, we can only take into account 1999–2000 and 2007–9. To allow the time trend of the outcome to vary with the initial conditions, we treat 1999 as the initial year and compare the GDP per capita in 2000 to that in 2007–9. Although the sample size
is not balanced between the pre- and posttreatment periods, the robustness test shows that the results are not sensitive to this problem.

Table 1 shows that the Qingzang railway has a substantial and significant effect on the log value of GDP per capita of the railway counties. In column 1, we apply model 1, in which the counterfactual trends of GDP per capita are assumed to be the parallel between the treated and untreated counties. The result shows that the railway increased GDP per capita by 27%, and the effect is significant at the 5% level. In column 2, we use model 2, which allows the trend of GDP per capita to vary with initial GDP per capita and population. In comparison to model 1, model 2 gives a larger and more significant estimate for the railway effect: the railway increased GDP per capita by 36%, which is significant at the 1% level.

Given that the average GDP per capita was approximately RMB 5,400 and the average population was about 2.1 million among the treated counties before the construction of the railway (year 2000), the estimate in column 2 implies that the railway raised the GDP by RMB 4.1 billion in 1 year, which means that the cost of the railway would be covered in 8 years. As the railway was opened to the public in July 2006, it is better to calculate the benefits on the basis of the counterfactual level of GDP in 2007. Assuming that the treated counties would have grown at the same rate as the untreated counties, the counterfactual GDP of the railway counties would have been RMB 34 billion in 2007. The increase in GDP due to the operation of the railway would then be about RMB 12.2 billion a year. Following this calculation, the cost of the railway would be covered in less than 3 years.

The coefficient of the posttreatment dummy indicates a significant growth in GDP per capita for all counties. This common time trend incorporates the railway effect on the GDP per capita of the untreated counties. If the untreated counties are positively (negatively) affected by the railway, the railway effect on the railway counties is underestimated (overestimated). We return to this issue in Section V.C, where we find that the GDP per capita of the untreated counties was not significantly affected by the Qingzang railway.

The difference in the estimates between models 1 and 2 indicates that the treatment counties grew at a slower rate than the control counties in the absence of the railway. This results from the fact that, on average, the treatment counties have higher initial GDP per capita, and the coefficient of the interaction term between the initial GDP per capita and the posttreatment dummy is negative. Given that the railway is unlikely to reduce GDP, we argue that model 1 tends to underestimate the railway effect on GDP per capita. Hence, hereafter, we focus on model 2 to estimate the railway effect on GDP per capita.
## Table 1


<table>
<thead>
<tr>
<th>Qinghai Province</th>
<th>Log (GDP per Capita)</th>
<th>Log (Population)</th>
<th>Log (GDP per Capita)</th>
<th>Log (Population)</th>
<th>Log (GDP per Capita)</th>
<th>Log (Population)</th>
<th>Log (GDP per Capita)</th>
<th>Log (Population)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post x railway counties</td>
<td>Post</td>
<td>Post</td>
<td>Post</td>
<td>Post</td>
<td>Post</td>
<td>Post</td>
<td>Post</td>
<td>Post</td>
</tr>
<tr>
<td>Log (GDP per capita)</td>
<td>0.36***</td>
<td>0.11***</td>
<td>0.10***</td>
<td>0.38***</td>
<td>0.046***</td>
<td>0.033***</td>
<td>0.28***</td>
<td>0.029***</td>
</tr>
<tr>
<td>Log (Population)</td>
<td>0.33***</td>
<td>0.064***</td>
<td>0.053***</td>
<td>0.31***</td>
<td>0.051***</td>
<td>0.037***</td>
<td>0.26***</td>
<td>0.022***</td>
</tr>
<tr>
<td>Growth rate of per capita GDP</td>
<td>-0.21***</td>
<td>-0.086***</td>
<td>-0.053***</td>
<td>-0.28***</td>
<td>-0.074***</td>
<td>-0.022***</td>
<td>-0.26***</td>
<td>-0.074***</td>
</tr>
<tr>
<td>Log (length of highways)</td>
<td>0.044***</td>
<td>0.033***</td>
<td>0.037***</td>
<td>0.044***</td>
<td>0.033***</td>
<td>0.037***</td>
<td>0.044***</td>
<td>0.033***</td>
</tr>
<tr>
<td>Log (provincial CPI)</td>
<td>1.59***</td>
<td>0.73***</td>
<td>0.73***</td>
<td>1.59***</td>
<td>0.73***</td>
<td>0.73***</td>
<td>1.59***</td>
<td>0.73***</td>
</tr>
<tr>
<td>Constant</td>
<td>8.89***</td>
<td>9.02***</td>
<td>9.02***</td>
<td>8.89***</td>
<td>9.02***</td>
<td>9.02***</td>
<td>8.89***</td>
<td>9.02***</td>
</tr>
<tr>
<td>Observations</td>
<td>452</td>
<td>452</td>
<td>452</td>
<td>452</td>
<td>452</td>
<td>452</td>
<td>452</td>
<td>452</td>
</tr>
<tr>
<td>R²</td>
<td>0.910</td>
<td>0.915</td>
<td>0.917</td>
<td>0.921</td>
<td>0.973</td>
<td>0.929</td>
<td>0.938</td>
<td>0.929</td>
</tr>
</tbody>
</table>

