Development Policies and Economic Geography in China: Transport Infrastructure and Natural Resource

著者	Xu　Hangtian
学位授与機関	Tohoku University
学位授与番号	11301甲第16146号
URL	http://hdl.handle.net/10097/59647

Development Policies and Economic Geography in China:
Transport Infrastructure and Natural Resource

by

Hangtian Xu

January 2015

Acknowledgements

First and foremost, I want to thank my supervisor Kentaro Nakajima. You have been the best Ph.D. supervisor I could have wished for, giving me all the freedom in following my own research interests and being always available and patient in hearing and commenting on my research ideas and drafts. Without your insight comments and encouragement, this research would certainly not have been what it is now. Also, I appreciate for your suggestions and supports concerning on my personal academic career.

In addition, I am indebted to Shiro Hioki, my sub-supervisor, and Dao-Zhi Zeng. You are always ready to listen to me when I visit your offices, and sincerely provide me advices regarding my Ph.D. research and academic career. In particular, I appreciate for the chances attending the weekly economics seminar of Zeng laboratory, in which I learn a lot on theoretical economics.

I also appreciate for the suggestions concerning on this dissertation made at meetings: ARSC and JEA (Japan), YES, CCE and RUSE (China), ASRS (Taiwan), ABCDE (US) and UEA (Russia); seminars at Tohoku University, Tongji University, Fudan University, GRIPS and Nagoya University; in particular, the comments made by Ching-mu Chen, Zhao Chen, Chishio Furukawa, Yoshitsugu Kanemoto, Tatsuaki Kuroda, Takashi Kurosaki, Zhigang Li, David McKenzie, Hikaru Ogawa, Xiaoma Tao, and Hao Zhou.

Further, I thank my fellow students at Tohoku University. I am particularly grateful to, Mengmeng Cao, Chuan Chen, Qinghua He, Jie Ni, Jie Li, Nan Li, Liang Liang, Yong Song, Shan Wang, you enriched my life in Sendai. Then, I am grateful for the financial supports from China Scholarship Council, and Japan Society for the Promotion of Science.

Finally, I would like to express my special appreciation to my parents.

Contents

Acknowledgements **i**

Contents **ii**

List of Figures **iv**

List of Tables **v**

1 Introduction **1**

2 Highways and Industrial Development in Rural China **5**
2.1 Introduction 5
2.2 Transport infrastructure and regional development 7
2.3 Empirical strategy and data 8
  2.3.1 Empirical strategy 8
  2.3.2 Data 11
2.4 Baseline results and robustness checks 14
  2.4.1 Baseline results 14
  2.4.2 Placebo test 18
  2.4.3 OLS approach with propensity score 19
2.5 Further estimates and discussion 23
  2.5.1 Highway effects in different geographical locations 23
  2.5.2 Highway effects on industry structure 25
  2.5.3 Net spillover effects or substitution effects? 27
  2.5.4 Highway effects or railway effects? 28
2.6 Conclusion 29

3 Speeding-Up of Railways and Exports: Evidence from Silk Road **30**
3.1 Introduction 30
3.2 The China railway speed-up project 34
3.3 Empirical strategy 36
  3.3.1 Speed-up effects on national railway traffic 36
  3.3.2 Exogeneity concern 37
  3.3.3 Quasi-experimental settings and data 39
3.4 Baseline results and robustness tests 43
  3.4.1 Baseline results 43
  3.4.2 Robustness tests 47

3.5 Mechanisms . . . 50
3.5.1 Channels of the speed-up effects on exports . . . 50
3.5.2 Heterogeneous effects with respect to unit price of exporting products and firm size . . . 54
3.6 Exports of Xinjiang . . . 56
3.7 Conclusion . . . 60

4 The Role of Coal Mine Regulation in Regional Development 62
4.1 Introduction . . . 62
4.2 Resource curse . . . 64
4.2.1 Resource curse and institutions . . . 64
4.2.2 Within-country resource curse . . . 65
4.3 China's coal mines and coal mine regulation . . . 66
4.4 Theoretical framework . . . 68
4.5 Empirical results . . . 71
4.5.1 Empirical strategy . . . 71
4.5.2 Coal mine regulation and the whole regional economy . . . 72
4.5.3 Coal mine regulation and non-coal based economy . . . 75
4.6 Robustness checks using mortality rate . . . 76
4.6.1 Coal mine regulation measurement: why mortality rate is a suitable proxy . . . 77
4.6.2 OLS relationship between mortality rate and GDP growth . . . 78
4.7 Mechanism . . . 82
4.8 Conclusion . . . 85

A Appendix of Chapter 2 86

B Appendix of Chapter 3 90

C Appendix of Chapter 4 93
C.1 Data source . . . 93

List of Figures

2.1 Mileage of highways in China 7
2.2 Geographic units of mainland China related to this study 12
2.3 Selected county samples from the matching procedure 14
2.4 Output (value-added) of industry in 1998 and 2007 20

3.1 The geographic proximity of China and Central Asia 31
3.2 China's railway network in 2009 36
3.3 Ratio of rail/non-rail freight exports from China to Central Asia 40
3.4 Ratio of rail/non-rail freight exports from China to Central Asia (2) . . . 41
3.5 Kernel density of export transaction distribution with respect to sped-up mileage 45
3.6 Ratio of railway freight exports from China (Xinjiang/other regions) to Central Asia 46

4.1 The relationship between entrepreneurs and productivity 69
4.2 The number of TVE coal mines in China 70
4.3 Employment in China's coal mining industry 70
4.4 The product share of small-size coal mines in total coal production 72
4.5 The mortality rate of coal mining in China 79

A.1 National trunk highway system 86
A.2 National highway system planned 7918 highway network 87
A.3 Kernel density of propensity score (1) 87
A.4 Kernel density of propensity score (2) 87

List of Tables

2.1 First stage results (Probit regression) 16
2.2 Estimation of highway treatment effects 18
2.3 Estimation of counterfactual highway treatment effects 19
2.4 OLS regression with propensity score and highway dummy 22
2.5 Number of highway toll stations in sample counties 23
2.6 OLS regression with propensity score and count data of toll stations . . . 23
2.7 Estimation of highway treatment effects in different geographical locations 25
2.8 Estimation of highway treatment effects within different industries 26
2.9 Estimations using labor productivity and employment 27
2.10 Estimation of the source of highway spillover effects 28

3.1 Schedule of the railway speed-up project 34
3.2 Achievements of the railway speed-up project 35
3.3 Railway turnover and traffic, with respect to the speed-up project 37
3.4 Baseline and placebo regressions 44
3.5 Effects of the speed-up on mileage 45
3.6 Effects of the speed-up on the number of shipments 47
3.7 Demand fluctuations from specific products 49
3.8 Effect of the “Sept. 11 Attacks” 50
3.9 Two kinds of diversion effects 52
3.10 Distinguishing between time and capacity effects 53
3.11 Distinguishing between extensive and intensive margins 54
3.12 Effect of products with different unit prices 56
3.13 Size of the exporters 56
3.14 Export performance of Xinjiang 59

4.1 DID approach with respect to whole regional economy 75
4.2 DID approach with respect to the non-coal based economy 77
4.3 Regressions with mortality rate 80
4.4 Regressions with additional mortality rates 82
4.5 Mechanism (intra-industry) 83
4.6 Mechanism (whole economy) 84

A.1 Mean of covariates used in the matching procedure 88
A.2 Mean of outcome indicators 89
A.3 Post-matching differences in covariates (bias %) 89

B.1 Descriptive summary of Central Asia economy 90
B.2 Rail transport infrastructure improvements in Central Asia 91

B.3 Descriptive summary of dataset 91
B.4 Sped-up mileage to Urumqi (Xinjiang) 92
B.5 Effect of unit price changes 92

C.1 Placebo tests 93

Chapter 1

Introduction

China's phenomenal economic rise is a key point of the world economy in recent three decades. China's GDP increases by 9.68% averagely from 1978 to 2006 (Data source: National Bureau of Statistics of China). Although most of the developed countries experienced rapid economic growth in the early 1900s, or after the Second World War, no other countries continued such a high GDP growth in so long a period. By the rapid economic growth, accompanied by industrialization and urbanization, China's social and economic structure changes greatly. Strong economic growth gives strong demands in both production and consumption sides. As a result, resource shortages on social infrastructure and natural resource, defects in institution on the rise of market economic system and the decline of planned economic system, urban-rural and coastal-inland conflicts become the topical issues to be solved to support China's lasting development. Under this background, a number of development policies are proposed.

National development policies are targeted to create comprehensive welfare gains in the country. Inevitably, they might induce differential outcomes across regions and people. Identifying the precise effects of a policy is a complex and challenging task, but it is important to give reference to policy-makers so as to improve the targeting and efficacy of policies. Now, comparing policy effects across different regions and countries is receiving greater attention, as projects target larger populations and become more ambitious in scope, and researchers acquire enough data to be able to test specific policy questions across localities (Khandker et al., 2010). Moreover, to evaluate the impact to development policies, we need take consideration of both their overall outcomes and their side (unintended) effects to specific regions/sectors/people. That is, given the target of relevant policies, the policy impacts to non-targeted regions/sectors/people should also be taken consideration. However, it is always been neglected in policy impact evaluation, due to data availability and difficulties on measurement.

