途上国における貧困削減と制度・市場・政策
比較経済発展論の試み
Poverty Reduction, Institutions, Markets, and Policies in Developing Countries:
Toward a Theory of Comparative Economic Development

PRIMCED Discussion Paper Series, No. 33

Highways and Development in the Peripheral Regions of China

Hangtian Xu and Kentaro Nakajima

January 2013

Research Project **PRIMCED**
Institute of Economic Research
Hitotsubashi University
2-1 Naka, Kunitatchi Tokyo, 186-8601 Japan
http://www.ier.hit-u.ac.jp/primced/e-index.html

Hitotsubashi University

Highways and Development in the Peripheral Regions of China*

Hangtian Xu[†] Kentaro Nakajima[‡]

January 2013

Abstract: This paper estimates the effects of highways (*Gaosu Gonglu*) on economic development in China's county-level cities from 1998 to 2007, a period in which China experienced sharp growth in highway mileage, using a micro level data set on industry and highway placement and the double difference propensity score matching method. After extracting the core regions, empirical estimates indicate that highway placement promotes industrial development in related cities with higher output and more investments, and these results are robust to two different checks. However, county-level cities more than 300 km away from large cities do not benefit from new highways. Furthermore, highways tend to promote the development of heavy industry but not that of light industry. Labor productivity exhibits few positive effects.

Keywords: transport infrastructure project; double difference propensity score matching (DD-PSM); regional development

JEL Classification: H54; R12

*We thank the participants of the Applied Regional Science Conference (Aomori), the China Economics Annual Conference (Ji'nan) and the IPEM workshop of Tohoku University for invaluable comments. In addition, we thank Shiro Hioki for his kind support on the data set and Ching-mu Chen for suggestions on ArcGIS. We gratefully acknowledge financial support from the Japan Society for the Promotion of Science (#22223003) and the 21st Century cultural and academic foundation. All errors are ours.

[†]Graduate School of Economics, Tohoku University, 27-1 Kawauchi, Aoba-ku, Sendai 980-8576, Japan. E-mail: hangtianxu@gmail.com

[‡]Graduate School of Economics, Tohoku University, 27-1 Kawauchi, Aoba-ku, Sendai 980-8576, Japan. E-mail: nakajima.kentaro@gmail.com

1 Introduction

China built its first limited access highway (*Gaosu Gonglu*)[1] in 1988, and since then, it has experienced a dramatic increase in highway mileage. By the end of 2011, the total highway mileage reached 85,000 km, giving China the second longest highway system in the world. Compared to ordinary roads, highways have advantages in road quality, congestion, and driving speed limit, which make transport more efficient and cheaper. It has been widely accepted that transport plays as an important role in industry location and economic development. Improvements in costs and transport efficiency promote inter-regional trade, help exploit comparative advantages, and improve population and firm mobility. Economic development is much more dependent on cross-border and inter-regional trade now than in any period in history, and highway networks play a core role in reducing shipping costs and affecting economic performance. This makes the case for an empirical investigation of highway effects. Furthermore, during the rapid process of industrialization, developing economies like China built most of their highway networks in the past two decades. This setting provides an opportunity to observe and assess the effects of road improvements on social and economic development in a modern emerging economy.

[**Figure 1**]

Thus, this paper quantitatively investigates the role of highway placement on regional growth. Access to China's micro level industry and highway placement databases allows us to estimate of the impact of the highway infrastructure on economic and social development from 1998-2007, when most of the major highway network was being developed. Then, by distinguishing between cities with and without highway placement and employing the propensity score matching method, we compare the economic performances of these two kinds of cities. Focusing on China's peripheral regions (county-level cities),[2] we find that newly placed highways stimulate industry growth in connected peripheral regions with more investments and higher output. The results are robust to a placebo test comparing the economic performances of cities without highways and cities planned to have highways in the near future and to an alternative estimate using the Mahalanobis distance matching method instead of propensity score match-

[1]For simplicity, hereafter we often refer to "limited access highway (*Gaosu Gonglu*)" as "highway" in the following sections.

[2]In some previous literature, "*Xian*" is considered to be China's county, and "*Xianjishi*" is considered to be China's county-level city. In this paper, they are classified as equivalent.

ing. Furthermore, we find that remote peripheral regions, that is, county-level cities more than 300 km from large cities, do not benefit from newly placed highways and that highway placement promotes the development of heavy industry but not that of light industry, These results suggest that highway effects are heavily dependent on geographical location and industry.

This paper contributes to the existing literature in three aspects. First, unlike Baum-Snow *et al.* (2012) and Baum-Snow and Turner (2012), which focus more on differences between core regions and peripheral regions, this paper focuses on two kinds of peripheral regions, those with and without highway placement. Although we use this sampling strategy in part because of restrictions on controlling for endogeneity, this strategy does provide another perspective on highway effects. Second, we employ a careful strategy to make causal inferences about highway treatment effects. We extend the experimentalist approach toward highway issues (Rephann and Isserman, 1994) by using the propensity score matching method and constructing a placebo test to check for potential counterfactual economic effects of highways in the absence of highway placement. Third, we provide new results. Industry is found to be sensitive to highway placement, which differs from the previous results such as Rephann and Isserman (1994) and Holl (2004b). This suggests that highway effects differ between developing and developed economies. The economies of richer countries rely less on shipping of raw materials and manufacturing goods and more on communication and human capital (Baum-Snow and Turner, 2012).

The rest of this paper is organized as follows. Section 2 contains a brief review of issues related to transport and regional development. In Section 3, we present our empirical strategy and a description of the data set. Section 4 gives the baseline results and robustness checks. Section 5 provides further estimates of the highway effects related to geographical location and industry structure. Section 6 concludes.

2 Transport and Regional Development

There has been considerable work on the role of transport infrastructure in promoting social and economic development. Along with imperfect competition and economies of scale, changing transport costs affect the locations of firms and workers (Krugman, 1991; Fujita, Krugman and Venables, 1999). By setting transport costs as endogenous, Behrens *et al.* (2009) point out that an increasing number of carriers, falling marginal costs in the transport sector, or both induce a gradual agglomeration of industrial firms, but the centripetal forces are partially balanced by price-setting behavior in the transport sector. Transport infrastructure improvements work like market integration and can change the relative importance of concentrating (market size and agglomeration economies) and dispersing forces (factor costs and competition) and consequently the spatial distribution of economic activity (Holl, 2004a). With mobility of firms, however, the distribution of benefits from infrastructure investment is not clear (Venables, 1996; Puga, 1999). On the other hand, because of decreasing transport costs and the increasing importance of non-material flows, some analysts have cast doubts on the importance of transport infrastructure on firm location. It has also been argued that in areas where the network is already very dense, the impacts on firm location might be small (Banister and Berechman, 2000).

The empirical literature on China's highway construction and its effects on social and economic development is still inconclusive. Baum-Snow *et al.* (2012) and Baum-Snow and Turner (2012) exploit the decentralization effects of highways on the urban population in the core cities of prefectures. Faber (2012) investigates the effects of highways on industrial performance and proves that, relative to core regions, highways have a negative effect on the connected peripheral regions, which is caused by *home market effects*. Descriptively, Fujita *et al.* (2004) insist that, with proper modern highway links and easier access to the coast, some inland Chinese provinces will in a sense become part of the coast. Although coastal provinces still have better access to international markets, inland provinces like Sichuan may be domestically competitive, relatively specializing in domestic products because of the enormous market potential of the large population.

In addition, highway placement may affect regional development in many other ways, and research on other countries is informative about these questions. Sanchis-Guarner and Lyytikäinen (2011) estimate the effects of road improvements on the individual labor market and find strong positive impacts of the accessibility of the work location on basic hourly wages in Great Britain. Baum-Snow (2007) assesses the extent to which the construction of new limited access highways in the United States has contributed to the declining population in city centers. By employing micro firm-level data, Holl (2004a, 2004b) proves that new firms prefer locations closer to new motorways in Spain and Portugal. Rothenberg (2011) suggests that road improvements were accompanied by a significant dispersion of manufacturing activity in Indonesia. Highway and railroad construction between Tokyo and Osaka in the 1960s was one of the main factors promoting the rapid concentration of many industries and workers in the three largest metropolitan areas (Tokyo, Kyoto-Osaka-Kobe, and Nagoya) and shaping the Pacific industrial belt in Japan (Fujita *et al.*, 2004; World Bank, 2008).