**Note.** Robust standard errors in parentheses. CPI1997 = 100. All regressions control for county fixed effects.

* p < .1
** p < .05
*** p < .01
The latent trend of per capita GDP may also vary with the initial status of other factors. Table 1 column 3 allows the trend to vary with the initial growth rate of GDP per capita by controlling for the interaction term between the posttreatment dummy and the initial growth rate. Since we only have pretreatment data in 1999 and 2000 for Tibet, we use the growth rate in 2000 as the initial growth rate. The estimate for the railway effect is the same as that in column 2. The initial growth rate itself shows a significant negative effect. However, we need to be cautious on the interpretation because the initial growth rate in this specification is for 2000, and GDP per capita in 2000 is on the right-hand side of the regression. This potential problem also compels us to exclude this variable in the regressions later.6

We are also interested in real GDP per capita. Since the county-level price index is not available, we control for the log value of the provincial price (CPI) in column 4. Compared with column 2, the estimate for the railway effect on GDP per capita decreases to 33%, which is significant at the 1% level. The regression of the deflated GDP per capita (using the provincial CPI to deflate) delivers a similar result: 34% in model 2 and significant at the 1% level. These results imply that the railway tends to have a positive effect on the price levels in the railway counties. However, we do not have information on the county-level CPI to directly verify this implication. The aggregate data show that after the completion of the railway (2007–9), the inflation rates in both provinces were higher than the national average (3.5% for Tibet and 6.5% for Qinghai vs. 3.3% for the national average).

Table 1 column 5 estimates the railway effect on the total amount of the GDP that is affected by the change in population. The change in GDP per capita concerns the change in the economic welfare of each person, while the change in population indicates whether more people have reaped the benefits of Qingzang railway. The result shows that the total GDP increased by 35% due to the railway. This increase is quite similar to the increase in GDP per capita, implying that only a slight change occurred in the population after the railway began to operate.

Column 6 confirms that the effect of the railway on the population of railway counties is insignificant. This result is not surprising given that the Hukou system is still strictly enforced and it is difficult to migrate permanently in China (Chan and Buckingham 2008). However, temporary mi-

---

6 We also tried other specifications, including allowing the trend of GDP per capita to depend on whether the county is a provincial or prefectural capital; allowing the time trends in 2007, 2008, and 2009 to be different; controlling for county areas; and including the observations in 1999. The results are all similar.
grants who do not have local Hukou are not considered in the population statistics. Therefore, it is possible that the railway attracted more people to the railway counties, but these migrants were mostly temporary.

Since Qinghai has more information available than Tibet, we present estimates for Qinghai in the last two columns of table 1. Column 7 uses the same specification as column 4 and shows that the railway stimulated the GDP per capita by 55% for the railway counties in Qinghai. The magnitude is much higher than that in column 4, implying that the railway effect in Qinghai is stronger than that in Tibet.\(^7\) Column 8 exploits more information from Qinghai, specifically, using data in earlier years (1997, 1998) and controlling for more initial conditions, including the average growth rate in 1998 and 1999, as well as the initial human capital, proxied by the proportion of middle school students over the population. We also add current highway length to control for road development. Results show that the railway effect remains strong (64%) and significant at 1% level.\(^8\) The results for GDP are similar to those for GDP per capita because the railway did not affect the population in Qinghai. Given that adding more control variables delivers similar estimates as the baseline model and these additional controls are all insignificant, we continue using the baseline specification.

**B. Different Industries**

As illustrated in our model and studies on the US transportation system, transportation can have different impacts across industries. To examine which sector benefits the most from Qingzang railway, we estimate the effect of the railway on the value added in the agriculture, manufacturing, and service industries. We focus on the railway effect in Qinghai province because of the lack of data on these industries in Tibet. We use the data for 1998–2000 and 2007–9 for the estimation and use the information in 1997 as the initial status. Table 2 reports the estimates for models 1 and 2.

For the agricultural industry, both models show that the railway had no significant effect (cols. 1 and 2). By contrast, the railway significantly stimulated the manufacturing industry (cols. 3 and 4). The magnitude is 78% in model 2 and significant at the 1% level. Columns 5 and 6 show that the railway effect on service was insignificant. Specifically, model 1 indicates a positive

---

\(^7\) The difference is significant, as confirmed by the regression that compares the railway effect between these two provinces (not reported due to space limitations). However, since the railway passes through only five counties in Tibet, we should be cautious in interpreting the significant difference between the two provinces.

\(^8\) If only the number of pretreatment periods increases, the estimate of the railway effect increases to 57%. 
railway effect significant only at the 10% level, and the more reliable model 2 shows that the effect is not significant at the 10% level.

These results are consistent with the prediction of our theoretical model. The railway effect on the local manufacturing industry is likely to be positive because the reduction in the transportation cost increases the competitiveness of the local firms in the national market. For the agriculture and service industries, since the competition is more likely to be in the regional market, the railway effect tends to be much smaller than that on the manufacturing industry. The reductions in the cost of intermediate goods, labor, and technology adoption all amplify the difference across industries. The tourism industry, one subindustry of the service industry, was expected to boom after the railway began to operate. However, the effect might be concentrated on particular destinations, such as Lhasa and Xining. As a result, the railway effect on the service industry was on average modest.