Giving the increasing data availability at the micro-level, there is a rich body of existing studies on the impact evaluation of development policies respect to both the comprehensive gains and benefit for targeted regions/sectors/people. However, for the policy side effects to non-targeted regions/sectors/people, empirical studies evaluating the impact are still far from adequate, and conclusive. Further, the development experience in China, as the largest developing economies, is a good reference for the policy-makers in other developing countries, suggesting the impact evaluation of China's development policies is important. This research thus contributes to the existing empirical studies on the impact evaluation of China's development policies by focusing on the non-targeted regions and sectors. Particularly, this dissertation gives a large body of discussions in the scope of space, namely, it is concerned about the heterogeneity of program impact on the distribution of economic activities, to present the dynamic evolution of economic geography with respect to development policies.

This dissertation consists of three parts, which is a collection of three independent essays in empirical impact evaluations on development policies and economic geography in China, precisely, examination of the (positive) side effects of China's three national development policies. It focuses on three kinds of stylized regions and investigates their regional development with respect to development policies: peripheral regions (rural counties/county-level cities) and highway network building, border regions (Xinjiang, the gateway of China and Central Asia) and railway speed-up project, natural resource-intensive regions (regions with rich coal mine endowment) and coal mine regulation.

Specifically, the first part centers on the highway placement for non-targeted counties, for which the industrial development is positively affected by the side effects of national highway network building. The second part involves the speed-up of China's West-East railway lines and its side effects to China-Central Asia exports, it examines the two-fold effects of exogenous domestic transport improvements on export performance. The last part discusses the coal mine regulation and its direct/indirect impacts to economic growth. Coal mine regulation is targeted only for the coal mining sector, however, its crowding-out effects to capital and entrepreneurs, from coal mining sector to other sectors, positively promote local economy in both coal mining sector and other sectors.

The methodology in this dissertation is based on econometrical and statistical tools, precisely, OLS estimation, difference-in-difference estimation, and propensity score matching estimation, using various datasets. The dissertation designs and sets up related transport infrastructure placement and policy change as quasi-experiments to conduct quantitative analysis and observe heterogeneous impacts among regions and sectors. The following is an overview of the chapters.

In Chapter 2, we estimate how highways (*Gaosu Gonglu*) affected industrial development in China's rural regions between 1998 and 2007, a period in which China experienced sharp growth in highway mileage. Highway routes between large cities pass through some rural counties, but not others. Based on this exogeneity of rural counties to highway routes (i.e., counties' highway connection status), we conduct a difference-in-difference propensity score matching (DID-PSM) estimation using counties' highway connection status and their industrial performance before and after being connected to a highway as treatment variables. Our empirical estimates indicate that highway access promotes industrial development in those counties with a higher output and greater level of investment. However, counties that are more than 300 km away from the large cities do not benefit from highway connections. Furthermore, highways tend to promote the development of heavy industry over light industry in rural regions. Labor productivity exhibits few positive effects. Finally, we investigate the source of the estimated highway effects. We find that they are the result of the net spillover from building the infrastructure, and not because a county is a substitute for its unconnected neighbors.

Chapter 3 focuses on the impact evaluation of railway speed-up project on export performance. China's main railway line, linking the east and the west, was sped-up on Oct. 21, 2000, which improved freight efficiency between Eastern China and Xinjiang (the gateway from China to Central Asia). This chapter tests the impact of exogenous domestic accessibility variation on China's exports to Central Asia. By employing a transaction-level database, empirical results find that benefited exporters (using rail freight) increase the export value by 30%, compared with exporters using other freight modes. The speed-up effect is due to mixed channels: net export creation, and export diversion in freight modes and exporters. Exports of medium-value products benefit most. Increase in export is exerted by export expansion of existing exporters, not by new entries. Additionally, on the exports of Xinjiang, the speed-up effect is two-fold. Speed-up weakens Xinjiang's locational advantage in China-Central Asia trade, but promotes its export to other international markets through better accessibility to the coast.

Finally, we turn to the economic development of resource-based regions in Chapter 4. In response to high mortality rates and low productivity in coal mining, China began regulating coal mines in the 1990s, reshaping its coal economy. We empirically investigate the relationship between coal mine regulation and economic growth in China. Using two difference-in-difference approaches to compare the pre- and post-regulation periods as well as regions heavily/lightly affected by regulatory policy, we find that regulation positively affects regional economy. This result is further illustrated using an OLS estimation that uses mortality rates in coal mining as a proxy for measuring the quality of regulation. The effects are not limited to the intra-coal industry but also apply to the economic spillover of related regions by relieving the crowding-out effects of coal abundance; that is, resource abundance tends to crowd out investment, human capital, and innovation in non-resource sectors and thus hinders economic growth.

Chapter 2

Highways and Industrial Development in Rural China[1]

2.1 Introduction

China built its first limited-access highway (*Gaosu Gonglu*)[2] in 1988, and since then, 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 (see in Figure 2.1). Compared to ordinary roads, highways have advantages in terms of road quality, congestion, and driving speed limits, which make transportation more efficient. It is widely accepted that transport matters to industry location and economic development, since economic development is much more dependent on cross-border and inter-regional trade now than in any period in history. This makes the case for an empirical investigation of highway effects, because highway networks play a core role in reducing shipping costs and improving transport efficiency. Furthermore, during the process of rapid industrialization, developing economies such as 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.

---

[1] This is a joint-work with Kentaro Nakajima. We thank Ching-mu Chen, Dao-Zhi Zeng, Shiro Hioki, Sho-ichiro Nakayama, Makoto Tsukai, Chishio Furukawa, Ruixiang Xiong, Zhigang Li, Zheng Michael Song, Dennis Tao Yang, Ming Lu, Zhao Chen, Hao Zhou, and Xiaoma Tao, as well as the participants of ARSC (Aomori), CEAC (Shandong), the JEA Spring Meeting (Toyama), YES-China (Jinan), CCE (Fudan), and seminars at Tohoku University and Tongji University for their many valuable comments that have significantly improved this research. We also gratefully acknowledge the financial support from the Graduate School of Economics, Tohoku University, Japan Society for the Promotion of Science (#22223003, #25380275, #24330072) and the 21st Century Cultural and Academic Foundation. All errors are ours.

[2] We refer to a limited-access highway simply as a highway in this chapter.

Given this background, we quantitatively investigate the role of highway placement on rural industrial development. Access to China's micro-level industry and highway placement databases allows us to estimate the impact of the highway infrastructure on economic and social development in the period 1998 to 2007, during which most of the major highway network was developed. Then, by distinguishing between counties with and without highway access, we use the difference-in-difference propensity score matching (DID-PSM) method to compare the industrial performance of these two kinds of counties. Focusing on China's rural counties[3], we find that newly placed highways stimulate industrial growth in connected counties that have a greater level of investment and a higher output.

The DID-PSM results are robust to a placebo test involving those counties with no highways in our research period (1998–2005). The test compares the industrial performance of the counties that still had no highways in the subsequent period, 2006–2011, to the performance of those counties that planned to have highways in the period 2006–2011 (planned highway cities). Our results show there is no highway effect between these two kinds of counties during the research period, which partially controls the self-selection problem between highway placement and industrial performance. As an alternative control mechanism, we enter the propensity score into the OLS regression investigating the highway effects. After controlling for the propensity score, we find that highway connections positively affect industrial performance, which is consistent with the DID-PSM results.

Furthermore, highway density is also positively correlated with industrial performance. In addition, we find that remote counties (i.e., counties more than 300 km from large cities) do not benefit from newly placed highways and that highway access promotes the development of heavy industry over that of light industry. These results suggest that highway effects depend heavily on geographical location and industry type. Finally, we investigate the source of the estimated highway effects, and find that they are the result of the net spillover from building the infrastructure, and not because a county is a substitute for its unconnected neighbors.

This study contributes to the existing literature in two respects. First, we develop a new way of identifying highway accessibility. Here, we view highway toll stations as indicating physical access to highways. In China, all vehicles must pay to use the highway (*Gaosu Gonglu*) system, and this is managed by having a toll station at every highway access point. To the best of our knowledge, we are the first to apply this idea to an empirical study. We define a county as being connected to the highway system

---

[3] We consider county-level cities (e.g., *Xianjishi*) and counties (e.g., *Xian*) to be equivalent in this research, as per Qin (2014).

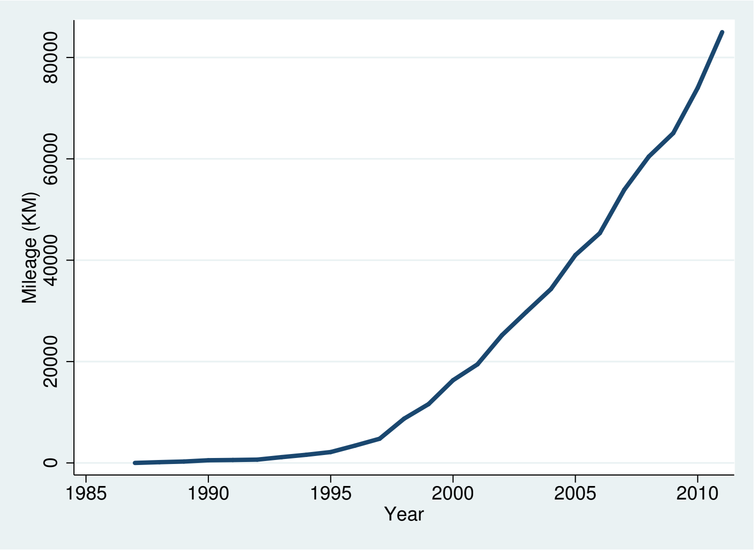


Figure 2.1: Mileage of highways in China

if there is at least one highway toll station located within in its borders. Second, we employ a careful strategy to make causal inferences about highway treatment effects, using the experimentalist approach toward highway issues. Here, we use the propensity score matching method and construct a placebo test to check for potential counterfactual economic effects in the absence of highway access.