In particular, Rephann and Isserman (1994) is the pioneer empirical work on highway issues in the US. They use the quasi-experimental matching method, which is similar to the strategy of this paper. By employing the Mahalanobis distance matching method, Rephann and Isserman (1994) examine the effects of interstate highways on counties that obtained links or were in close proximity to the new links and find that the areas that benefited most were those in close proximity to large cities or with some degree of prior urbanization. Furthermore, the population, tertiary industry, and the local government were affected most by highways, whereas manufacturing was not.

3 Empirical Strategy and Data

3.1 Empirical strategy

A highway route is not usually designed and constructed at random. Instead, it depends on social production, regional trade, national defense, political issues, existing principal truck roads, etc. (see e.g., Rietveld and Bruinsma, 1998; Holl, 2004a; Baum-Snow, 2007). Thus, most empirical research on highway construction faces the endogeneity problem. Regions with highway placement usually dominate other cities in many other aspects, so that even without highway placement, such regions would still have better economic performance because of advantages in political status, geographical location, population density, or similar factors. As a result, instruments such as historical highway planning, historical road lines, and geographical features (for example, elevation range, terrain ruggedness, groundwater, or heating and cooling degree days) are used as a source of exogenous variation to relieve the bias (see e.g., Baum-Snow, 2007; Baum-Snow *et al.*, 2012; Duranton and Turner, 2012; Faber, 2012).

A sensible instrument is often hard to find, and weak instruments cannot efficiently eliminate the bias. When the selection is due to observables, however, the methods available for controlling the selection bias are relatively straightforward (Zhao, 2004). Counterfactual analysis by matching, that is, selecting treated observations and comparison observations with similar covariates to correct the selection bias, may be efficient. Furthermore, compared to the structural approach, matching methods need not impose strong restrictions on a specific linear or non-linear relationship among the variables.

When focusing on units at the MSA (Metropolitan Statistical Area), prefecture, or higher administrative division level (for example, Baum-Snow, 2007; Duranton and Turner, 2012; Baum-Snow *et al.*, 2012), matching is still difficult to apply, because at these levels, highways may be an inherent infrastructure project. It is hard to find effective and large comparison samples for matching. However, at the micro level, it may be feasible to construct treated and comparison observations with similar values of the covariates. Motorway project placements can be assumed to be exogenous to changes at the municipality level since such decisions are mainly determined by factors and forces above the local level (Rietveld and Bruinsma, 1998). Sanchis-Guarner and Lyytikäinen (2011) point out that transport projects are aimed at a higher spatial scale to improve safety or reduce congestion within a wider area and are not aimed at specific individuals or wards (a ward is a micro administration unit, and there are 10,500 wards in Great Britain). Donaldson (2010) also gives similar evidence regarding India's railroads. To further check this assumption, Holl (2004a) employs a logistic regression model to assess whether the government's decisions of where to build new motorways were influenced by prior manufacturing plant locations at the municipality level. The results are insignificant. Thus, matching methods could be a reasonable approach for investigating highway effects in micro level units.

To do so, we apply propensity score matching (PSM). First, we need to select covariates that can affect both treatment participation and the potential outcome. Then, we run a probit or logit regression with participation as the dependent variable (1: treated; 0: untreated), and the predicted value from the probit or logit regression becomes the propensity score. Based on the given propensity score of each sample, PSM tries to match each treated observation with the most similar untreated observation (or some weighted observations).

The preferred estimate that receives the most attention in the PSM evaluation literature is the average treatment effect on the treated (ATT), which is defined as:

$$ATT = E[Y_1 - Y_0|D = 1] = E[Y_1|D = 1] - E[Y_0|D = 0] + E[Y_0|D = 0] - E[Y_0|D = 1],$$

where D is a dummy variable for participation, and Y_i represents the outcome indicator (1: treated; 0: untreated). PSM aims to ensure that the selection bias goes to 0 after conditioning on the propensity score, so the counterfactual part $E[Y_0|D=1]$ can be dropped:

$$E[Y_0|D=0,P(X)] - E[Y_0|D=1,P(X)] = 0,$$

where $P(X)$ is the Propensity Score with covariates X. Therefore, given that the required assumptions hold, the post-PSM ATT estimator can be written as:

$$ATT_{PSM} = E[Y_1|D=1,P(X)] - E[Y_0|D=0,P(X)],$$

which is the preferred form that can be directly estimated using the observable data.

However, when the sample size is too small, PSM does not perform well. On the other hand, Mahalanobis distance matching (MDM) (see e.g., Rubin, 1980) is relatively robust under different settings (Zhao, 2004). In the following matching procedure, we apply PSM as the main method to estimate highway effects on the regional growth of peripheral regions, and we use MDM as a supplement. Smith and Todd (2005) find that the double difference (DD) PSM estimate is more robust than traditional cross-section matching estimators, so we present the outcomes in the form of DD to relieve the bias due to unobservables. Further, we obtain the standard errors of the treatment effects by bootstrapping 200 times.

3.2 Data

If highway improvements affect economic performance, then there will be a difference between county-level cities with highway access (treated) and those with no highway access (controlled) in the long run. We choose the county-level city as the basic geographical unit to construct treatment and control observations with sufficient sample size. The existing highway network of China was mainly undertaken based on two national highway plans. The first was issued in 1992 and called *5-Zong-7-Heng* National Trunk Highway System, which includes five longitudinal roads and seven latitudinal roads. The planned length was 34,422 km, and 25,765 km of that is limited access highway. This plan aimed to connect the country's largest cities, main industrial zones, main transportation hubs and ports, all the cities with a population above one million, and most of the cities with a population above 500 thousand. This plan was completed by the end of 2007. The second extending plan, called the National Highway System Planned 7918 Highway Network, was issued in 2004 and aimed to shape the final national highway network with 7 capital lines, 9 longitudinal lines, and 18 latitudinal lines. The planned mileage of the main body was 85,000 km. Maps of these two highway plans are presented in Figures 2 and 3. The finished part of the National Highway System Planned 7918 Highway Network, together with provincial highways, connection lines, and ring lines of large cities, forms China's current highway network.

[Figure 2 and Figure 3]

County level socio-economic and highway data used in this analysis comes from several sources:

(1) We collected the highway toll station data from the homepage of the provincial public travel information service system, and we geo-coded[3] the location at the level of the county-level city. The highway is not free to access in China, so a toll station could be seen as an approximation of highway placement and highway access. This data set includes all the toll stations that can be geo-coded by Google Maps (updated to May 2012) in the designated provinces. In addition, a small number of ordinary roads are also subject to road tolls. They are not included in this data set. County-level cities with at least one highway toll are classified as highway cities.

(2) The earliest available micro level industry data set is the Chinese Industry Statistical Database (CISD) from 1998. In 1998, the highway mileage was 8,733 km, and it increased to 53,913 km by 2007. In the period from 1998 to 2007, China's highways experienced sharp development, and thus, we can suppose they had a significant impact on industry. In addition, China's market economy gained momentum in 1997, as the 15th Congress of the Communist Party of China officially endorsed an increase in the role of private firms in the economy (Song, Storesletten, and Zilibotti, 2011). Thus, 1998 is a suitable starting point for the analysis.

In the CISD (1998), there are detailed data from 165,119 firms, including all the state-owned industrial firms and other industrial firms with annual sales of more than 5 million Chinese Yuan. Of these 165,119 firms, 151,908 were established before 1996, when there was less than 4,000 km of highways. Thus, we can suppose that in 1998, firm location in China was not substantially affected by highways. In fact, before 1998, highways just connected some of the largest cities like Shenyang, Beijing, Shanghai, Guangzhou, and Wuhan, most of which are located in the coastal regions. For precision, we further exclude cities with highways placed before 1998 in the following analysis. CISD includes the ZIP code and administrative division code (*Xingzheng*

[3]Geo-coding is achieved through the Google Maps-based program: http://www.gpsspg.com/latitude-and-longitude.htm.

Quhua Daima) of each sample, so in this case, it is possible to locate a firm at the ZIP level. Furthermore, I employ the CISD (2007) to construct a double difference estimate. 2007 is the latest year in which the CISD is available with high quality data, and it is also the year when China finished its construction of the National Trunk Highway System.

There are three drawbacks of the CISD. First, it does not cover small, non-state-owned firms with annual sales less than 5 million Chinese Yuan, which will induce a bias since the distribution of such small firms is different across regions. Second, this data set locates a firm by its geographical location, but some firms may not belong to the local economy, and their output may contribute to other cities with higher hierarchy. This is particularly common for state-owned firms, which would cause a bias in statistical research. Third, this data set does not deflate the price index, which lowers the threshold to be included in the data in 2007 compared with 1998.

(3) Population data comes from the 5th national population census at the county level, which records the population, educational condition, and employment structure as of the end of 2000.