C. Spillover Effects

Previous analyses all assume that the untreated counties are not affected by the railway. However, the railway could have positive or negative spillover effects on the off-railway counties within the same region as the railway counties. As
illustrated in our theoretical model, negative spillover effects are attributed to the business-displacement effect, in which businesses and other economic activities shift from the off-railway counties to the railway counties. It could also result from a brain-drain effect, in which the railway counties attract human capital from the off-railway counties.

The negative spillover effect implies that the economic growth in the treated counties could occur at the cost of the untreated neighbors. By contrast, the positive spillover effect implies that the railway could also benefit the counties that are not directly connected to the railway. The existence of the spillover effect also means that our baseline estimates for the railway effect on the treated counties could have bias. By gauging the magnitude of the spillover effect, we can measure the potential bias of the baseline results.

To estimate the spillover effect, we distinguish four types of counties (fig. 3): the dark gray railway counties (20 counties), the hatched off-railway counties (23 counties) that neighbor the railway counties and are within the same prefecture as the railway counties, the dotted off-railway counties within Qinghai and Tibet (73 counties), and the light gray counties (27 counties) that are located in provinces other than Qinghai and Tibet and are adjacent to off-railway counties but not to railway counties. Presumably, the spillover effects decrease as the distance to the railway increases. The light gray counties are not only farther away from the railway but also under different provincial governance. Therefore, they are least likely to be affected by the railway, and we use them as our control group in the spillover analysis. The light gray counties are economically similar to the adjacent counties. Specifically, the GDP per capita in 2000 was 5,564, 3,032, 3,376, and 3,472 yuan for the dark gray, hatched, dotted, and light gray counties, respectively.

Column 1 of table 3 estimates the spillover effect of the Qingzang railway on off-railway counties. Specifically, we apply the DD model similar to the baseline specification, with the hatched and the dotted counties being the treatment group and the light gray counties being the control group. The result shows a nonsignificant average spillover effect of the Qingzang railway on the GDP per capita of off-railway counties. Column 2 analyzes whether the hatched counties that are closer to the railway counties experienced a significant spillover effect. By examining the double difference between the hatched and the light gray counties, we still find no significant spillover effect of the railway on GDP per capita. Using the dotted counties as the control group delivers similar results.

Columns 1 and 2 increase our confidence in the baseline model that uses the dark gray railway counties as the treatment group and the hatched and the dotted counties as the control group. The argument is further supported by
Figure 3. Counties and the path of Qingzang railway. Thick (thin) lines are boundaries of provinces (counties). Counties in dark gray are the railway counties or counties along the railway. Counties in hatched area are adjacent to railway counties in Qinghai and Tibet. Dotted area contains the remaining counties in the provinces of Qinghai and Tibet. Light gray counties are in other provinces and adjacent to off-railway counties (hatched or dotted) and not to railway counties (dark gray). Circles stand for county seats. Triangles represent railway stations.
column 3, in which we reestimate the baseline model and use the light gray untreated counties as the control group for the dark gray railway counties. The estimate of the railway effect on these counties is 31\% and significant at the 1\% level, which is almost the same as the baseline estimate.

Columns 4–6 show that no spillover effect is found on the registered population either, and the baseline estimate is reliable. Specifically, column 4 shows that the railway on average did not change the population in the hatched and the dotted counties. Column 5 shows that even in the hatched counties, the railway had no significant effect on the population. Column 6 confirms that the railway effect on the population in the railway counties is still not significant when we use the light gray counties as the control group.

Table 4 examines the spillover effect for different industries. Columns 1 and 2 (cols. 4 and 5) indicate that Qingzang railway had no significant spillover effect on the agricultural (manufacturing) industry, and columns 3 and 6 suggest that the baseline estimates in table 2 are reliable. For the service industry, we find a significant negative spillover effect of the railway (table 4 col. 7). As expected, this negative spillover effect is stronger in the hatched counties than in the dotted counties (col. 8). When the light gray counties are used as the control group, the estimate of the railway effect on the service industry becomes negative for railway counties (col. 9). The estimate is still insignificant, implying that the estimate in the baseline model is qualitatively correct.

To summarize, we find no evidence for the spillover effect of the Qingzang railway on GDP per capita, population, agriculture, and manufacturing in
TABLE 4


<table>
<thead>
<tr>
<th></th>
<th>Log (Value Added in Agriculture)</th>
<th>Log (Value Added in Manufacturing)</th>
<th>Log (Value Added in Service)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hatched versus Light Gray</td>
<td>Hatched versus Light Gray</td>
<td>Dark Gray versus Light Gray</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td></td>
<td>Hatched versus Light Gray</td>
<td>Hatched versus Light Gray</td>
<td>Dark Gray versus Light Gray</td>
</tr>
<tr>
<td></td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td></td>
<td>Hatched versus Light Gray</td>
<td>Hatched versus Light Gray</td>
<td>Dark Gray versus Light Gray</td>
</tr>
<tr>
<td></td>
<td>(7)</td>
<td>(8)</td>
<td>(9)</td>
</tr>
<tr>
<td>Post x treated counties</td>
<td>-.010 (.048)</td>
<td>-.19 (.13)</td>
<td>-.19 (.074)</td>
</tr>
<tr>
<td></td>
<td>-.018 (.063)</td>
<td>.064 (.16)</td>
<td>-.28 (.087)</td>
</tr>
<tr>
<td>Post</td>
<td>2.13*** (.60)</td>
<td>7.01*** (.33)</td>
<td>1.31* (.71)</td>
</tr>
<tr>
<td></td>
<td>3.44*** (.68)</td>
<td>7.51*** (.59)</td>
<td>2.50*** (.77)</td>
</tr>
<tr>
<td></td>
<td>3.43*** (1.22)</td>
<td>5.53*** (1.62)</td>
<td>2.56*** (.66)</td>
</tr>
<tr>
<td>Observations</td>
<td>278</td>
<td>278</td>
<td>278</td>
</tr>
<tr>
<td></td>
<td>188</td>
<td>188</td>
<td>188</td>
</tr>
<tr>
<td></td>
<td>.963</td>
<td>.942</td>
<td>.965</td>
</tr>
<tr>
<td></td>
<td>.967</td>
<td>.945</td>
<td>.962</td>
</tr>
<tr>
<td></td>
<td>.932</td>
<td>.911</td>
<td>.974</td>
</tr>
</tbody>
</table>