The rest of this chapter is organized as follows. Section 2.2 contains a brief review of issues related to transport and regional development. In Section 2.3, we present our empirical strategy and a description of the data set. Section 2.4 gives the baseline results and robustness checks, and Section 2.5 provides further estimates of the highway effects. Finally, Section 2.6 concludes.

2.2 Transport infrastructure and regional development

There has been a considerable amount of theoretical research 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). On the other hand, because of decreasing transport costs and the increasing importance of non-material flows, others doubt the importance of transport infrastructure to firm location (Banister and Berechman, 2000). In addition, the distribution of the benefits from infrastructure investment is not clear once we assume mobility of firms (Venables, 1996; Puga, 1999).

Empirically, Chandra and Thompson (2000) find that highways have a differential impact across industries: certain industries grow as a result of reduced transportation costs, while others shrink as economic activity relocates elsewhere. Boarnet (1998) find that infrastructure investment (i.e., street-and-highway investment) has positive output spillover effects within the same county and negative output spillover effects in other counties. Holl (2004a, 2004b) proves that new firms prefer locations close to new motorways in Spain and Portugal. Rephann and Isserman (1994) examine the effects of interstate highways on counties in the U.S., and find that the areas that benefited most are those in close proximity to large cities or with some degree of prior urbanization. Furthermore, population and tertiary industry are affected most by highways, while manufacturing is least affected. Rothenberg (2011) suggests that road improvements are accompanied by a significant dispersion of manufacturing activity in Indonesia.

The empirical literature on China's transport infrastructure construction and its effects on social and economic development is still inconclusive. Baum-Snow *et al.* (2012) and Baum-Snow and Turner (2012) explore the decentralization effects of highways on the urban population in the core cities of prefectures. Banerjee, Duflo, and Qian (2012) show that proximity to transportation networks has a moderate, positive causal effect on per capita GDP levels across sectors, but no effect on per capita GDP growth. Roberts *et al.* (2012) find that aggregate Chinese real income was approximately 6% higher than it would have been in 2007 without the highway network. Faber (2014) investigates the effects of highways on industrial performance and finds that, relative to core regions, highways have a negative effect on the connected peripheral regions. Xu and Zhou (2014) find highway connections help attract new firms in rural China and that this effect results from a net spillover, and not because a county is a substitute for its unconnected neighbors. In summary, evidence from previous research consistently indicates that transport infrastructure has a significant effect on economic development.

2.3 Empirical strategy and data

2.3.1 Empirical strategy

If road improvements affect economic performance, then there will be a difference in performance between regions with and without highway access in the long run. Thus, we first need to develop a strategy to identify road accessibility. Many studies have been forced to work with partial proxies, such as the road distance to the nearest interstate highway (highway) (e.g., Baum-Snow, 2007; Holl, 2004b), the number of highway "rays" emanating from central cities (e.g., Baum-Snow *et al.*, 2012), the shortest travel time to places of work (e.g., Sanchis-Guarner and Lyytikäinen, 2011), binary variables that capture the presence of an airport or a port (e.g., Costa-Campi and Viladecans-Marsal, 1999), and regional transport infrastructure density (e.g., Nicolini, 2003). In this study, we identify highway accessibility using highway toll stations. A payment is required to access all highways in China. Vehicles access a highway through a toll station, making toll stations a straightforward approximation of highway placement and highway access.

Second, we are concerned about the self-selection problem in highway placement. 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, and so on (see Rietveld and Bruinsma, 1998; Holl, 2004a; Baum-Snow, 2007). Thus, most empirical research on highway construction faces the self-selection problem. Regions with highway access usually dominate other regions in many other aspects, so even without highway access, such regions would still exhibit 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 (e.g., elevation range, terrain ruggedness, groundwater) are used as a source of exogenous variation to relieve the bias (e.g., Baum-Snow, 2007; Baum-Snow *et al.*, 2012; Duranton and Turner, 2012; Faber, 2014).

A sensible instrument is often hard to find, and weak instruments cannot efficiently eliminate the bias. However, when the self-selection is based on observable factors, the methods available for controlling the selection bias are relatively straightforward (Zhao, 2004). Counterfactual analysis by matching selects treated and comparison observations with similar covariates to correct the selection bias, and may be efficient. When focusing on micro county-level samples, it is feasible to construct balanced treated and comparison observations by conducting a counterfactual analysis, since highway placement at this level is supposed to be exogenous. Highways are always built to enable quick and efficient movement between large cities, so smaller counties tend not to have a say in the placement or shape of a highway route.

Evidence from previous research is informative on this assumption. Motorway project placements can be assumed to be exogenous to changes at the municipality level, because 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). Holl (2004a) and Donaldson (2014) provide similar evidence for India's railroads and Spain's motorways, respectively. As a result, whether small geographical units such as wards and counties are connected by highways might be determined by exogenous factors such as geographical location.

To achieve the counterfactual analysis, we apply propensity score matching. First, we select covariates that affect both highway treatment participation and the potential economic/industrial outcome. Although we assume that highway placement at the county level does not depend on the local economic/industrial situation, some economy, industry, or population covariates can, to some extent, predict the probability of highway placement. Then, we run a probit regression using highway treatment as the dependent variable (1: if the county has access to highways; 0: if it does not) and the selected related covariates as independent variables. Then, the predicted value from the probit regression becomes the propensity score (0 <propensity score <1). Based on the propensity score of each sample county, the PSM method matches each treated highway county with the most similar control counties (weighted).

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

$$\text{ATT} = E[Y_1 - Y_0|D = 1] = E[Y_1|D = 1] - E[Y_0|D = 0] + E[Y_0|D = 0] - E[Y_0|D = 1],$$

where D is a dummy variable for the treatment ($D = 1$: treated; $D = 0$: control), and Y_i represents the outcome indicator (Y_1: the outcome if it is treated; Y_0: the outcome if it is a control). PSM aims to ensure that the selection bias goes to 0 after conditioning on the propensity score, so that the counterfactual part, $E[Y_0|D = 1]$, can be dropped in the ATT estimation:

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

where $P(X)$ is the propensity score with covariate X. Therefore, given that the required assumptions hold, the post-PSM ATT estimator can be written as

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

This is the preferred form, and can be estimated directly using the observable data.

Smith and Todd (2005) find that the difference-in-difference propensity score matching (DID-PSM) estimate is more robust than traditional cross-section matching estimators. As such, we apply the DID estimate to the PSM estimate to construct a modified estimation strategy. Specifically, we set $Y_{i,j}$ as the output of county j at time i (i=0: pre-highway treatment, i=1: post-highway treatment; j=0: control county, j= 1: county gained highway access at time 1). The DID highway treatment effect is estimated as follows:

$$\text{Highway effect (DID)} = \frac{E[Y_{1,1} - Y_{0,1}|D = 1, P(X)] - E[Y_{1,0} - Y_{0,0}|D = 0, P(X)]}{E[Y_{1,0} - Y_{0,0}|D = 0, P(X)]}.$$

In addition, to check whether the required assumption is satisfied in the PSM estimate, we present the pre-treatment difference estimation. Based on the above discussion, we know that if PSM relieves the selection bias well, then conditional on the propensity score, the pre-treatment difference between a treated highway county and a control county is expected to go to 0, since neither had the treatment at the beginning (i=0); that is,

$$\text{Pre-treatment difference} = \frac{E[Y_{0,1}|D = 1, P(X)] - E[Y_{0,0}|D = 0, P(X)]}{E[Y_{0,0}|D = 0, P(X)]} = 0.$$

Furthermore, based on the same PSM strategy with respect to the highway effect (DID) and pre-treatment difference, we conduct a placebo test by applying alternative matching samples. That is, we test whether the estimated highway effects (DID) are also observed in the absence of highway placement. This strategy might relieve the self-selection problem in highway placement.

2.3.2 Data

The existing highway network in China was mainly undertaken based on two national highway plans. The first was issued in 1992 and called the *5-Zong-7-Heng* National Trunk Highway System, and included five longitudinal lines and seven latitudinal lines. The planned length was 34,422 km. This plan aimed to connect the hubs, main ports, all cities with a population above one million, and most of the cities with a population above 500,000. This plan was completed by the end of 2007. The second extension plan, called the National Highway System Planned 7918 Highway Network, was issued in 2004. This plan aimed to extend the first plan by shaping the final national highway network with seven capital lines, nine 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 A.1 and A.2. The finished part of the second plan, together with the provincial highways, connection lines, and ring lines of large cities, form China's current highway network.