(4) County level GDP and per capita GDP come from the Provincial Statistical Yearbooks of the designated years.

The ZIP code and administrative division code are used to combine these four data sets. Because of the administrative division changes[4] in the past two decades, the ZIP code may fail to locate all the county-level units in some cases. We therefore use the administrative division code as a supplement. The administrative division code is another recording system in the national administration division that is similar to the ZIP code but is more regular and easier to track and employ for distinguishing and classifying. County-level units in all the data sets are corrected to match 1998 administrative divisions. To compare real growth in GDP/per capita GDP, industrial output (value-added), and investment (net value of fixed assets) in the industry, we use the provincial level consumer price index, the producer price index for manufactured goods, and the price index for the net value of fixed assets, respectively, to deflate the nominal value in 2007. The indices are taken from the price indices section of the annual China Statistical Yearbook. We assume comparability among provinces based on nominal prices in the base year, 1998.

[4]Some county-level cities were upgraded to prefectural-level units, placed under the jurisdiction of other prefectures, or classified as the core districts of prefectures.

In China, the county-level city is the fourth hierarchy of administration units following the central government, provincial unit, and prefecture-level unit (see Figure 4). Specifically, we choose the sample of county-level cities in four steps. First, we exclude 11 provinces due to data restriction. Second, the four provincial cities (Beijing, Tianjin, Shanghai, and Chongqing) and all the provincial capitals are excluded. These cities have a much higher highway density because of policy and economic superiority. Their social and economic characteristics are also different from other cities. On the other hand, county-level cities in these areas tend to be affected by the spillover effects of the large cities. Third, all of the core districts of relating prefectures are excluded. Finally, we exclude all the samples in prefectures that had highway placement before 1998. The related data comes from *Gonglu* (a journal on traffic study in Chinese; see 2002(4): 132-140), which summarizes the detailed building progress of China's limited access highways from 1988 to 2001.

To achieve the empirical procedures in the following sections, we divide the sample cities into three categories:

∗ *highway cities: received their first highway placement between 1998 and 2005;*

∗ *potential highway cities: received their first highway placement between 2006 and 2011;*

∗ *control cities: have no highway placement by the end of 2011.*

[Figure 4 and Figure 5]

Figure 5 shows that in the coastal regions (Eastern China), highways have a higher density than in the inland regions (Western and Central China); most of the sample cities in the coastal regions have highway placement. This is not strange based on China's economic geography. The coastal regions have higher populations and economic densities. Thus, we further divide the samples into coastal regions and inland regions and perform matching separately. Following the traditional classification of China's economic geography, we define six of the 16 sample provinces as coastal regions (Zhejiang, Jiangsu, Fujian, Guangdong, Liaoning and Hebei) and the remaining ten sample provinces as inland regions (Jilin, Anhui, Jiangxi, Hu'nan, Guangxi, Gansu, Shanxi, Sichuan, Guizhou, He'nan). The final prepared sample consists of 983 county-level cities, belonging to 16 provinces, with a total population of 460.7 million, which accounts for 36.4% of the Chinese population (5th national population census in 2000), and a total GDP of 1958.28 billion Chinese Yuan (24.7% of the national GDP in 1998).

4 Baseline Results, Placebo Check, and MDM Approach

4.1 Baseline results

In this section, we consider just highway cities and control cities. The post-treatment data set is from 2007, and the identification of highway cities is based on the status of highway placement as of the end of 2005, so we assume the period from 2005-2007 is the lag for industry to respond to the new highway placement. We choose covariates that may affect regional development and investment in infrastructure based on traditional literature, and we decide the final number and form (such as intersection, logarithm) of the covariates based on three criteria. First, we require sufficiently high predictive value from the propensity score to the actual placement in the samples (Heckman *et al.*, 1998). Second, the conditional independence assumption (CIA) based on the propensity score should be satisfied. Specifically, the post-matching samples should have a smaller pseudo R^2 than the pre-matching samples, so that the covariates are no longer able to provide new information about treatment participation conditional on the propensity score. Third, the difference in the covariates between the treatment and post-matching control samples should be insignificant. In addition, to apply PSM without bias, the covariates selected to use in matching should be either fixed over time or measured before participation in the treatment.

The details of the selected covariates are explained in Table A.1. Although highway planning is not subject to population and economic development at the county level, the population, economic, and geographical characteristics of the county-level cities are good predictors of highway placement. We find that highway cities are on average closer to large cities and main ports and higher in population density, firm density, and per capita GDP than cities with no highway placement. The explanation is that large cities tend to receive more highways, so the county-level cities close to them have a higher probability of being connected by highway compared to more remote county-level cities. In addition, these cities naturally perform better in terms of economic development because of their proximity to big markets and suppliers. Thus, we can assume that the covariates should have high correlations with both the participation indicator (highway placement) and regional development, which is the requirement for applying PSM.

We use the fraction of treated subjects (including both highway and potential highway cities) in the total number of observations as the cutoff value for the predicted probability. The cutoff values for our sample are 0.6055 for coastal cities and 0.3907 for inland cities. Using these cutoff values, the prediction rates for the treatment and control observations are 65.2% and 60.4% in coastal city groups and 90.9% and 52.0% in inland city groups, respectively. Using the same propensity score calculation, we find in Figure A.1 that both highway cities and potential highway cities can be well distinguished from control cities. The mean of the related probability for the treatment (propensity score) is greater in treated cities than in control cities. Thus, the propensity score is a suitable predictor of actual treatment participation. Then, the matched pseudo R^2 is given for each PSM treatment effect estimate in each column of Tables 1, 2, 4, 5, and 6. We can see a satisfactory reduction in the pseudo R^2 between the raw and matched samples. Furthermore, the covariates bias (%) is reduced, and there is no significant difference in most covariates between the treatment and post-matching control groups even at the 10% confidence level.[5]

The highway data set does not include the exact foundation year of each toll station, so it is impossible to precisely estimate the average yearly highway effect on industry. As a compromise, we compare the difference in the total growth from 1998 to 2007 between treatment and control cities using the DD-PSM method. DD-PSM is applied between the highway cities and the control cities. We present the estimation results with the kernel and radius matching algorithms, which are the most common matching algorithms. Furthermore, matching is achieved with replacement, so the absence of potential highway cities, which will be used in the placebo check (see section 4.2), does not affect the matching results in this step. Columns 1-4 of Table 1 give the estimates of the effects of highway placement on per capita GDP and industrial growth (firm density, investment, and output). In column 1, the indicators of output and investment in industry show significant treatment effects relative to their pre-treatment values. For example, in 1998 (before treatment) the average industrial investment in the treated cities is 101.79% of that in control cities, and the difference is not significant, but the DD results increase to 176.01% and are significant at the 1% confidence levels. Columns 2-3 present similar results, and the only exception is in column 4. However, the treatment effects on per capita GDP and firm density are insignificant. We attempt to explain this result further in the following sections.

[Table 1]

[5] Detailed information is presented in Table A.2.

4.2 Placebo check

The treatment effects of the baseline results for industrial indicators could be caused by the disturbance of unobservable variables. One possible test is to check for a *pseudo treatment effect* using the same propensity score calculation and matching strategy but with potential highway cities (treated after 2005) instead of highway cities (treated before 2005).

China continued its highway construction in the Eleventh Five-Year (2006-2010) period. Its highway mileage was extended to more than 70,000 km in 2010 and reached 80,000 km by the end of 2011. In this period, more than 30,000 km was newly constructed all over the country, and many county-level cities received highway placements for the first time. In this respect, if the improved performance in the highway cities presented in Table 1 is indeed attributable to highway placement, we can expect to find no significant difference between potential highway cities and control cities, because in the estimated period, potential highway cities have no actual highway placement. If this is true, we can also prove that economic performance in the county-level cities does not dominate the highway placement. To maintain comparability, we use the same matching strategies as in Section 4.1.

In Figure A.1 we can see that the propensity score also distinguishes well between potential highway cities and control cities; the average propensity score of potential highway cities is larger than that of control cities. The pseudo treatment effect estimates are consistent with the above assumption. Compared to the results in Table 1, the average difference between the industrial indicators in potential highway cities and control cities is much smaller. Almost all of the pre-treatment differences and double differences of the industrial indicators are insignificant (with only two exceptions in firm density, but they have negative values that are compatible with our opinion). To be clear, the trend in industrial output (value added of industry) during 1998 to 2007 is presented in Figure 6. Highway cities tend to have faster growth in industry output than control cities, but the potential highway cities do show a similar trend with control cities. For the coastal potential highway cities, the growth in output is even slower than control cities. The figures are based on kernel matching estimates, but under radius matching, the result is also highly consistent. Similarly, in Table 2, we do not find a significant difference in per capita GDP growth, and the double difference result is similar to the pre-treatment difference.