Note. Robust standard errors in parentheses. CPI1997 = 100. All regressions use the specification of model 2, in which the trend of the outcome is allowed to vary with the initial per capita GDP and population. All regressions control for county fixed effects and provincial CPI.

* p < .1.
** p < .05.
*** p < .01.
off-railway counties in Qinghai and Tibet. Some displacements are found in the service industry, but the baseline model still delivers the right qualitative conclusion. These results are mostly consistent with the prediction of our theoretical model. The manufacturing industry in the off-railway counties may lose consumers to railway counties in the regional market but gain consumers in the national market. The gain in the national market is likely missing for agricultural and service industries. For agriculture in China, the existence of a labor surplus and the immobility of the crucial input (land) help explain the insignificance of the displacement.

VI. Robustness
A. Counterfactual Tests
The main concern of the DD method is that different secular trends or contemporary events could have led to different levels of economic growth in the treated and untreated counties. For instance, Qingzang Road has a persistent influence, as well as the other newly built roads, such as Xihuang Road, Huangdao Road, Xima Road, and Xida Road. To examine the potential bias, we conduct the counterfactual tests.

More specifically, we expect little or small change in the railway effect on GDP within the three periods of 1999–2000, 2002–4, and 2007–9, while the influence of other events remained present throughout these periods. Specifically, from 1999 to 2000, no railway effect existed because the construction of the railway had not started. From 2002 to 2004, the railway remained under construction, such that the change in GDP per capita should be weak if the influence of the construction did not change substantially over time. From 2007 to 2009, the Qingzang railway was in operation. Hence, the change in GDP per capita should also be weak if the effect of the operation of the railway grew slowly over time.

The first three columns of table 5 apply model 2 to each of the respective periods and find no significant difference in GDP per capita between the railway counties and other counties. These results increase our confidence that the baseline estimates of the railway effect are reliable.

9 There are six subindustries in the service industry. We do not have county-level information on the subindustries to analyze what drives the displacement. The theoretical analysis predicts that the displacement effect in an off-railway county is likely to be driven by its subindustries that compete within the region and have a higher MC than the corresponding industries in the railway counties.

10 These four roads are from Xining to Huangyuan (completed in 2003), Huangyuan to Daotanghe (completed in 2003), Xining to Machangyuan (completed in 2003), and Xining to Datong (completed in 2004). Most parts of these regional roads do not overlap with the route of the railway.

11 For all the robustness tests, we only show the results for model 2 to save space. The results based on model 1 are similar, but they are relatively less robust.
### TABLE 5
COUNTERFactual TESTS AND ROBUSTNESS TESTS

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) (2) (3) (4) (5) (6) (7) (8)</td>
</tr>
<tr>
<td>Post × railway counties</td>
<td>.028 .043 .089</td>
</tr>
<tr>
<td>Post</td>
<td>.82 .50 .23</td>
</tr>
<tr>
<td>Constant</td>
<td>−92.1 62.1** 7.41</td>
</tr>
<tr>
<td>Observations</td>
<td>225 227 228</td>
</tr>
</tbody>
</table>

**Note.** Robust standard errors in parentheses. Dependent variable = log (GDP per capita). CPI1997 = 100. All regressions use the specification of model 2, in which the trend of the outcome is allowed to vary with the initial per capita GDP and population. All regressions control for county fixed effects and provincial CPI.

** p < .05.

*** p < .01.
B. More Comparable Treatment and Control Groups

The foregoing analyses mostly define all the off-railway counties as the control group. For the robustness tests, we restrict the control group to untreated counties that are similar to the treated counties. Columns 4–7 in table 5 show that the estimates based on model 2 are quite robust to different choices of the control group.

As many economic policies are formulated at the prefecture level in China, column 4 restricts the control group to the off-railway counties that are in the same prefecture as the railway counties. The result shows that the effect of the railway on GDP per capita is 0.29 and still significant at the 5% level.

Next, we use the propensity score matching method to select the control group. We first estimate the probability of being treated by applying the probit model to regress the binary variable of a county being along the railway on its characteristics in 2000. The probit estimation is shown in table B2. Table 5 column 5 excludes the counties whose probability of being treated is out of the common support of the probability. In our case, only one county is excluded. The estimate for the railway effect is 0.32 and significant at the 1% level.

Column 6 uses the five nearest neighbors matching with replacement to estimate the railway effect. This method gives a smaller sample: 18 counties in the treated group, 31 counties in the control group. The reduction in bias achieved by matching is substantial, as shown in table B2. The estimate is still similar to that in the baseline model. Specifically, the railway effect on GDP per capita is now 0.31 and significant at the 1% level.