We chose the county as the basic geographical unit to construct treatment and control observations with a sufficient sample size. In China, the county is fourth in the hierarchy of administration units, following the central government, provincial unit, and prefecture

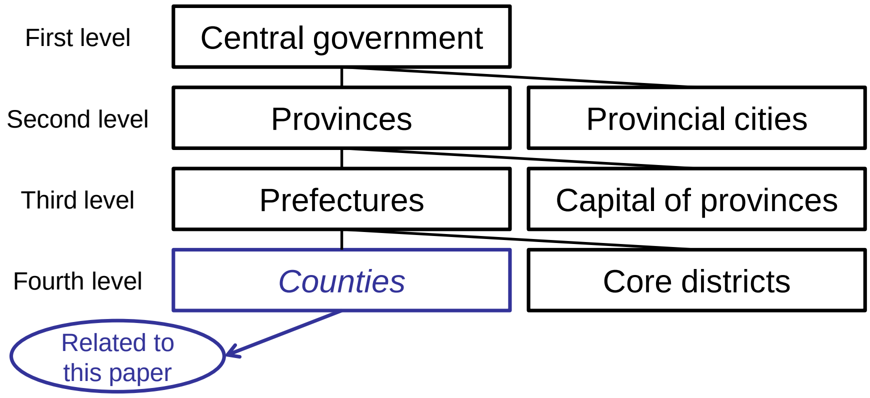


FIGURE 2.2: Geographic units of mainland China related to this study

unit (see Figure 2.2). As a result, our samples are able to indicate the rural regions in China.

County-level socio-economic and highway data used in this analysis come from several sources.

(1) We collected the highway toll station data from the provincial public travel information service system, and geo-coded the location at the county-level. In China, all vehicles must pay to use the highway (*Gaosu Gonglu*) system. Thus, highway toll stations indicate highway placement and highway access. This data set includes all 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, although these were not included in the data set. Any county with at least one highway toll station within its borders is classified as a highway county (i.e., a county with access to a highway).

(2) Although the first highway was built in 1988, micro-level data in that period is not readily available. The earliest available micro-level industry data set is the Chinese Industry Statistical Database (CISD) from 1998. The CISD includes all state-owned industrial firms and other industrial firms with annual sales of more than five million Chinese Yuan. In addition, China's market economy gained momentum in 1997, when 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, we set 1998 is as the starting point for the analysis. For precision, we exclude counties with highway placement before 1998 in the following analysis. After 2007, the CISD no longer reports on firm-level annual value-added, so this date marks the end of our analysis. China's highways experienced significant development between 1998 and 2007, growing from 8,733 km to 53,913 km. Therefore, we assume that they had a significant economic impact on related regions.

(3) Population data come from the 5th national population census at the county-level, which recorded the population, level of education, and employment status at the end of 2000.

(4) GDP and per capita GDP data at the county-level come from the Provincial Statistical Yearbooks of the designated years.

We merged the data sets based on the unique administrative division code (Xinzheng Quhua Daima). County-level units in all data sets are corrected to match the 1998 administrative divisions. To compare real growth in GDP/per capita GDP, industrial output (value-added), and investment (net value of fixed assets) in an 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.

Specifically, we choose the sample counties in four steps. First, we exclude counties in 11 provinces because of data restrictions. Second, counties in the four provincial counties (Beijing, Tianjin, Shanghai, and Chongqing) and all provincial capitals are excluded. These large counties 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, counties in these areas tend to be affected by the spillover effects of the large cities and urban sprawl. Finally, we exclude all counties in prefectures that had highway access before 1998. The related data come from *Gonglu* (a journal on traffic studies in Chinese; see 2002(4): 132-140), which summarizes the progress of China's highways from 1988 to 2001.

To achieve the empirical procedures in the following sections, we further divide the sample counties into three categories:
∗ *highway counties: gained access to highways between 1998 and 2005;*
∗ *planned highway counties: gained access to highways between 2006 and 2011;*
∗ *control counties: no highway access as of the end of 2011.*

Figure 2.3 and A.2 show that highways are denser in the coastal regions (Eastern China) than in the inland regions (Western and Central China). Most of the sample counties in the coastal regions have highway access. This is not strange based on China's economic geography. The coastal regions have a higher population and economic density. Thus, we further divide the samples into two sub-groups: coastal regions and inland regions, and perform the matching separately. Following the traditional classification of China's

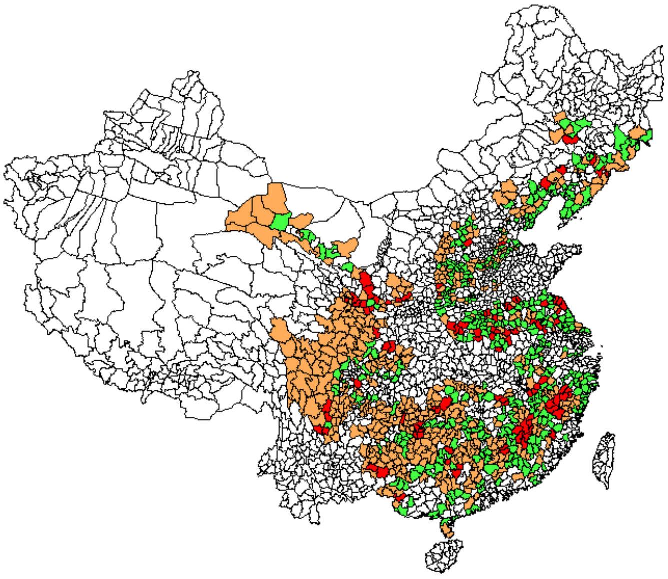


Figure 2.3: Selected county samples from the matching procedure
Note: Green units: highway counties; red units: planned highway counties; orange units: control counties. Distinguishing between highway counties and planned highway counties is achieved using the digitalized map of China's highway lines in 2005 from Baum-Snow and Turner (2012) and the Georeferencing module of ArcGIS.

economic geography, we define six of the 16 sample provinces as coastal regions (Zhejiang, Jiangsu, Fujian, Guangdong, Liaoning, and Hebei) and the remaining 10 sample provinces as inland regions (Jilin, Anhui, Jiangxi, Hu'nan, Guangxi, Gansu, Shanxi, Sichuan, Guizhou, and He'nan). The final prepared sample consists of 983 counties belonging to 16 provinces, accounting for 36.4% of the Chinese population (5th national population census in 2000), and 12.2% of national industrial output (value-added) in 1998. In 2007, the latter figure increases slightly to 12.6%.

2.4 Baseline results and robustness checks

2.4.1 Baseline results

As mentioned in Section 3, we have data on the highway building status for two years, 1998 and 2005, and industrial outcome information for two years, 1998 and 2007. We assume that the period from 2005 to 2007 is the lag for the industry to respond to the recently gained highway access, so we set the post-treatment data set to 2007, and identify highway counties based on their highway placement status at the end of 2005. The highway data set does not include the precise date each toll station was installed, so it is not possible to estimate the average yearly highway effect per industry. As a compromise, we compare the difference in the total growth from 1998 to 2007 between the treated and control counties using the DID-PSM method. We choose covariates that may affect regional development and transport infrastructure placement, based on prior literature, and decide on the final number and form (such as intersection and logarithm) of the covariates based on three criteria. First, we require a sufficiently high predictive value from the propensity score for 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 calculated propensity score. Third, the difference in the covariates between the treated and post-matching control samples should be insignificant. In addition, to apply PSM without bias, the covariates used in the matching should either be fixed over time or measured before participation in the treatment.

Although highway planning and building are not subject to population and economic development at the county level, the population, economic, and geographical characteristics of the counties are good predictors of highway placement. Specifically, we choose nine covariates on population density, initial per capita GDP, location, et al., and apply these to the propensity score calculation. The first stage of the DID-PSM estimation is as shown in Equation 2.1, which provides the propensity score for each county. Here, *Dummy for highway* is a dummy variable that takes the value 1 for highway counties and planned highway counties, and 0 for control counties. The results are presented in Table 2.1 (the descriptive summary for these nine covariates are presented in Table A.1).

$$Dummy\_for\_highway_i = Probit(area\_size_i, pop.dens._i, initial\_p.c.GDP_i, et\ al.) + \epsilon_i. \quad (2.1)$$

We find that covariates matter more to the highway status in inland regions than in coastal regions. Furthermore, in both regions, highway counties have, on average, a lower share of tertiary industries. For coastal regions, highway counties tend to be closer to large ports. For inland regions, highway counties tend to have a higher initial per capita GDP, higher population density, and higher share of secondary industries.