[Table 2 and Figure 6]

4.3 MDM approach

Another bias in the treatment effects estimation may come from the matching strategy. Zhao (2004) proves that PSM (propensity score matching) does not perform well when the sample size is too small (n<500) and that MDM (Mahalanobis distance matching) is relatively robust under different settings. MDM finds the pairs with the smallest Mahalanobis distance. The Mahalanobis distance between two points $x = (x_1, ..., x_n)^T$ and $y = (y_1, ..., y_n)^T$ in the n-dimensional space $\mathbb{R}^n$ is defined as:

$$D_M(xy) = \sqrt{(x-y)^T S^{-1} (x-y)}.$$

If the covariance matrix S is replaced by the identity matrix I_n, then the Mahalanobis distance reduces to the general Euclidean distance.

To better check the stability of the PSM results, we apply MDM with the same covariates as in the PSM. The results presented in Table 3 are highly consistent and stable. For both coastal and inland cities, the pre-treatment differences between highway cities and control cities are not large and are insignificant. However, after highway placement, industrial investment and output have significant and more considerable growth compared to the control cities. For example, the average pre-treatment difference in output is 13.11% (coastal regions) and 3.86% (inland regions), and neither of these differences is significant, but in the period from 1998-2007, the difference in growth of these two indicators increased to 67.72% and 213.92%, respectively. On the other hand, columns 3 and 4 of Table 3 give the estimated results for potential highway cities. We find similar results to Table 2. That is, potential highway cities do not have significantly better performance in industry compared to the control cities.

[Table 3]

5 Further Estimates and Discussions

5.1 Highway effects in different geographical locations

The effects of highways on regional development and industrial performance rely on access to consumer and supplier markets. Xu *et al.* (2010) point out that scale effects and other external economies related to spatial agglomeration drive large cities to absorb economic resources from their surroundings, which is the significant centripetal force. Thus, the closer to the central cities, the faster a city grows. They estimate that the turning point for Chinese cities is 300 km. When the distance exceeds 300 km, the centrifugal force instead plays a major role. Rephann and Isserman (1994) also find that the beneficiaries of new interstate highways in terms of economic growth are interstate counties in close proximity to large cities. Highway placement plays a role in drawing connected cities closer because the average transport efficiency on the highway is much higher than on ordinary roads. Thus, we can expect that highways can help make the corresponding cities closer, at least in the economic sense, to large cities. It is therefore persuasive to check whether highway effects differ between county-level cities with different distances to large cities.

From Table A.1, we can see that in the coastal regions, the average minimum distance from highway county-level cities to large cities is less than 200 km, and most of them are close to large cities. The average minimum distance of inland cities is 312 km, which is similar to the distance (300 km) mentioned by Xu *et al.* (2010), so it is suitable to use only inland samples to check the highway effects in different geographical locations. We divide the cities into two parts by geographical location: close or remote to large cities. The cutoff is set at 300 km, and the control cities in the inland regions are also divided into two parts by this criterion. The related propensity score distribution shows that the propensity score can still predict highway placement (Figure A.2).

The DD-PSM results are presented in Table 4. Evidently, cities close to large cities are more active in industrial activities. Columns 2 and 4 show that most industrial indicators are not significant before treatment but become significant after treatment. To correct the large pre-treatment bias in industrial indicators in column 1, we insert two outcome variables, industrial output and investment, into the set of matching covariates and present an alternative matching estimate in column 5. We also present a similar estimate related to column 2 in column 6. Then, the results are consistent. The highway treatment effect is not evident in county-level cities remote to large cities, but it is highly positive and significant in cities close to large cities.

[Table 4]

The centripetal and centrifugal forces mentioned by Xu *et al.*(2010) could form a core explanation of the results in Table 4. On the other hand, they could follow another logic. Highway construction always starts in large cities and then gradually extends to the peripheral regions.

That is, county-level cities close to large cities tend to receive highway placement earlier. Thus, in highway cities remote to large cities, industrial performance needs more time to respond to the highway placement. This explanation does not weaken the robustness of the estimated results for the *Potential Highway Cities* (Table 2), because the average minimum distance to large cities among the inland potential highway cities is 306 km (see column 6 of Table A.1), which is even smaller than that of the highway cities (312 km). These *insignificant* estimates are due to the absence of highway placement before 2005, not their remoteness to large cities.

5.2 Highway effects on the structure of industry

All of the estimates in Sections 4.1, 4.2, 4.3, and 5.1 do not show significant effects of highway placement on per capita GDP. As a further check, we pool all the possible outcomes and estimate the highway treatment effects. Table 5 shows that there is a significant differentiation of highway effects even in the internal structure of industry. That is, heavy industry, not light industry, is heavily affected by highway placement.

[Table 5]

This result is consistent with China's process of industrialization. Although light industry in China also developed rapidly in the period 1998-2007, heavy industry performed as the pillar of the economy. In 1998, heavy industry was 49% of all firms and 55% of the value added to national industry, but in 2007, these shares were raised to 57% and 68%, respectively (China Statistical Yearbook 1999 & 2008). This suggests that heavy industry has been more active than light industry recently in China and that the average firm size of heavy industry is larger than that of light industry, which could partly explain why the estimated growth of firm density in the highway cities is not as significant as the growth in investment and output. It is also straightforward that heavy industry is much more dependent than light industry on transport and will thus benefit more from highway placement.

The indicator of labor productivity in industry is insignificant, but it is significant in terms of employment (Table 6). Although in some estimates the treated cities have a greater increase in firm density, neither increase is significant. Thus, the significance of treatment effects in employment could potentially explain why the highway cities have more employment and a larger average firm size, which leads to higher industrial output and investments. In addition, this could be a potential explanation for the insignificant treatment effects on per capita GDP. That is, highway placement may promote the upgrading of the industrial structure from a primary sector driven economy to a secondary sector driven economy or from a light industry driven economy to a heavy industry driven economy, but it is helpless to improve social productivity. On the other hand, another two factors could help the interpretation. First, the CISD records firm level data according to the firm's geographical location, but in some cases, a firm's output may not be attributed to the local economy but instead to a higher administration unit. This is particularly common for state-owned firms. Thus, the statistical rules between the two data sets (industrial indicator and GDP indicator) are inconsistent. Second, China's regional convergence in economic development is observed throughout the literature. Taking the growth (1998-2007) of per capita GDP as the dependent variable and per capita GDP in 1998 as an independent variable, an OLS estimate using the samples in this paper gives a significant negative coefficient under different settings.

[**Table 6**]

5.3 Highway effects or railway effects?

In addition, railways could be a potential disturbance of highway effects on industry performance since they are also an important transport infrastructure. The following facts support our results. First, freight stations in the railway network did not increase in the past decade. Most of them were built before the 2000s and even before the 1960s. The limited new railway lines and stations are mostly for passengers, such as the China Railway High-speed (CRH). Second, coal, coke, petroleum, steel and iron, metal ores and nonmetal ores, cement, timber, grain, cotton, and salt account for 83.2% of the total railway freight ton-kilometers and 91.2% of the total railway freight traffic in 2007. That is, raw materials, not industrial products, are mainly carried by railway, so we can expect that the road network plays a role in the shipping of industrial products and further affects industrial firm location. Third, more and more manufacturing firms are located along highways all over the country.

6 Conclusion

By employing a micro level data set, we estimate the highway effects on regional growth in the peripheral regions of China. We find that highway placement promotes the industrial development of related cities with higher output and more industrial investment. However, highways do not help remote peripheral regions, which in this case are county-level cities more than 300 km from large cities. Furthermore, in the county-level cities, highways only promote the development of heavy industry and do not promote light industry. Moreover, highway placement does not improve industrial productivity or per capita GDP. We use another matching method and robustness checks to ensure and improve precision, and we provide some potential explanations for the insignificant highway effects on industrial productivity and per capita GDP. In the future, an interesting extension of this research may be to examine highway effects in conjunction with urbanization.

References

Banister, D., J. Berechman (2000) *Transport Investment and Economic Development.* UCL Press, London.

Baum-Snow, N. (2007) Did Highways Cause Suburbanization?, *Quarterly Journal of Economics* 122, 775-805.

Baum-Snow, N., L. Brandt, J.V. Henderson, M.A. Turner, Q. Zhang (2012) Roads, Railroads and Decentralization of Chinese Cities, *mimeo*, Brown University.

Baum-Snow, N., M.A. Turner (2012) Roads, Transportation and the Decentralization of Chinese Cities, *mimeo*, Brown University.