Finally, we use the weighted least squares method, which combines matching with regression. In particular, the method assigns a weight to each county: 1 for the treated counties and propensity score / (1 − propensity score) for the control counties. Imbens and Wooldridge (2009) highlight that this method attains “double robustness.” As long as the parametric model for either the propensity score or the regression function is specified correctly, the resultant estimate for the average treatment effect on the treated group will be consistent. This method does not reduce sample size but concurrently considers the similarity between the treated and control counties. The result in table 5 column 7 confirms that the Qingzang railway stimulated the economic growth of the railway counties. The coefficient is 0.42 and significant at 1% level, larger than that in column 5. The increase in the magnitude results from the fact that big counties in the control group are now assigned a large weight because of their high propensity of being treated. As a result, the average GDP growth of the control group is reduced because large counties generally have a low growth rate (as indicated by the negative coefficient of the interaction term between initial GDP per capita and post).
C. Sensitivity to Treatment Definition

The main concern of the nonrandom placement of a railway comes from the fact that the railway is more likely to pass through major cities, such as the provincial and prefectural capitals, as shown in table B2. As suggested in the literature, we can reduce the bias by excluding these major cities and focusing on nonmetropolitan or nontargeted counties in the regression. Column 8 of table 5 shows the result. The estimate of the railway effect on the local GDP per capita actually increases to 43\% (significant at the 5\% level), implying that if our baseline estimate is biased, the estimate tends to be downward.\textsuperscript{12}

Previous analyses mostly define the treatment status of a county on the basis of whether the railway passes through the county. However, it is possible that the railway passes through the edge of a county, and the railway is far from the center of the county. Meanwhile, the center of an off-railway county may be very close to the railway and potentially to a station. Therefore, a better definition of the treatment status can be the distance from a county seat (circles in fig. 3) to the nearest railway station (triangles in fig. 3). We use GIS data and satellite maps to obtain the distance. The average distance to a station in the 116 counties in Tibet and Qinghai is 246 kilometers, with a maximum of 1,138 kilometers and a minimum of less than 1 kilometer.

We first define the treatment to be the continuous distance from the county capital to the nearest station. The first row in table 6 shows that when the distance increases by 1\%, the positive railway effect on GDP per capita declines significantly by 0.11\%. This elasticity is larger than the estimate in Banerjee, Duflo, and Qian (2012). Next, we choose a cutoff value of the distance, say 10 kilometers, and define counties whose capital is within 10 kilometers from the nearest station as the treatment group and other counties as the control group. Then, we apply model 2 to estimate the railway effect on GDP per capita for the treated counties. Table 6 shows the results for different cutoff values. Not surprisingly, the estimate of the railway effect decreases as the cutoff value increases. The estimates decrease from 44\% for counties within 20 kilometers to 19\% (still significant) for counties within 300 kilometers. The railway effect remains significant until the distance increases to 400 kilometers.

The estimate in the baseline model is similar to the result using 30 kilometers as the cutoff, which shows a 33\% railway effect on GDP per capita with a significance level of 1\%. Given that the average distance to the nearest

\textsuperscript{12} The coefficient of the interaction term "provincial or prefectural capitals $\times$ post $\times$ treated" in col. 2 of table 7 shows that the railway effect on major cities is weaker than that on other railway counties, but the difference is not statistically significant.
station is 31 kilometers for railway counties in the baseline model, this similarity suggests that the baseline model is reliable.

VII. Heterogeneity and Dynamics of the Railway Effect

A. Different Effects among Treated Counties

The effects of transportation infrastructure can vary with the characteristics of local counties. Table 7 examines how the treatment effect differs among the railway counties. We first examine how the railway effect varies with the initial economic status of the counties. We divide all 116 counties into two groups according to their GDP per capita in 2000, with the top half comprising the rich group and the bottom half the poor group. Among the railway counties, 14 are in the rich group and six are in the poor group. For the off-railway counties, the corresponding numbers are 44 and 52. Column 1 interacts the railway effect with an indicator of whether a county is rich. We also add an interaction term between “being rich” and “posttreatment” to allow the two income groups to have different time trends. The result shows no significant difference in the railway effect between the rich and poor counties.

Second, the railway effect could vary across counties with different administrative levels. Column 2 adds an interaction term between the indicator for provincial or prefectural capitals and the railway effect. We find that the
railway effect tends to be weaker in the provincial or prefectural capitals than in other counties; however, the difference is insignificant.

We then allow the railway counties along the second section of the Qingzang railway to have a railway effect different from the railway counties in the first section. The conjecture is that the latter may have been more developed to attract new investments than the former group. Column 3 shows that the railway effect is slightly weaker in counties along the second section, but the difference is not significant. This result indicates that the early construction (or development) of first section did not bring in significant advantages for the first section.