The propensity score is then obtained by Equation 2.2, where $\widetilde{\epsilon_i}$ are the estimated residuals from Equation 2.1. We use the fraction of treated subjects (including both highway and planned highway counties) 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 counties and 0.3907 for inland counties. Using these cutoff values, the prediction rates for the treated and control observations are 65.2% and 60.4% in coastal county groups, and 90.9% and 52.0% in inland county groups, respectively. Using the same

Table 2.1: First stage results (Probit regression)

	(1) Dummy for highway	(2) Dummy for highway
Area size	0.0001	-0.0000
	(1.06)	(-1.61)
Minimum distance to large cities	-0.0010	0.0002
	(-0.75)	(0.57)
Minimum distance to large ports	0.0016**	-0.0004**
	(1.98)	(-2.38)
Employment share of secondary sector	1.0493	2.9835**
	(0.83)	(2.32)
Employment share of tertiary sector	-4.1602**	-3.2001**
	(-2.24)	(-2.01)
Firm density	2.8372	0.2824
	(1.14)	(0.10)
Population density	0.0006	0.0009***
	(1.11)	(2.78)
Population share (education⩾high school)	2.1627	1.1818
	(0.68)	(0.57)
Initial per capita GDP	0.1044	0.6837***
	(0.47)	(4.88)
Constant	-1.0611	-5.3759***
	(-0.55)	(-4.99)
Region	Coastal	Inland
No. obs.	276	707
Pseudo R^2	0.0559	0.1287

Note: Z-values in parentheses: **, *** indicate significance at the 5% and 1% levels, respectively.

propensity score calculation, we find in Figure A.3 that both highway counties and planned highway counties are easily distinguished from control counties. The mean of the related probability for the treatment (propensity score) is greater in treated highway counties than in control counties. Thus, the propensity score is a suitable predictor of actual treatment participation.

$$Propensity\_Score_i = Dummy\_for\_highway_i - \widetilde{\epsilon}_i. \tag{2.2}$$

Then, the matched pseudo R^2 is presented for each PSM treatment effect estimate in Tables 2.2, 2.3, 2.7, 2.8, and 2.9. We can see a satisfactory reduction in the pseudo R^2 between the raw and matched samples. Furthermore, the covariate bias (%) is reduced, and most covariates show no significant difference between the treated and post-matching control groups, even at the 10% confidence level. Detailed information is presented in Table A.3. Thus, we can assume that the covariates are highly correlated with both the participation indicator (highway placement) and industrial development, which is a requirement for applying PSM.

In the second stage of DID-PSM, we set the outcome growth ($Growth_i$) as the difference-in-difference with respect to the growth in 1998–2007 and the matched control observations. Here, $Growth_i$ is defined in Equation 2.3. Outcome indicators refer to the per capita GDP of the overall economy, and the firm density, investment (net value of fixed assets), and output (value-added) per industry. Then, $\Delta Outcome_i$ reflects the growth in absolute value for treated county i, and $\Delta Outcome_{matchedcontrol,i}$ shows the growth in absolute value for the control county matched to county i. We perform the matching using kernel matching algorithms. As a result, the related control county is a weighted average of many control counties. In other words, it represents a pseudo county rather than an actual county.

$$Growth_i = \Delta Outcome_i / \Delta Outcome_{matchedcontrol,i} - 1. \tag{2.3}$$

$$\textit{Pre-treatment}_i = \textit{Outcome1998}_i / \textit{Outcome1998}_{matchedcontrol,i} - 1. \tag{2.4}$$

The number of treated counties is set to A, and each treated county is numbered as $\mathbf{A} = \{1, 2, ..., A\}$. We then test whether the obtained $Growth$=$\{Growth_i | i \in \mathbf{A}\}$ and $\mathbf{0}$ are significantly different. If *Growth* is significantly different to **0** and positive, we interpret this as showing positive highway effects. To relieve the limitation caused by the sample size, we report the significance using bootstrapped standard errors (200 times). This strategy also applies in the following DID-PSM estimations. As a reference, we also test whether *Pre-treatment* (as in Equation 2.4) and **0** are significantly different. Here, *Pre-treatment* is symmetrical with *Growth*, but we replace the growth between 1998–2007, $\Delta Outcome_i$, with the initial value in 1998, $Outcome1998_i$. Since we test the post-matching difference, the pre-treatment difference between the treated and control samples should be nonsignificant. Therefore, we test whether the difference between *Pre-treatment* and **0** is nonsignificant.

Table 2.2 gives the estimates of the highway effect (DID) and pre-treatment difference for four county-level outcome indicators (see the descriptive statistics in Table A.2): per capita GDP and the industrial performance indicators (firm density, investment, and output). In the first four columns of panel A, we report the DID results for four indicators for coastal regions and the last four columns show the results for the inland regions. Then, as a reference, we present the pre-treatment differences (i.e., the estimations from Equation 2.4) in Panel B. For both the coastal and inland regions, 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/output in the treated counties is 101.79%/118.67% of that in the control counties, and the differences are nonsignificant. However, the DID results increase to 176.01%/139.58% and are significant at the 1%/5% confidence level. We also tested the results using radius matching algorithms rather than kernel matching algorithms, with the results being highly consistent (results are available upon request). However, the treatment effects on per capita GDP and firm density are nonsignificant. We explain this result in the following sections.

TABLE 2.2: Estimation of highway treatment effects

	(1) Growth in p.c. GDP	(2) Growth in firm density	(3) Growth in investment	(4) Growth in output	(5) Growth in p.c. GDP	(6) Growth in firm density	(7) Growth in investment	(8) Growth in output
Panel A: Highway effects (difference-in-difference)								
Highway	0.0829	0.0987	0.7601***	0.3958**	-0.0065	-0.1174	0.7230**	0.2818**
	(0.75)	(0.30)	(2.90)	(2.13)	(-0.09)	(-0.70)	(2.36)	(2.09)
Panel B: Pre-treatment difference								
Highway	0.0947	0.1456	0.0179	0.1867	0.0365	0.2169*	0.0523	0.1259
	(1.57)	(1.06)	(0.09)	(1.28)	(1.01)	(1.80)	(0.26)	(0.84)
Description								
Region	Coastal				Inland			
No. matched obs. (Treated/Control)	150/101				242/379			
Pseudo R^2 (Raw/Matched)	0.062/0.009				0.187/0.019			

Note: Samples that do not satisfy the *common support* are dropped; Z-values in parentheses: *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

2.4.2 Placebo test

In the baseline highway treatment effect estimation, we control for the observable covariates through PSM. However, the significant estimation results could be caused by the disturbance of unobservable variables if such variables can affect both highway treatment and industrial performance. In other words, this is the source of self-selection problem, but we omitted them because of data restrictions. One possible test is to check for a pseudo treatment effect using the same propensity score calculation and matching strategy, but using planned highway counties (treated after 2005) rather than highway counties (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. This period saw more than 30,000 km of newly constructed highways across the country, and many rural counties gained highway access for the first time. In this respect, if the improved performance in highway counties presented in Table 2.2 is indeed attributable to highway access, we can expect to find no significant difference between planned highway counties and control counties. This is because, in the period 1998–2005, planned highway counties had no highway access (they only gained highway access in 2006–2011). If this is true, we can also prove that whether or not gain access to a highway in the rural counties is not affected by their prior economic performances. To maintain comparability, we use the same matching strategies as in the baseline estimation.

Table 2.3: Estimation of counterfactual highway treatment effects

	(1) Growth in p.c. GDP	(2) Growth in firm density	(3) Growth in investment	(4) Growth in output	(5) Growth in p.c. GDP	(6) Growth in firm density	(7) Growth in investment	(8) Growth in output
Panel A: Highway effects (difference-in-difference)								
Highway	-0.1980	-0.1445	-0.2747	-0.1663	0.0337	-0.4250***	1.0301	0.1055
(Pseudo)	(-1.62)	(-0.59)	(-1.26)	(-0.63)	(0.32)	(-2.89)	(1.53)	(0.67)
Panel B: Pre-treatment difference								
Highway	-0.1475*	-0.3899**	-0.2397	-0.0392	0.0037	0.1572	0.2741	0.1060
(Pseudo)	(-1.73)	(-2.18)	(-0.79)	(-0.15)	(0.09)	(1.21)	(0.70)	(0.53)
Description								
Region	Coastal				Inland			
No. matched obs. (Treated/Control)	20/101				85/379			
Pseudo R^2 (Raw/Matched)	0.212/0.190				0.033/0.011			

Note: Samples that do not satisfy the *common support* are dropped; Z-values in parentheses: *, **, *** significant at the 10%, 5%, and 1% levels, respectively.