Behrens, K., C. Gaigné, J.-F. Thisse (2009) Industry Location and Welfare When Transport Costs Are Endogenous, *Journal of Urban Economics* 65(2), 195-208.

Donaldson, D. (2010) Railroads of the Raj: Estimation the Impact of Transportation Infrastructure, *mimeo*, Massachusetts Institute of Technology.

Duranton, G., M.A. Turner (2012) Urban Growth and Transportation, *Review of Economic Studies* 79, 1407-1440.

Faber, B. (2012) Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System, *mimeo*, London School of Economics and Political Science.

Fujita, M., P. Krugman, A.J. Venables (1999) *The Spatial Economy: Cities, Regions and International Trade.* MIT Press, Cambridge.

Fujita, M., T. Mori, J.V. Henderson, Y. Kanemoto (2004) Spatial Distribution of Economic Activities in Japan and China, in: J.V. Henderson, J.-F. Thisse (eds.) *Handbook of Urban and Regional Economics* 4(65) 2911-2977, North-Holland, Amsterdam.

Heckman, J., H. Ichimura, P. Todd (1998) Characterizing Selection Bias Using Experimental Data, *Econometrica* 66, 1017-1098.

Holl, A. (2004a) Manufacturing Location and Impacts of Road Transport Infrastructure: Empirical Evidence from Spain, *Regional Science and Urban Economics* 34, 341-363.

——. (2004b) Transport Infrastructure, Agglomeration Economies, and Firm Birth: Empirical Evidence from Portugal, *Journal of Regional Science* 44, 693-712.

Krugman, P.R. (1991) Increasing Returns and Economic Geography, *Journal of Political Economy* 99, 483-499.

Puga, D. (1999) The Rise and Fall of Regional Inequalities, *European Economic Review* 43, 303-334.

Rephann, T., A. Isserman (1994) New Highway as Economic Development Tools: An Evaluation Using Quasi-Experimental Matching Methods, *Regional Science and Urban Economics* 24, 723-751.

Rietveld, P., F. Bruinsma (1998) *Is Transport Infrastructure Effective? Transport Infrastructure and Accessibility: Impacts on the Space Economy.* Springer, Berlin.

Rothenberg, A.D. (2011) Transport Infrastructure and Firm Location Choice in Equilibrium: Evidence from Indonesia's Highways, *mimeo*, University of California, Berkeley.

Rubin, D.B. (1980) Bias Reduction Using Mahalanobis-Metric Matching, *Biometrics* 36, 293-298.

Sanchis-Guarner, R., T. Lyytikäinen (2011) Driving Up Wages: The Effects of Road Improvements in Great Britain, *mimeo*, London School of Economics and Political Science.

Smith, J.A., P.E. Todd (2005) Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?, *Journal of Econometrics* 125, 305-353.

Song, Z., K. Storesletten, F. Zilibotti (2011) Growing Like China, *American Economic Review* 101, 202-241.

Venables, A. (1996) Equilibrium Locations of Vertically Linked Industries, *International Economic Review* 37, 341-359.

World Bank (2008) *World Development Report 2009—Reshaping Economic Geography*. World Bank, Washington, DC.

Xu, Z., Z. Chen, M. Lu (2010) Core-Periphery Model of Urban Economic Growth: Empirical Evidence from Chinese City-Level Data (1990-2006) (in Chinese), *The Journal of World Economy* 7, 144-160.

Zhao, Z. (2004) Using Matching to Estimate Treatment Effects: Data Requirements, Matching Metrics, and Monte Carlo Evidence, *Review of Economics and Statistics* 86, 91-107.

Table 1: Estimation of Highway Treatment Effects

	(1)	(2)	(3)	(4)
Double difference				
Per capita GDP	0.0829	-0.0065	0.2198	-0.0036
	(0.75)	(-0.09)	(0.94)	(-0.03)
Firm density	0.0987	-0.1174	0.4197	0.0434
	(0.30)	(-0.70)	(1.37)	(0.14)
Investment	0.7601***	0.7230**	1.6474***	0.5289
	(2.90)	(2.36)	(2.83)	(1.38)
Output	0.3958**	0.2818**	0.8894***	0.2536
	(2.13)	(2.09)	(2.61)	(1.60)
Pre-treatment				
Per capita GDP	0.0947	0.0365	0.1862	0.0090
	(1.57)	(1.01)	(1.43)	(0.15)
Firm density	0.1456	0.2169*	0.0652	-0.0285
	(1.06)	(1.80)	(0.29)	(-0.18)
Investment	0.0179	0.0523	0.3633	0.3469
	(0.09)	(0.26)	(1.36)	(0.96)
Output	0.1867	0.1259	0.3702	0.3496
	(1.28)	(0.84)	(1.38)	(1.25)
Description				
Region	Coastal	Inland	Coastal	Inland
Matching method	PSM,kernel	PSM,kernel	PSM,radius	PSM,radius
Highway placement	Yes	Yes	Yes	Yes
No. matched obs. (Treated/Control)	150/101	242/379	64/101	107/379
Pseudo R^2 (Raw/Matched)	0.062/0.009	0.187/0.019	0.062/0.040	0.187/0.027

Note: Definition of the estimated ATT coefficient: Coefficient=(Treated-Control)/Control. Samples that cannot satisfy the *common support* are dropped. Industrial investment (net value of fixed assets) and output (value-added) indicate the total amount of a city. These also apply in the following tables.
We choose the caliper based on the matched pseudo R^2 and a t-test of bias in covariates.
Z values in parentheses: * Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level.
Highway placement: "Yes" means highway cities.

Table 2: Pseudo Highway Treatment Effects Estimation in Potential Highway Cities

	(1)	(2)	(3)	(4)
Double difference				
Per capita GDP	-0.1980	0.0337	-0.0901	-0.0093
	(-1.62)	(0.32)	(-0.26)	(-0.09)
Firm density	-0.1445	-0.4250***	0.1870	-0.3156
	(-0.59)	(-2.89)	(0.43)	(-1.29)
Investment	-0.2747	1.0301	-0.3412	0.5572
	(-1.26)	(1.53)	(-0.74)	(0.87)
Output	-0.1663	0.1055	0.1652	-0.0736
	(-0.63)	(0.67)	(0.46)	(-0.39)
Pre-treatment				
Per capita GDP	-0.1475*	0.0037	-0.1535	0.0061
	(-1.73)	(0.09)	(-0.78)	(0.10)
Firm density	-0.3899**	0.1572	-0.2415	0.1369
	(-2.18)	(1.21)	(-0.56)	(0.75)
Investment	-0.2397	0.2741	-0.2467	-0.0919
	(-0.79)	(0.70)	(-0.67)	(-0.27)
Output	-0.0392	0.1060	0.1096	0.0529
	(-0.15)	(0.53)	(0.27)	(0.22)
Description				
Region	Coastal	Inland	Coastal	Inland
Matching method	PSM,kernel	PSM,kernel	PSM,radius	PSM,radius
Highway placement	Potential	Potential	Potential	Potential
No. matched obs. (Treated/Control)	20/101	85/379	13/101	82/379
Pseudo R^2 (Raw/Matched)	0.212/0.190	0.033/0.011	0.212/0.206	0.033/0.017

Note: Z values in parentheses: * Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level.
Highway placement: "Potential" means potential highway cities.

Table 3: Estimation of Mahalanobis Distance Matching

	(1)	(2)	(3)	(4)
Double difference				
Per capita GDP	0.1509	0.0910	-0.0551	0.0169
	(1.38)	(0.86)	(-0.19)	(0.16)
Firm density	0.3239	0.1850	0.0664	-0.1043
	(1.46)	(0.77)	(0.27)	(-0.43)
Investment	0.6772**	2.1392**	-0.2266	0.5613
	(2.18)	(2.03)	(-0.54)	(0.82)
Output	0.4426**	0.5490**	-0.1287	-0.0498
	(2.55)	(2.42)	(-0.49)	(-0.28)
Pre-treatment				
Per capita GDP	0.0439	0.0316*	0.0235	0.0710**
	(1.19)	(1.84)	(0.30)	(2.20)
Firm density	0.0704	0.0685	-0.1784	0.0818
	(0.70)	(1.05)	(-0.93)	(1.44)
Investment	0.1311	0.0386	-0.0469	0.3681
	(0.93)	(0.09)	(-0.13)	(0.82)
Output	-0.0129	-0.0275	0.0271	0.2090
	(-0.09)	(-0.11)	(0.07)	(0.87)
Description				
Region	Coastal	Inland	Coastal	Inland
Matching method	MDM,radius	MDM,radius	MDM,radius	MDM,radius
Highway placement	Yes	Yes	Potential	Potential
No. matched obs. (Treated/Control)	69/101	115/379	12/101	75/379

Note: For comparison with the PSM results, in the MDM we use the same covariates as in PSM and choose a matched sample size corresponding to the PSM with radius in Tables 1 and 2.
Z values in parentheses: * Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level.