Finally, we distinguish the railway counties that have no passenger stations from other railway counties. Column 4 shows, as expected, that the railway

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rich Counties versus Poor Counties</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Post × treated counties</td>
<td>.40***</td>
</tr>
<tr>
<td>Rich × post × treated</td>
<td>−.086</td>
</tr>
<tr>
<td>Rich × post</td>
<td>−.15</td>
</tr>
<tr>
<td>Provincial or prefectural capitals × post × treated</td>
<td>−.34</td>
</tr>
<tr>
<td>Provincial or prefectural capitals × post</td>
<td>.31**</td>
</tr>
<tr>
<td>Second section × post × treated</td>
<td>−.19</td>
</tr>
<tr>
<td>No passenger stations × post × treated</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1.75</td>
</tr>
<tr>
<td></td>
<td>(1.52)</td>
</tr>
<tr>
<td>Observations</td>
<td>452</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.923</td>
</tr>
</tbody>
</table>

Note. Robust standard errors in parentheses. Dependent variable = log (GDP per capita). CPI$_{1997} = 100$. All regressions use the specification of model 2, in which the trend of the outcome is allowed to vary with the initial per capita GDP and population. All regressions control for county fixed effects and provincial CPI. ** p < .05. *** p < .01.
counties without passenger stations experienced a significantly smaller railway effect than other railway counties. For the railway counties with passenger stations, the railway stimulated the GDP per capita by 43%. In contrast, the railway had no significant effect on the GDP per capita of the railway counties without passenger stations.

B. Dynamics of the Railway Effect

Previous analyses focus on the average railway effect after the railway has begun operating. However, even before the railway opened to traffic, people’s expectations had changed. The expectations of new opportunities for investment and employment after the railway began operating may have induced early capital investment or labor migration. Therefore, it is worth examining the dynamics of the railway effect after the Qingzang railway project was officially approved in 2001.

We start with the construction periods. The construction of the Qingzang railway took more than 4 years and involved more than 20,000 workers and over 6,000 pieces of industrial equipment (Ministry of Railways 2001). Construction of such a big project definitely demanded local services and could have stimulated the local economy. We apply the DD method on data from 2000 and 2002 to 2004 to estimate the construction effect of the railway on GDP per capita.

The first four columns in table 8 show that during the construction period, the GDP per capita in railway counties was on average higher than that in other counties. Column 1 uses the data for 2000 and 2002. The result shows that the effect of the construction on GDP per capita was positive but insignificant at the 10% level in 2002. Column 2 uses the data for 2000 and 2003, and shows that the construction effect was about 18% and significant at the 1% level in 2003. Column 3 uses the data for 2000 and 2004, and shows that the GDP per capita of the railway counties increased by 16% in 2004, which is significant at the 10% level.

In column 4, we test whether the change over time is significant. We pool all these years together and allow the railway effect to differ over time by adding the interaction terms between the year dummies and the indicator of treated counties. We also allow the common time trend to differ in 2002, 2003, and 2004. The coefficients of “post-2003 × treated” and “post-2004 × treated” show no significant variation over time in terms of the construction effect of the railway on GDP per capita.

We notice that the effect of construction was much smaller than the railway effect in the operation stage, implying that the construction effect was not the driver of the railway effect in the operation stage. Moreover, the short-run
<table>
<thead>
<tr>
<th></th>
<th>Construction Period</th>
<th>Operation Period</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td></td>
<td>(5)</td>
</tr>
<tr>
<td>Post-2002 × treated counties</td>
<td>.12*</td>
<td>.18***</td>
<td>.16*</td>
<td>.14**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.070)</td>
<td>(.070)</td>
<td>(.097)</td>
<td>(.068)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-2003 × treated counties</td>
<td>.063</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.056)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-2004 × treated counties</td>
<td>−.065</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.074)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-2007 × treated counties</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.35***</td>
<td>.43***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(.11)</td>
<td>(.13)</td>
</tr>
<tr>
<td>Post-2008 × treated counties</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.18*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(.096)</td>
<td></td>
</tr>
<tr>
<td>Post-2009 × treated counties</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>−.038</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(.11)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>14.6**</td>
<td>16.7***</td>
<td>27.8***</td>
<td>20.8***</td>
<td>24.4***</td>
<td>15.8***</td>
</tr>
<tr>
<td></td>
<td>(7.18)</td>
<td>(4.85)</td>
<td>(7.53)</td>
<td>(4.96)</td>
<td>(5.51)</td>
<td>(4.37)</td>
</tr>
<tr>
<td>Observations</td>
<td>225</td>
<td>224</td>
<td>224</td>
<td>451</td>
<td>225</td>
<td>224</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.926</td>
<td>.931</td>
<td>.862</td>
<td>.858</td>
<td>.940</td>
<td>.942</td>
</tr>
</tbody>
</table>

**Note.** Robust standard errors in parentheses. Dependent variable = log (GDP per capita). CPI<sub>1997</sub> = 100. All regressions use the specification of model 2, in which the trend of the outcome is allowed to vary with the initial per capita GDP and population. All regressions control for county fixed effects and provincial CPI. Dummies for each year are included to capture potential trend changes.

* $p < .1$.

** $p < .05$.

*** $p < .01$. 

This content downloaded from 128.252.067.066 on July 26, 2016 06:41:16 AM
All use subject to University of Chicago Press Terms and Conditions (http://www.journals.uchicago.edu/t-and-c).
stimulation of the local economy resulting from the construction was likely to have disappeared after the completion of the railway.

Next, we examine whether the railway effect changed over time after the railway began operating. To this end, we estimate the railway effect separately for 2007, 2008, and 2009. Table 8 column 5 shows that the railway effect on GDP per capita in 2007 was 0.35. The coefficient increases to 0.43 in 2008 (col. 6) and 0.44 in 2009 (col. 7). All the coefficients are significant at the 1% level. To test whether the difference over time is significant, we pool the observations together and add interactions between the railway effect and year dummies. Column 8 shows that the effect of the railway was stronger in 2008 (only significant at the 10% level). However, the effect stopped increasing in 2009.