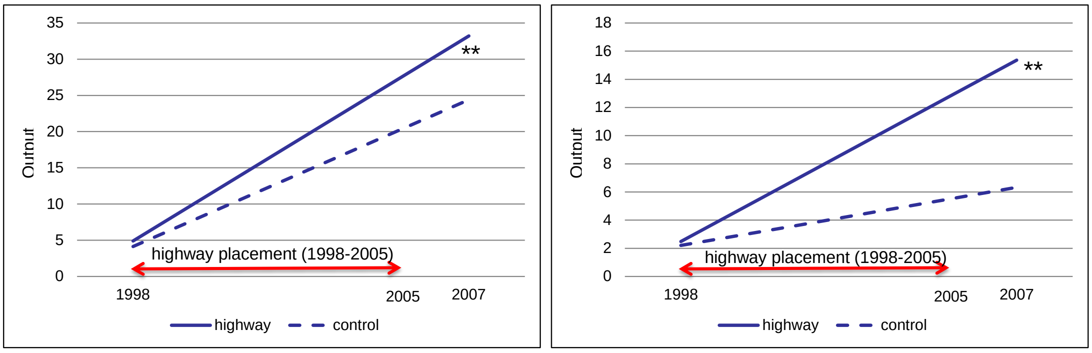
In Figure A.3, we can see that the propensity score is also able to distinguish between planned highway counties and control counties. Here, the average propensity score of planned highway counties is larger than that of control counties. The pseudo treatment effect estimates are consistent with the above assumption. Compared to the results in Table 2.2, the average difference between the industrial indicators in planned highway counties and the control counties is much smaller. Almost all of the pre-treatment differences and DID values of the industrial indicators are nonsignificant (the only exceptions being in firm density and per capita GDP, but their negative values are compatible with our opinion). In addition, we tested the results using radius matching algorithms rather than kernel matching algorithms, and the results were again similar (results are available upon request). For clarity, the trend in industrial output (value-added of industry) between 1998 and 2007 is presented in Figure 2.4. Highway counties tend to have faster growth in industry output than control counties, but planned highway counties show a similar trend to control counties. For the coastal planned highway counties, the growth in output is even slower than in the control counties. These figures are based on kernel matching estimates in industrial output, but the results are highly consistent under radius matching or when using data on industrial investment. Similarly, in Table 2.3, we do not find a significant difference in per capita GDP growth.[4]

2.4.3 OLS approach with propensity score

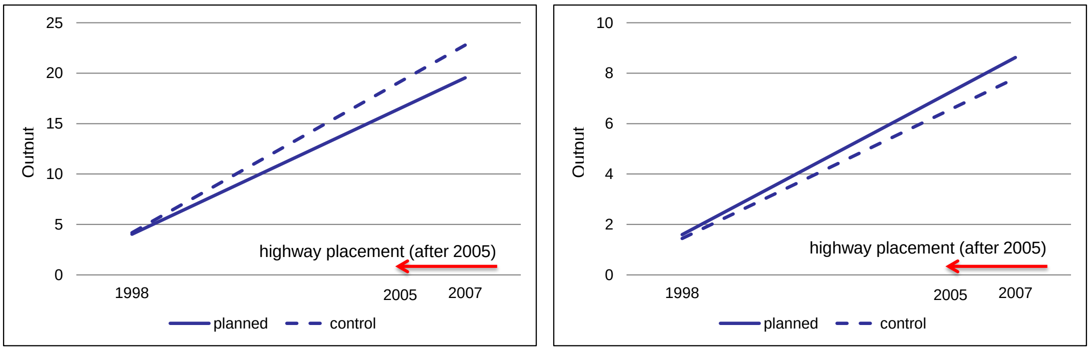
Our sample unit is the county in China. However, counties in China are divided into rural counties (*Xian*) and county-level cities (*Xianjishi*). Owing to national urbanization,

---

[4] PSM has a shortcoming in that, when the sample size is too small, it does not perform well. On the other hand, Mahalanobis distance matching (MDM) (see Rubin, 1980) is relatively robust under different settings (Zhao, 2004). See also Rephann and Isserman (1994) for an application of MDM. We thus also tested our baseline and placebo estimations using difference-in-difference Mahalanobis distance matching, with the results remaining unchanged. These results are available upon request.



(1) Coastal highway counties (2) Inland highway counties



(3) Coastal planned highway counties (4) Inland planned highway counties

FIGURE 2.4: Output (value-added) of industry in 1998 and 2007

Note: **: the difference in value-added between treatment and control groups is significant at the 5% level. The monetary unit in the figures is one billion Yuan, after deflating the price index to 1998 values.

some counties with a high population density or that perform well economically play, at least to some extent, the role of a city. These counties are called the *Xianjishi*. In the DID-PSM estimation, we treat counties (*Xian*) and county-level cities (*Xianjishi*) symmetrically (see also Qin (2014)) because matching them separately might cause an insufficient sample size in a quasi-experiment. Although the heterogeneity between *Xian* and *Xianjishi* might be reflected in the covariates we employ in the matching process, it causes a potential disturbance in the randomness of highway placement. Therefore, it might be informative if we were able to treat them separately. In addition, counties with a single highway toll station and counties with multiple toll stations are seen as equivalent in the DID-PSM estimations, which might lead to the omission of an important variation. Counties with multiple toll stations have a higher highway density than counties with single toll station, and the relationship between highway density and industrial performance is informative.

PSM makes few assumptions about the nature of the relationship between the confounders and the outcome, which makes matching the preferred method of applying the propensity score. However, when the sample size is not large enough and heterogeneity among observations needs to be well controlled, matching in sub-samples is not efficient.

As an alternative, entering the propensity score in a regression model for the outcome, just as we would any independent variable, is another way to investigate the treatment effects (see, Indurkhya, Mitra, and Schrag (2006); Basu, Polsky, and Manning (2008)). More specifically, to further understand how highway density and the heterogeneity in administrative units affect the estimates of highway effects, we run an OLS regression while controlling for the features of the *Xianjishi* and highway density.

First, we employ the growth rate of four output indicators in the period 1998–2007 as the dependent variable (ln (*Growth*)=ln (*Outcome2007*)-ln (*Outcome1998*)), the propensity score as the independent variable, and two dummy variables: *Highway* (takes the value 1 if there is at least one highway toll station in the area in the specific period); *Xianjishi* (takes the value 1 if it is a *Xianjishi*, 0 if it is a *Xian*). To control the initial output difference among counties, we insert the related initial values, η, as the fixed effects of a province or prefecture:

$$\ln(Growth_i) = \alpha_0 + \alpha_1 Highway_i + \alpha_2 Propensity\_Score_i + \alpha_3 Xianjishi_i + a_4 Initial_i + \eta_i + \epsilon_i. \tag{2.5}$$

The results are presented in Table 2.4. Note that we only show the results for industrial investment and output because the results for per capita GDP and industrial firm density are all nonsignificant (results are available upon request). In the first four columns, we estimate the highway effects of the outcome indicators using the full sample. The highway98–05 dummy variable is set to 1 for highway counties (counties that obtained a highway during the period 1998–2005)and 0 for planned highway counties and control counties. In columns 5–8, we run the same regression, but only include the planned highway counties (counties that obtained a highway during the period 2006–2011) and control counties (excluding those counties that had gained highway access by the end of 2005). Here, the highway06–11 dummy variable is set to 1 for planned highway counties and 0 for control counties. The results for highway effects are consistent with the baseline results. A significant highway effect on industrial investment and output is observed in columns 1–4 with province and prefecture fixed effects. However, in the counterfactual regression (column 5–8), the coefficients for the highway06–11 dummy are nonsignificant. The *Xianjishi* dummy captures the variation caused by sample heterogeneity, but the results are nonsignificant in all cases. We also find in most of the cases that the coefficients of the propensity score in Table 2.4 are significant at the 1% level. This suggests that the propensity score does well in controlling some of the factors that affect industrial performance. Baum-Snow *et al.* (2012) find that their instrumental variable (IV) decreases the effects of highways. This is consistent with the fact that highways tend to be located in places with better unobservables. Similarly, to ascertain whether the propensity score effectively corrects some endogeneity problems, we run the regression without controlling for the propensity score. The results show that the propensity score, like the IV in Baum-Snow *et al.* (2012), decreases the effects of highways. The coefficients for growth in industrial investment and output decreased after controlling for the propensity score. For example, in column 1 the relevant coefficient increased from 0.203 to 0.258 without the control for the propensity score, and in column 3 the coefficient increased from 0.167 to 0.220. In both cases, the level of significance remained at the 1% level (detailed results are available upon request).

Table 2.4: OLS regression with propensity score and highway dummy

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ln (Growth in output)		ln (Growth in investment)		ln (Growth in output)		ln (Growth in investment)	
Highway98–05	0.203***	0.179***	0.167***	0.168**				
	(3.38)	(3.05)	(2.59)	(2.39)				
Highway06–11					-0.049	-0.132	0.022	0.079
					(-0.46)	(-1.09)	(0.19)	(0.54)
Propensity Score	1.608***	1.246***	1.269***	1.131***	1.970***	1.499**	1.365***	1.341*
	(5.44)	(3.50)	(4.20)	(2.75)	(4.72)	(2.53)	(3.13)	(1.96)
Xianjishi	0.091	0.091	0.109	0.153	0.056	0.025	0.118	0.140
	(1.12)	(1.17)	(1.25)	(1.63)	(0.41)	(0.17)	(0.81)	(0.78)
ln (output initial)	-0.415***	-0.413***			-0.397***	-0.414***		
	(-13.39)	(-13.00)			(-9.64)	(-8.89)		
ln (invest. initial)			-0.401***	-0.410***			-0.403***	-0.422***
			(-12.24)	(-11.07)			(-9.12)	(-7.81)
Constant	5.296***	6.062***	5.215***	5.983***	5.083***	6.492***	5.330***	6.613***
	(16.45)	(14.22)	(14.15)	(11.39)	(11.24)	(8.31)	(10.23)	(7.00)
Sample	All	All	All	All	Planned& Control	Planned& Control	Planned& Control	Planned& Control
Province F.E.s	Yes		Yes		Yes		Yes	
Prefecture F.E.s		Yes		Yes		Yes		Yes
Observations	910	910	913	913	540	540	541	541
R^2	0.273	0.526	0.211	0.360	0.252	0.527	0.190	0.373

Note: T-values in parentheses: *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.
For some counties, the outcome information in 1998 in related industrial outcome indicators (investment and output) is 0. This means we were not able to calculate the growth rate, which led to a reduced sample size. Note that this also applies to Table 2.6.