Table 4: Estimation of Highway Treatment Effects in Different Geographical Locations

	(1)	(2)	(3)	(4)	(5)	(6)
Double difference						
Per capita GDP	-0.0446	-0.0207	0.0164	0.0856	0.0084	-0.0139
	(-0.34)	(-0.24)	(0.11)	(0.64)	(0.08)	(-0.16)
Firm density	-0.3721	0.1525	0.0129	-0.1061	-0.4072	0.0192
	(-1.44)	(0.86)	(0.03)	(-0.47)	(-1.43)	(0.10)
Investment	0.2788	1.1681**	0.5345	0.8395*	0.3226	0.7297***
	(0.83)	(2.42)	(0.96)	(1.72)	(1.12)	(2.61)
Output	0.1697	0.3183*	0.3394	0.1548	0.1473	0.3223*
	(0.75)	(1.81)	(1.14)	(0.59)	(0.77)	(1.83)
Pre-treatment						
Per capita GDP	0.0096	0.0017	0.0064	0.0544	0.0276	-0.0016
	(0.19)	(0.03)	(0.08)	(0.58)	(0.54)	(-0.02)
Firm density	0.5345***	-0.0827	0.4029*	-0.1124	0.1831	-0.0007
	(2.85)	(-0.44)	(1.84)	(-0.54)	(1.08)	(-0.00)
Investment	0.5922**	-0.1385	-0.0416	0.1218	0.0958	0.1615
	(2.34)	(-0.49)	(-0.09)	(0.46)	(0.43)	(0.74)
Output	0.6451***	-0.0419	0.2831	-0.0727	0.2352*	0.1218
	(2.95)	(-0.18)	(1.00)	(-0.22)	(1.71)	(0.46)
Description						
Region	Inland	Inland	Inland	Inland	Inland	Inland
Distance	Remote	Close	Remote	Close	Remote	Close
Matching method	PSM,kernel	PSM,kernel	PSM,radius	PSM,radius	PSM,kernel	PSM,kernel
Highway placement	Yes	Yes	Yes	Yes	Yes	Yes
No. matched obs.						
(Treated/Control)	139/208	98/171	68/208	52/171	133/208	95/171
Pseudo R^2						
(Raw/Matched)	0.212/0.033	0.188/0.020	0.212/0.090	0.188/0.043	0.223/0.025	0.192/0.010

Note: Columns 5 and 6 are estimated with matching covariates including the nine variables presented in Table A.1 and another two covariates, industrial output and investment, to eliminate the significant pre-treatment difference in column 1.

Z values in parentheses: * Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level.

Distance: “Close” means cities have a minimum distance to large cities less than 300 km, “Remote” means cities have a minimum distance to large cities more than 300 km.

Table 5: Estimation of Highway Treatment Effects with Different Industries

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Double difference								
Firm density	0.4986	0.1276	0.4354	-0.0246	-0.0763	0.1647	0.4114	-0.1486
	(1.61)	(0.49)	(0.70)	(-0.06)	(-0.18)	(0.91)	(0.90)	(-0.61)
Investment	1.3135***	0.1561	0.7455	-0.0202	0.6384**	1.6352**	1.9460***	1.3201*
	(2.79)	(0.56)	(1.51)	(-0.05)	(2.10)	(2.36)	(2.59)	(1.73)
Output	0.3176	-0.0265	0.2715	-0.1613	0.4379**	0.5383**	1.2611***	0.4068
	(1.15)	(-0.12)	(0.43)	(-0.44)	(2.18)	(2.30)	(3.09)	(1.13)
Pre-treatment								
Firm density	0.1259	-0.1569	-0.0348	-0.1727	0.1701	0.0003	0.1882	-0.0438
	(0.70)	(-0.74)	(-0.11)	(-0.64)	(1.32)	(0.00)	(0.95)	(-0.21)
Investment	-0.0505	-0.0066	0.8700**	-0.0182	0.0714	-0.1873	0.1533	0.1983
	(-0.15)	(-0.03)	(2.75)	(-0.06)	(0.37)	(-0.52)	(0.43)	(0.59)
Output	0.2415	-0.1086	0.4179	-0.3615	0.1349	0.0121	0.3314	0.2631
	(1.32)	(-0.43)	(1.28)	(-0.97)	(0.75)	(0.04)	(0.86)	(0.52)
Description								
Region	Coastal	Inland	Coastal	Inland	Coastal	Inland	Coastal	Inland
Distance	All	Close	All	Close	All	Close	All	Close
Matching method	PSM,	PSM,	PSM,	PSM,	PSM,	PSM,	PSM,	PSM,
	kernel	kernel	radius	radius	kernel	kernel	radius	radius
Highway placement	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry	Light	Light	Light	Light	Heavy	Heavy	Heavy	Heavy
No. matched obs.								
(Treated/Control)	150/101	98/171	64/101	52/171	150/101	98/171	64/101	52/171
Pseudo R^2								
(Raw/Matched)	0.062/0.009	0.188/0.020	0.062/0.040	0.188/0.043	0.062/0.009	0.188/0.020	0.062/0.040	0.188/0.043

Note: Based on Table 4, we only include samples with a minimum distance to large cities less than 300 km in columns 2, 4, 6 and 8. To maintain comparability, we choose a caliper in the radius matching consistent with column 3 of Table 1 and column 4 of Table 4. Z values in parentheses: * Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level. Distance: "Close" means cities have a minimum distance to large cities less than 300 km, "All" means all of the related cities.

Table 6: Industry Structure, Labor Productivity, and Employment

	(1)	(2)	(3)	(4)
Post-treatment				
Labor productivity	0.0323	-0.0059	0.1907	0.0476
in industry	(0.47)	(-0.07)	(1.45)	(0.22)
Employment in	0.4014***	0.2397**	0.9710***	0.1438
industry	(2.69)	(2.11)	(3.59)	(0.88)
Double difference				
Share of	0.1247	-0.0597	0.2015	-0.0508
primary sector	(-0.86)	(0.42)	(-0.87)	(0.29)
Share of	0.3389	-0.2094	0.3163	-0.1028
secondary sector	(0.90)	(-1.11)	(0.64)	(-0.34)
Share of	-0.1410	0.5643	0.0700	0.0660
tertiary sector	(-0.45)	(0.99)	(0.13)	(0.13)
Pre-treatment				
Share of	-0.0155	-0.0104	-0.0521	-0.0129
primary sector	(-0.27)	(-0.23)	(-0.54)	(-0.21)
Share of	0.0104	0.0306	0.0375	0.0214
secondary sector	(0.21)	(0.67)	(0.46)	(0.33)
Share of	-0.0010	-0.0264	0.0024	-0.0072
tertiary sector	(-0.03)	(-0.71)	(0.05)	(-0.11)
Description				
Region	Coastal	Inland	Coastal	Inland
Distance	All	Close	All	Close
Matching method	PSM,kernel	PSM,kernel	PSM,radius	PSM,radius
Highway placement	Yes	Yes	Yes	Yes
No. matched obs.				
(Treated/Control)	150/99	98/170	62/99	52/170
Pseudo R^2				
(Raw/Matched)	0.069/0.010	0.182/0.020	0.069/0.040	0.182/0.043

Note: Due to data restrictions, we do not present the pre-treatment indicator of labor productivity and employment in the industry (these two indicators are only available for firms in the CISD).
To maintain comparability, we choose a caliper in the radius matching equivalent to column 3 of Table 1 and column 4 of Table 4. The difference in sample size is due to the missing data.
The share of primary sector is reduced by industrialization, so there is a converse plus-or-minus sign between the coefficient and the Z value in the double difference of "Share of primary sector".
Z values in parentheses: * Significant at the 10 percent level, ** Significant at the 5 percent level, *** Significant at the 1 percent level.
The description of the covariates is as follows:
Labor productivity: per capita value-added (value-added divided by number of employees in the firms above the designated size);
Employment: average number of employees in the industrial firms above the designated size;
Share of primary sector: GDP share of primary sector;
Share of secondary sector: GDP share of secondary sector;
Share of tertiary sector: GDP share of tertiary sector.