VIII. Conclusion and Policy Implications

The natural experiment of the Qingzang railway provides empirical evidence that transportation infrastructure can have a significant causal effect on economic development for the undeveloped regions. The results of this study indicate that the Qingzang railway increased the GDP per capita of the railway counties by about 33%. The economic benefits of the railway could be as high as RMB 12 billion a year (11% of total GDP in Qinghai and Tibet in 2007), in which case less than 3 years would be needed to recover the cost of the construction of the railway. The institutional background of the Qingzang railway and results of the counterfactual and robustness tests all indicate that the estimates are reliable.

The positive effect of the Qingzang railway on GDP worked mainly through the positive railway effect on the manufacturing industry that is likely to compete in the national market. The agricultural sector and service industry, which are likely to compete mainly in the regional market, were barely affected by the new railway. These results indicate that the railway effect depends on the industry structure of the local economy. In contrast, we find that the railway effect did not vary with the initial economic status of the local economy.

We also find that the railway did not affect the number of permanent residents, maybe because of the Hukou system. The GDP per capita of the off-railway counties was not significantly affected by the railway. The construction of the railway also had a positive effect on the local economy. The changes in the railway effect within the construction and the postoperation periods were modest and mostly insignificant.

13 These coefficients are larger than those in the baseline results because of the change in the sample size.
Although we find that the Qingzang railway has a positive effect on the local economy, two reasons caution our generalizing the conclusion. First, selection bias can still occur on time-variant unobservables. Other contemporary shocks may also contaminate the estimates. These problems are always encountered outside a randomized experiment.

Second, the large positive effect of the Qingzang railway depends on the facts that Tibet and Qinghai are among the most underdeveloped provinces in China and Qingzang railway is the only railway connecting Tibet with other provinces. Although we find no significant differences in the railway effect between counties with different initial economic conditions, this finding may not hold when the scope of the variation in economic conditions is expanded beyond this region. The effect of the transportation infrastructure may be nonlinearly correlated with the initial economic conditions of a region. The effect of new infrastructure can be very strong when the existing infrastructure lags behind local economic development. However, if the investment in infrastructure exceeds the level of economic development, the new infrastructure can have little effect on economic growth.

The Chinese government plans to add six more railway lines to the Qinghai-Tibet region to further boost the economy of the region. The results of our study provide economic justification for this kind of government-initiated investment. Nonetheless, the economic effects of these additional railways warrant further research.

Appendix A
Proof of the Theoretical Model

The First Case
Denote the distance between A and B as $D_{AB}$. There exists a consumer $i$ who locates between A and B and is indifferent between purchasing from these two firms. Denote the distance between A and $i$ as $D_{Ai}$. Obviously, firm A will attract all the consumers whose distance from firm A is less than $D_{Ai}$. Hence, $D_{Ai}$ represents the location of the market boundary and $dD_{Ai}/dt_1$ represents the negative of the railway effect on firm A’s output.

For the indifferent or marginal consumer $i$, the following equation must be true:

$$P_A + t_1 \times D_{Ai} = P_B + t_1 (D_{AB} - D_{Ai}).$$  \hspace{1cm} (A1)

Taking the derivative of equation (A1) with regard to $t_1$, we obtain

$$\frac{dD_{Ai}}{dt_1} = \frac{D_{AB} - 2D_{Ai}}{2t_1}. \hspace{1cm} (A1')$$
Equation (A1) implies that \( D_{A_i} \geq D_{AB} / 2 \) if and only if \( P^S_A \leq P^S_B \); that is, the firm with a lower MC acquires a larger market size. Therefore, equation (A1') implies that the railway increases the output of firm A (\( dD_{A_i} / dt \leq 0 \)) if and only if firm A has a lower MC (and thus \( P^S \)) than firm B.

**The Second Case**

This case only occurs when the consumer at point O prefers firm C or \( P^S_A + t_1 \times D_{AO} \geq P^S_C + t_2 D_{CO} \). For the marginal consumer \( j \) that locates at route 1, we have

\[
P^S_A + t_1 \times D_{A_j} = P^S_C + t_1(D_{AO} - D_{A_j}) + t_2 D_{CO}.
\]

Taking the derivative of equation (A2) with regard to \( t_1 \), we obtain

\[
\frac{dD_{A_j}}{dt_1} = \frac{D_{AO} - 2D_{A_j}}{2t_1}.
\]

Reorganizing equation (A2), we get \( D_{A_j} \geq D_{AO} / 2 \) if and only if \( P^S_A \leq P^S_C + t_2 D_{CO} \). Notice that \( P^S_C + t_2 D_{CO} \) can be interpreted as the effective MC of firm C that takes into account its transportation cost to the railway. For firms along the railway, their effective MC is equivalent to the MC of production. Therefore, the second case essentially has the same prediction as the first case; that is, the railway increases the output of a local firm A if and only if firm A has a lower effective MC than firm C. This implies that firm C can lose consumers even if its MC of production is lower than the MC of firm A.