Second, we run the regression again using the number of toll stations (count data) instead of the highway dummy. The equation is as follows:

$$\ln(Growth_i) = \alpha_0 + \alpha_1 \ln(number\_of\_toll\_station_i) + \alpha_2 Propensity\_Score_i + \alpha_3 Xianjishi_i + \alpha_4 Initial_i + \eta_i + \epsilon_i. \tag{2.6}$$

In Table 2.5 we present the number of toll stations for our sample counties in 2011. Since the related data from 2005 is not available, we use the data from 2011 as a proxy for that of 2005. We assume a proportional growth in the number of toll stations for the sample counties within the highway county group. However, this assumption is not reasonable for the full sample because, for one highway line, counties connected in the first period will not be connected again in next period. Within the highway county samples, or within the planned highway and control county samples, we assume that the probability of receiving new toll stations in the period 2006–2011 is proportional to

Table 2.5: Number of highway toll stations in sample counties

Number of toll stations	0	1	2	3	4	5	6	7	>7
Number of highway counties	0	131	94	82	41	20	15	8	7
Number of planned highway counties	0	28	44	15	8	5	3	0	2
Number of control counties	480	0	0	0	0	0	0	0	0

the number of existing toll stations in 2005. Actually, for planned highway and control county samples, they have no highway toll station before 2006. For most of the highway counties and planned highway counties have 1–3 toll stations in 2011.

Table 6 presents our modified estimations. Unsurprisingly, highway density (number of toll stations) has a positive relationship with industrial outcomes within the highway county sample (columns 1–4), with the exception of column 2. After controlling for prefecture fixed effects, the effects of highway density on growth in industrial output become nonsignificant. In the robustness check, using only planned highway counties (columns 5–8, the results of the counterfactual highway effects are nonsignificant, which is consistent with previous estimations.

Table 2.6: OLS regression with propensity score and count data of toll stations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ln (Growth in output)		ln (Growth in investment)		ln (Growth in output)		ln (Growth in investment)	
ln (number of toll stations)	0.113**	0.0193	0.147**	0.122*	0.015	-0.131	0.276	0.288
	(2.01)	(0.32)	(2.45)	(1.74)	(0.11)	(-0.64)	(1.34)	(0.88)
Propensity Score	0.941**	0.363	1.059***	0.499	1.716	0.882	2.510*	3.738
	(2.34)	(0.66)	(2.66)	(0.82)	(1.63)	(0.42)	(1.72)	(1.28)
Xianjishi	0.176*	0.177*	0.084	0.202	0.190	-0.127	0.145	0.017
	(1.97)	(1.68)	(0.86)	(1.64)	(0.77)	(-0.31)	(0.38)	(0.03)
ln (output initial)	-0.488***	-0.425***			-0.361***	-0.224		
	(-10.18)	(-7.49)			(-3.56)	(-1.30)		
ln (invest. initial)			-0.419***	-0.468***			-0.592***	-0.616***
			(-8.45)	(-7.47)			(-5.44)	(-3.40)
Constant	6.589***	6.930***	5.533***	7.164***	3.697***	4.328***	6.121***	5.777**
	(13.63)	(9.87)	(10.03)	(8.47)	(2.98)	(3.06)	(3.67)	(2.15)
Sample	Highway	Highway	Highway	Highway	Planned	Planned	Planned	Planned
Province F.E.s	Yes		Yes		Yes		Yes	
Prefecture F.E.s		Yes		Yes		Yes		Yes
Observations	370	370	372	372	96	96	96	96
R^2	0.381	0.662	0.306	0.565	0.471	0.708	0.383	0.625

Note: T-values in parentheses: *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

2.5 Further estimates and discussion

2.5.1 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 a significant centripetal force. Thus, the closer a city is to a central city, the faster the city grows. Xu *et al.* (2010) estimate that the turning point for Chinese cities is 300 km. When the distance exceeds 300 km, a centrifugal force 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 access plays a role in drawing connected cities closer because the average transport efficiency of a highway is much higher than that of an ordinary road. Thus, we can expect that highways can help bring cities closer, at least in the economic sense, to large cities. Therefore, we check whether highway effects differ between counties with different distances to large cities.

From Table A.1, we can see that in the coastal regions, the average minimum distance between a highway county and a large city is less than 200 km (see the "minimum distance to large cities"), with most highway counties being close to large cities. The average minimum distance for inland counties is 312 km, which is similar to the distance (300 km) mentioned by Xu *et al.* (2010). As such, it is suitable to use only inland samples to check for highway effects in different geographical locations. We divide the counties into two groups, according to geographical location: close to or remote from large cities. The cutoff is set at 300 km, and the control counties in the inland regions are also divided into the same two groups using the same criterion. The related propensity score distribution shows that the propensity score still performs well in predicting highway placement (Figure A.4).

The DID-PSM results are presented in Table 2.7. The highway treatment effect is not evident in those counties that are remote from large cities, as shown in columns 1–4. However, the effect is highly positive and significant in counties close to large cities. Evidently, counties close to large cities are more active in industrial activities. Columns 7 and 8 show that most industrial investment and output are nonsignificant before treatment, but become significant after treatment. In addition, we still find no evidence of highway effects on per capita GDP and firm density.

The centripetal and centrifugal forces mentioned by Xu *et al.* (2010) could explain the results in Table 2.7, although there may also be some other underlying logic. Highway construction always starts in large cities and then gradually extends to the peripheral regions. That is, counties close to large cities tend to gain highway access sooner than those that are more remote. Thus, in highway counties that are distant from large cities, industrial performance needs more time to respond to highway access. However, this explanation does not weaken the robustness of the estimated results for the *planned highway counties* (Table 2.3). This is because the average minimum distance to large cities among the inland planned highway counties is 306 km (see column 6 of Table A.1),

Table 2.7: Estimation of highway treatment effects in different geographical locations

	(1) Growth in p.c. GDP	(2) Growth in firm density	(3) Growth in investment	(4) Growth in output	(5) Growth in p.c. GDP	(6) Growth in firm density	(7) Growth in investment	(8) Growth in output
Panel A: Highway effects (difference-in-difference)								
Highway	0.0084	-0.4072	0.3226	0.1473	-0.0207	0.1525	1.1681**	0.3183*
	(0.08)	(-1.43)	(1.12)	(0.77)	(-0.24)	(0.86)	(2.42)	(1.81)
Panel B: Pre-treatment difference								
Highway	0.0276	0.1831	0.0958	0.2352*	0.0017	-0.0827	-0.1385	-0.0419
	(0.54)	(1.08)	(0.43)	(1.71)	(0.03)	(-0.44)	(-0.49)	(-0.18)
Description								
Region	Inland, Remote				Inland, Close			
No. matched obs. (Treated/Control)	133/208				98/171			
Pseudo R^2 (Raw/Matched)	0.223/0.025				0.188/0.020			

Note: Samples that do not satisfy the *common support* are dropped; Z-values in parentheses: *, ** indicate significance at the 10% and 5% levels, respectively.

which is less than that of highway counties (312 km). These *nonsignificant* estimates are due to the absence of highway access before 2005, not the distance from large cities.

2.5.2 Highway effects on industry structure

Here, we examine the relationship between within-industry heterogeneity and highway effects. The CISD classifies each firm-level record as either heavy industry or light industry.[5] Based on this classification, Table 2.8 shows that there is a significant difference in highway effects, even in terms of the internal structure of an industry. Heavy industry is heavily affected by highway access, but light industry is not. We find that highway connections promote heavy industry in the inland counties (here we refer only to the "close" samples, as described in Section 5.1, because the highway effects in "remote" counties are weak), but that the effect of these connections on light industry is unclear. For firm density, the results are all nonsignificant (the estimation results for coastal counties are consistent with those of inland counties; results are available upon request).

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 accounted for 49% of all firms and 55% of the value-added of the national industry. By 2007, these figures had increased to 57% and 68%, respectively (China Statistical Yearbooks, 1999 & 2008). This suggests that, in recent years, heavy industry has been more active than light industry in China, and that the average firm size within heavy industry is larger than in light industry. This could partly explain why the estimated growth of firm density in highway counties is not as significant as the growth in investment and output. In addition, heavy industry

---

[5] Heavy industry refers to chemistry, general equipment, the steel industry, etc.; light industry refers to textiles, food, the tobacco industry, etc.

Table 2.8: Estimation of highway treatment effects within different industries

	(1) Growth in firm density	(2) Growth in investment	(3) Growth in output	(4) Growth in firm density	(5) Growth in investment	(6) Growth in output
	Panel A: Highway effects (difference-in-difference)					
Highway	0.1276	0.1561	-0.0265	0.1647	1.6352**	0.5383**
	(0.49)	(0.56)	(-0.12)	(0.91)	(2.36)	(2.30)
	Panel B: Pre-treatment difference					
Highway	-0.1569	-0.0066	-0.1086	0.0003	-0.1873	0.0121
	(-0.74)	(-0.03)	(-0.43)	(0.00)	(-0.52)	(0.04)
	Description					
Industry	Light			Heavy		
Region	Inland, Close					
No. matched obs. (Treated/Control)	98/171					
Pseudo R^2 (Raw/Matched)	0.188/0.020					

Note: Z-values in parentheses: ** indicates significance at the 5% level.

is much more dependent on transport than light industry, and so will thus benefit more from highway access.