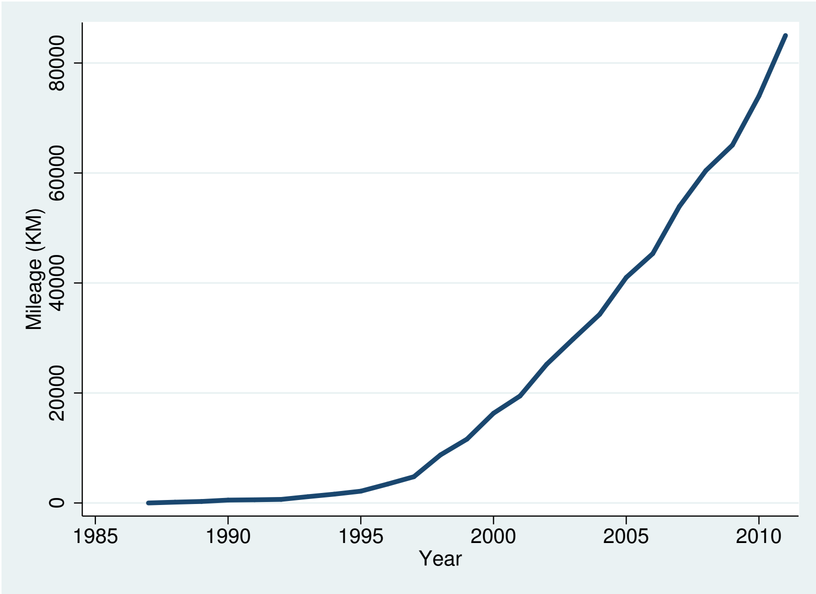

Figure 1: Mileage of China's Limited Access Highways

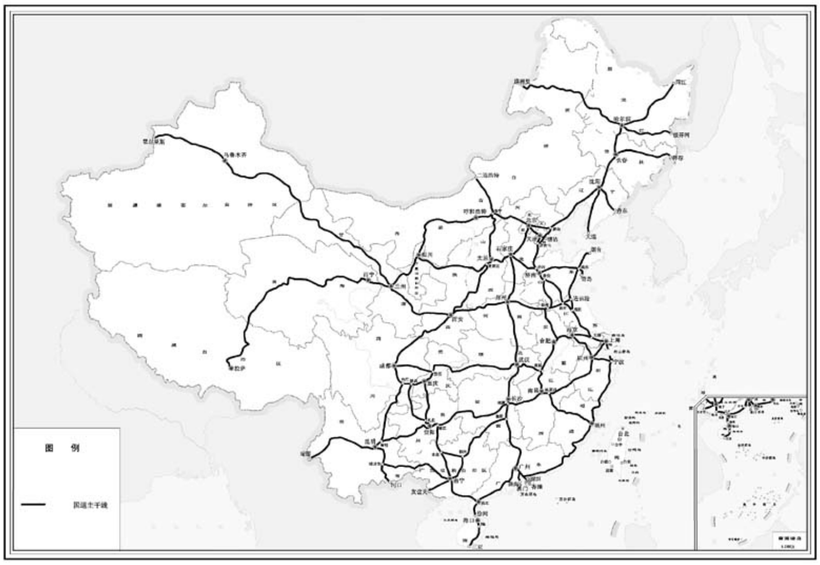

Figure 2: National Trunk Highway System

Figure 3: National Highway System Planned 7918 Highway Network

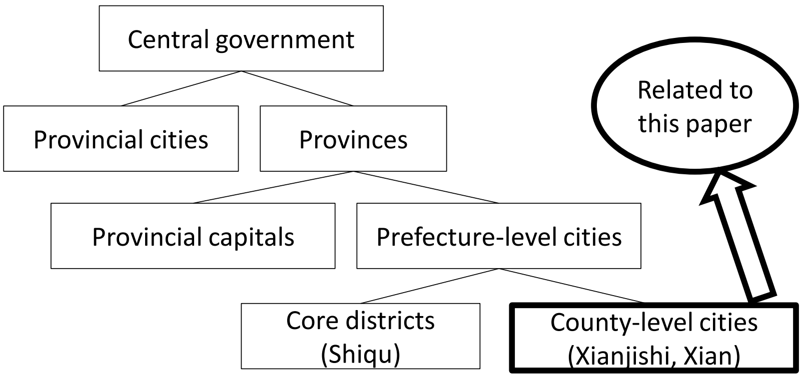

Figure 4: Geographic Units of Mainland China Related to This Paper

Note: For the five autonomous regions Xinjiang, Tibet, Ningxia, Guangxi and Inner Mongolia, they are classified as equivalent to provinces.

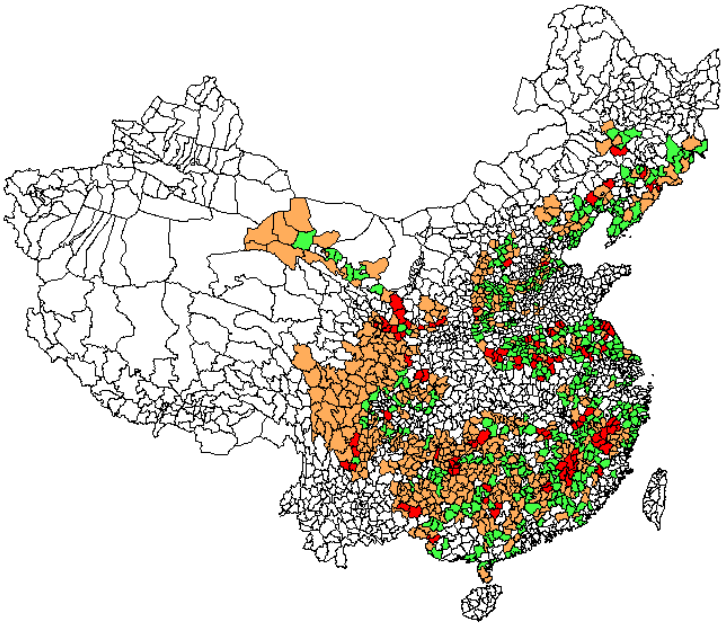

Figure 5: Selected City Samples for the Matching Procedure
Note: Green units: highway cities; red units: potential highway cities; orange units: control cities. Distinguishing between highway cities and potential highway cities is achieved using the digitalized map of China's highway lines in 2005 from Baum-Snow and Turner (2012) and the Geo-referencing module of ArcGIS.