**The Third Case**

The third case occurs only when the consumer at point O prefers firm A or \( P^S_A + t_1 \times D_{AO} \leq P^S_C + t_2 D_{CO} \). For the marginal consumer \( k \), we must have

\[
P^S_A + t_1 \times D_{AO} + t_2 \times D_{Ok} = P^S_C + t_2(D_{CO} - D_{Ok}).
\]

Taking the derivative of equation (A3) with regard to \( t_1 \), we obtain

\[
\frac{dD_{A_k}}{dt_1} = \frac{dD_{Ok}}{dt_1} = -\frac{D_{AO}}{2t_2}.
\]

This case implies that firm A always expands after the arrival of the railway. Combining the discussions for the second and third cases, we derive that when competing with an off-railway firm, firm A expands because of the railway if and only if \( P^S_A \leq P^S_C + t_2 D_{CO} \).
### TABLE B1
DESCRIPTIVE STATISTICS OF KEY VARIABLES

<table>
<thead>
<tr>
<th></th>
<th>Treated (20)</th>
<th>Untreated (96)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td><strong>A. Full sample (Qinghai and Tibet)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999–2000:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP per capita (1,000 RMB)</td>
<td>5.24</td>
<td>3.70</td>
</tr>
<tr>
<td>Population (1,000)</td>
<td>104</td>
<td>101</td>
</tr>
<tr>
<td>2002–4:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP per capita (1,000 RMB)</td>
<td>8.84</td>
<td>7.10</td>
</tr>
<tr>
<td>Population (1,000)</td>
<td>110</td>
<td>105</td>
</tr>
<tr>
<td>2007–9:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP per capita (1,000 RMB)</td>
<td>25.13</td>
<td>28.04</td>
</tr>
<tr>
<td>Population (1,000)</td>
<td>122</td>
<td>113</td>
</tr>
<tr>
<td>Rich</td>
<td>.75</td>
<td>.44</td>
</tr>
<tr>
<td>Provincial or prefectural capitals</td>
<td>.60</td>
<td>.50</td>
</tr>
<tr>
<td>Distance to stations (100 km)</td>
<td>.31</td>
<td>.68</td>
</tr>
<tr>
<td><strong>B. Qinghai province</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998–2000:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP per capita (1,000 RMB)</td>
<td>5.61</td>
<td>3.73</td>
</tr>
<tr>
<td>Population (1,000)</td>
<td>116</td>
<td>111</td>
</tr>
<tr>
<td>Value added in agriculture (million RMB)</td>
<td>70</td>
<td>74</td>
</tr>
<tr>
<td>Value added in manufacturing</td>
<td>244</td>
<td>335</td>
</tr>
<tr>
<td>Value added in service</td>
<td>324</td>
<td>381</td>
</tr>
<tr>
<td>Number of middle school students (1,000)</td>
<td>3.87</td>
<td>6.33</td>
</tr>
<tr>
<td>Length of highways (100 km)</td>
<td>3.91</td>
<td>2.88</td>
</tr>
<tr>
<td>2007–9:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP per capita (1,000 RMB)</td>
<td>29.05</td>
<td>30.36</td>
</tr>
<tr>
<td>Population (1,000)</td>
<td>137</td>
<td>122</td>
</tr>
<tr>
<td>Value added in agriculture (million RMB)</td>
<td>166</td>
<td>166</td>
</tr>
<tr>
<td>Value added in manufacturing</td>
<td>2,059</td>
<td>2,908</td>
</tr>
<tr>
<td>Value added in service</td>
<td>1,550</td>
<td>2,007</td>
</tr>
<tr>
<td>Number of middle school students (1,000)</td>
<td>5.79</td>
<td>7.57</td>
</tr>
<tr>
<td>Length of highways (100 km)</td>
<td>6.99</td>
<td>6.25</td>
</tr>
</tbody>
</table>

**Note.** Rich is a dummy that equals 1 if the county ranks in the top 50% in per capita GDP in 2000. Several counties have no information for the key variables, which are thus missing from the sample.
TABLE B2
EFFECTIVENESS OF FIVE NEAREST NEIGHBORS PROPENSITY SCORE MATCHING

<table>
<thead>
<tr>
<th></th>
<th>Probit Mean</th>
<th></th>
<th></th>
<th>t-Test</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treated</td>
<td>Control</td>
<td>%</td>
<td>Reduce</td>
<td>t</td>
</tr>
<tr>
<td></td>
<td>(2000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provincial or</td>
<td>.58 .17 .91 .6</td>
<td>.8 .0 .0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>prefectural capital</td>
<td>(.38)</td>
<td>Matched</td>
<td>.56 .50 12.6 86.3</td>
<td>.66 .51</td>
<td></td>
</tr>
<tr>
<td>Province dummy</td>
<td>.79 .30 111.2</td>
<td>.8 .51 .0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log (population)</td>
<td>.78 .80 −5.1 95.4</td>
<td>−32 .75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log (county area)</td>
<td>.12 .11 15.6 80.5</td>
<td>1.0 .32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log (GDP per capita)</td>
<td>.42 .83 71.0</td>
<td>6.46 .0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>−9.18**</td>
<td>.29</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Observations 111

Postmatching bias 17.4
% change in bias through matching −76.7
Postmatching pseudo $R^2$ 0.03
% change in pseudo $R^2$ through matching −90.3
Postmatching prob value of $\chi^2$ 0.314

Note. Robust standard errors in parentheses.

** $p < .05$.
*** $p < .01$.

References