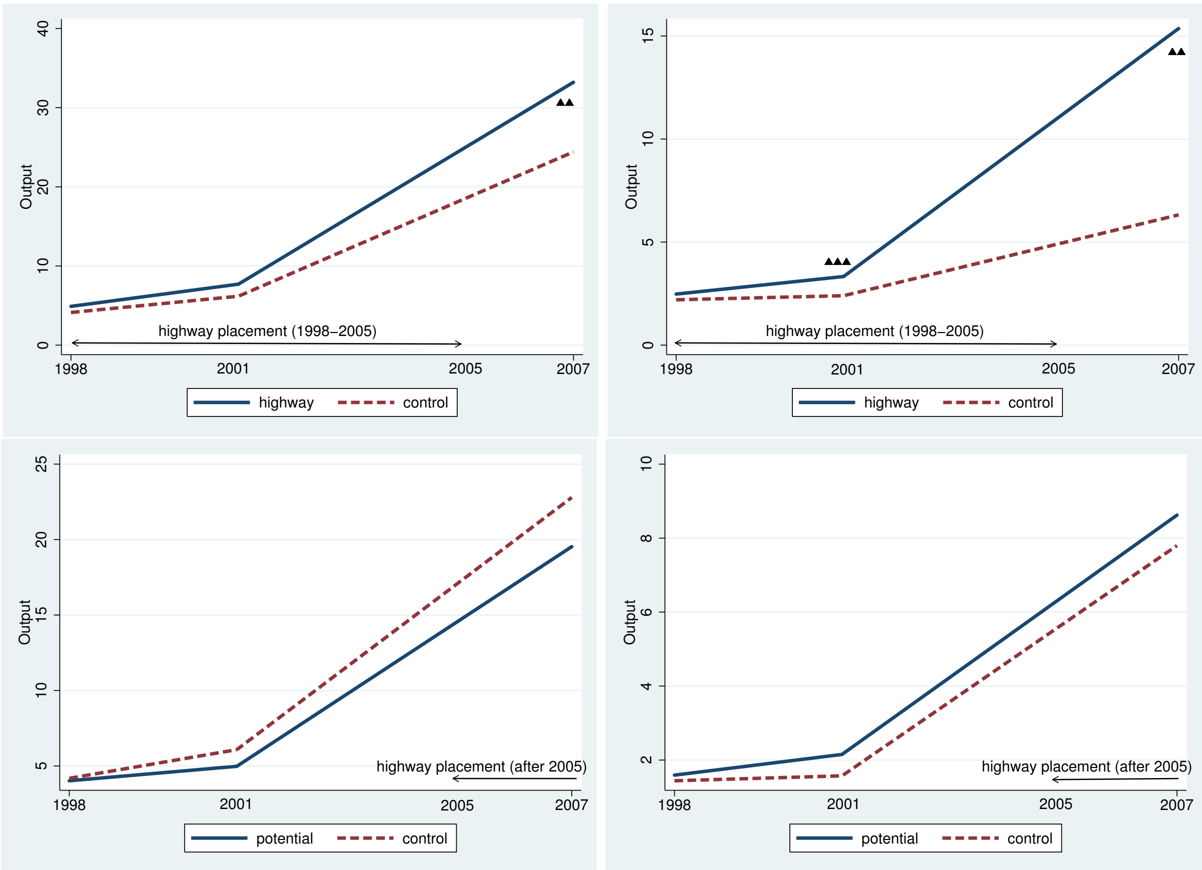

Figure 6: Output (Value Added) of Industry in 1998, 2001, 2007

Note: Top left is coastal highway cities; top right is inland highway cities; bottom left is coastal potential highway cities; and bottom right is inland potential highway cities. One, two and three triangles: difference in value added between treatment and control groups is significant at the 10, 5, and 1 percent level, respectively. The monetary unit in the figures is one billion yuan after deflating the price index to 1998.

[APPENDIX]

Table A.1: Mean of Covariates Used in the Matching Procedure

Variate	(1)	(2)	(3)	(4)	(5)	(6)
area	1766	1988	2393	2057	3286	2287
mindistto19	170.48	185.49	220.73	312.21	335.56	306.12
mindistto4	270.59	278.09	389.62	645.76	813.49	712.83
share2	0.1710	0.1372	0.1200	0.0948	0.0589	0.0649
share3	0.1593	0.1661	0.1627	0.1340	0.1176	0.1129
firmdensity	0.0682	0.0401	0.0292	0.0291	0.0120	0.0173
popudensity	446.7390	321.3732	297.5525	343.6198	205.1906	258.2149
educ	0.1117	0.1112	0.1239	0.1094	0.0929	0.0937
lnpergdp98	8.6785	8.5594	8.4442	8.1967	7.8208	7.9013
Region	Coastal	Coastal	Coastal	Inland	Inland	Inland
Highway placement	Yes	No	Potential	Yes	No	Potential
No. obs.	155	101	20	243	379	85

Note: Because of data restrictions, population, employment, and education data come from the 5th national population census in 2000. In PSM, it is required to measure the covariates before participation or fix them over time, so we assume these indicators are relatively fixed from 1998 to 2000. Because of China's "*Hukou*" system, population mobility in peripheral regions is fairly limited.
Description of the covariates is the follows:
area: land area (km^2);
mindistto19: The minimum spherical distance to 3 provincial cities (Shanghai, Beijing, and Tianjin) and 16 sub-provincial cities (Harbin, Changchun, Shenyang, Dalian, Qingdao, Nanjing, Wuhan, Hangzhou, Guangzhou, Shenzhen, Chengdu, Chongqing, Ji'nan, Ningbo, Xiamen, Xi'an) in 1998;
mindistto4: The minimum spherical distance to Dalian, Tianjin, Shanghai, and Guangzhou. These four cities are the largest port cities of the Northeast, North, East, and South regions of mainland China (Dongbei, Huabei, Huadong, Hua'nan);
share2: employment share of the secondary sector in total employment;
share3: employment share of the tertiary sector in total employment;
firmdensity: No. of industrial firms above the designated size per km^2;
popudensity: Resident population per km^2;
educ: Share of population with a high school or higher education;
lnpergdp98: log of per capita GDP (Yuan) in 1998.

Table A.2: Post-matching Differences in Covariates (Bias%)

Covariate	1.1,5.1,5.5	1.2	1.3,5.3,5.7	1.4	2.1	2.2	2.3	2.4	3.1	3.2	3.3
area	-12.9	5.9*	-15.5	10.0*	29.2	-4.5	-9.4	-1.5	1.3	1.9	24.3
mindistto19	-7.5	4.0	1.2	13.4	50.3	-8.7	27.4	-5.2	-2.3	-6.3	-1.5
mindistto4	-11.5	9.6	-6.0	1.6	74.7**	0.2	63.8	-5.1	5.5	0.5	-7.1
share2	9.7	5.5	22.9	3.3	-27.9	-5.2	-14.1	-6.7	6.8	5.7	20.0
share3	1.2	5.0	13.5	7.1	10.6	-11.0	9.0	-10.9	2.9	0.9	13.3
firmdensity	11.3	15.2	4.7	-1.5	-47.0	12.7	-20.8	10.3	3.6	3.3	-14.8
popudensity	7.9	0.9	9.2	4.5	-32.1	7.2	-30.1	2.9	8.9	6.9	-9.0
educ	-6.0	-1.5	4.9	-1.0	42.9	-10.9	18.6	-14.9	15.0	1.5	24.3
lnpergdp98	10.4	8.8	29.6*	4.6	-38.0	-5.2	-31.4	-4.6	6.5	6.7	1.7

Covariate	3.4	4.1	4.2,5.2,5.6	4.3	4.4,5.4,5.8	4.5	4.6	6.1	6.2	6.3	6.4
area	3.9	3.4	2.8	2.5	11.8	4.0	0.2	-14.0	-2.9	-16.0	11.8
mindistto19	-6.8	9.6	-19.5	24.2*	-31.9*	5.1	-11.9	-3.7	-19.5	2.8	-31.8*
mindistto4	0.5	8.3	16.6	13.6	26.6	2.1	12.2	-9.7	16.6	-4.8	26.6
share2	9.3	-1.8	2.0	0.2	10.6	-5.1	-2.1	5.4	2.2	21.0	11.8
share3	7.3	1.9	-2.4	20.2	15.4	5.0	-6.5	-2.8	-2.5	12.6	15.7
firmdensity	5.8	28.9**	-9.1	12.4*	-2.7	10.8	-0.1	8.2	-9.1	8.7	-2.7
popudensity	7.0	12.0	4.0	17.0	-12.0	-8.8	7.2	11.2	4.0	4.0	-11.9
educ	3.3	-3.2	-16.2	18.8	-2.1	0.7	-12.2	-11.6	-16.3	4.8	-2.1
lnpergdp98	9.3	3.8	2.9	3.2	12.2	8.2	3.1	8.7	2.9	32.1*	12.2
investment						7.4	7.7				
output						17.0	8.2				

Note: The number in the first and eleventh line means the related table and column in Table 1-6. For example, 2.1 means data related to column 1 of Table 2.
* Significant at the 10 percent level, ** Significant at the 5 percent level.

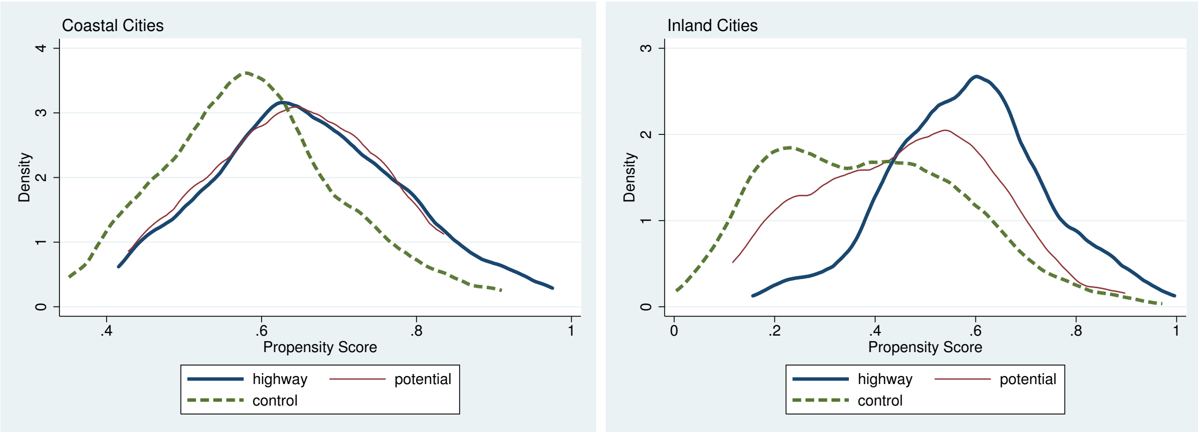

Figure A.1: Kernel Density of Propensity Score (1)

Figure A.2: Kernel Density of Propensity Score (2)