Next, we estimate labor productivity and employment within industry. The indicator for labor productivity in industry is nonsignificant, but is significant in terms of employment (see Table 2.9). The significance of the treatment effects in employment could potentially explain why highway counties have a higher industrial output and level of investment, while at the same time, firm density does not increase. In other words, the average firm size in highway counties tends to be larger than in control counties. The nonsignificant results for labor productivity seem to be consistent with the nonsignificant results for per capita GDP. Two factors may help interpret these findings. First, the CISD records firm-level data according to the firm's geographical location. However, in some cases, a firm's output may be attributed a higher administration unit rather than to the local economy. 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 has been observed in many prior studies. Using the growth (for the period 1998–2007) of per capita GDP as the dependent variable and per capita GDP in 1998 as the independent variable, an OLS estimate using the samples from this study gives a significant and negative coefficient under different settings. Using the industrial structure, we also test whether highway access promotes upgrading the industrial structure from that suited to a primary sector-driven economy to that suited to a secondary sector-driven economy. However, the results are unclear (results are available upon request).

Table 2.9: Estimations using labor productivity and employment

	(1) Lab. Prod.	(2) Empl.	(3) Lab. Prod.	(4) Empl.
	Highway effects (Post-treatment difference)			
Highway	0.0323	0.4014***	-0.0059	0.2397**
	(0.47)	(2.69)	(-0.07)	(2.11)
	Description			
Region	Coastal, All		Inland, Close	
No. matched obs. (Treated/Control)	150/99		98/170	
Pseudo R^2 (Raw/Matched)	0.069/0.010		0.182/0.020	

Note: Owing to data restrictions, we do not present the pre-treatment indicators for labor productivity and employment within the industry (these two indicators are only available for firms in the CISD).
Z-values in parentheses: **, *** indicate significance at the 5% and 1% levels, respectively.
The description of the covariates is as follows:
Lab. Prod. (Labor productivity): per capita value-added (value-added divided by the number of employees in firms above the designated size);
Empl. (Employment): average number of employees in industrial firms above the designated size.

2.5.3 Net spillover effects or substitution effects?

As stated in Boarnet (1998), infrastructure investment has positive output spillover effects within the same county and negative output spillover effects in other counties. This finding is relevant to the results of our identification strategy. Counties within a prefecture have similar industrial policies and geographical proximity. As a result, for highway connected counties, our estimated highway effects might stem from a county being used as a substitute for its unconnected neighbors. In this case, the presence of a highway does not promote industrial development in rural regions, instead having a negative impact on unconnected counties. For a planned highway county or a control county, if other counties within a prefecture are connected to a highway in the period 1998–2005, we refer to the planned highway county or control county as being spillovered by its neighboring highway counties. We thus test the potential negative effects (i.e., the substitution effects) of highways on unconnected counties by comparing two kinds of counties: (1) a county that did not receive a highway in 1998–2005, but within its prefecture, at least one other county did receive a highway during this period (spillovered); and (2) a county that did not receive a highway in the period 1998–2005, and no other counties within its prefecture received highways either (unspillovered). In this estimation we employ only the planned highway counties and control counties, and treat them as equivalent. The estimation equation is as follows:

$$\ln(Growth_i) = \alpha_0 + \alpha_1 Spillover + \eta_i + \epsilon_i. \tag{2.7}$$

Here, *Spillover* refers to the county dummy (takes the value 1 for a county if there is at least one county within the prefecture that is connected to a highway in 1998–2005, and 0 otherwise). The results are presented in columns 1 and 3 of Table 2.10. The difference between spillovered and unspillovered counties is nonsignificant. Then we employ another indicator instead of *Spillover*, namely the density of spillover, represented as *number of highway neighbors*. Here, *number of highway neighbors* refers to the number of highway counties within the same prefecture as the spillovered planned highway or control county.[6] The results in columns 2 and 4 remain unchanged. Furthermore, in columns 5–8, we insert other controls for the *Xianjishi* dummy, namely the propensity score and initial values. In both cases, the results are nonsignificant. This suggests that the spillover effects from a highway county to neighboring counties, whether positive or negative, are weak and unclear.[7]

Table 2.10: Estimation of the source of highway spillover effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ln (Growth in output)		ln (Growth in invest.)		ln (Growth in output)		ln (Growth in invest.)	
Spillover	-0.00		0.04		-0.04		-0.02	
	(-0.01)		(0.31)		(-0.37)		(-0.20)	
ln (number of highway neighbors+1)		0.03		0.06		-0.02		-0.00
		(0.29)		(0.62)		(-0.20)		(-0.04)
Propensity Score					1.99***	1.99***	1.37***	1.36***
					(4.76)	(4.73)	(3.13)	(3.10)
Xianjishi					0.05	0.05	0.12	0.12
					(0.40)	(0.39)	(0.82)	(0.82)
ln (output initial)					-0.40***	-0.40***		
					(-9.74)	(-9.72)		
ln (invest. initial)							-0.40***	-0.40***
							(-9.15)	(-9.12)
Constant	1.64***	1.60***	1.20***	1.14***	5.14***	5.12***	5.35***	5.33***
	(7.88)	(7.27)	(5.33)	(4.82)	(11.04)	(10.96)	(9.95)	(9.85)
Sample	Planned& Control	Planned& Control	Planned& Control	Planned& Control	Planned& Control	Planned& Control	Planned& Control	Planned& Control
Province F.E.s	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	540	540	541	541	540	540	541	541
R^2	0.110	0.110	0.054	0.054	0.252	0.252	0.190	0.190

Note: T-values in parentheses: *** indicates significance at the 1% level.

2.5.4 Highway effects or railway effects?

Railways could be another potential disturbance of the highway effects on industry performance, because they are also an important transport infrastructure. The following facts support our results. First, the number of freight stations in the railway network

---

[6] For the unspillovered counties, *number of highway neighbors* is 0. As a result, we use ln (*number of highway neighbors*+1) in the estimation.

[7] Here, we do not consider the highway situation in a core district of the prefecture because it is symmetrical. That is, in the geographical sense, if a highway goes through a core district of a prefecture, it must pass through one or more counties within this prefecture (in most cases). In addition, we calculate the geographical distance from the toll station to a county to measure the spillover effects more precisely, because some cross-prefecture counties may be spillovered by highways in other prefectures. However, since we do not know the exact year a toll station was built, the geographical distance is unable to reflect highway accessibility for the overall period from 1998 to 2007. For example, if a county is relatively close to a highway toll station, but did not receive its own toll station until 2005, the highway effects will be weak, even though the highway is highly accessible. Second, there are cross-prefecture political barriers within industries, so we believe it is reasonable to assume that substitution effects only occur within a prefecture.

did not increase in the past decade. Most were built before 2000, and some even before 1960. The limited number of 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 accounted for 83.2% of the total railway freight ton-kilometers and 91.2% of the total railway freight traffic in 2007. That is, railways mainly carry raw materials, and not industrial products. Hence, we expect that the road network plays a larger role in shipping industrial products and so affects the location of an industrial firm. Third, an increasing number of manufacturing firms are located along highways all over China.

2.6 Conclusion

By employing a micro-level data set, we estimated the effects of highways on industrial development in rural China. We found that highway access promotes industrial development in counties with higher output and greater levels of industrial investment. However, highways do not help remote rural counties, which are counties located more than 300 km from a large city. Furthermore, at the county level, highways tend to promote the development of heavy industry more than light industry. Moreover, highway access does not improve industrial productivity or per capita GDP. We use various robustness checks to ensure and improve the precision of our findings. We also provided potential explanations for the nonsignificant highway effects on industrial productivity and per capita GDP. In addition, we investigated the source of the estimated highway effects, and show that they stem from a net spillover from building the infrastructure, and not because a county is a substitute for its unconnected neighbors. An interesting extension of this research may be to examine highway effects in conjunction with urbanization.

Chapter 3

Speeding-Up of Railways and Exports: Evidence from Silk Road[1]

3.1 Introduction

Improvement of domestic transport infrastructure is important to international trade, since domestic freight costs account for a considerable proportion of the total trade cost, particularly in large countries. In fact, it has been shown that domestic freight costs accounted for 36.6% of international transport costs in the United States a few decades ago (Rousslang and To, 1993). Transport burdens are magnified for firms in interior cities of China seeking to import low- or medium-value inputs from the coast (World Bank, 2006). These suggest that domestic transport costs can be an important barrier to trade[2]. This feature is strengthened in China-Central Asia (Kazakhstan, Kyrgyzstan, Tajikistan, Uzbekistan, and Turkmenistan) trade. First, China has a lot of landlocked territory in the western regions, while its production mainly agglomerates in the eastern regions. As a result, to move goods to Central Asia (west of China) (see in Figure 3.1),

---

[1] I would like to thank Kentaro Nakajima and Takashi Kurosaki for their careful reading and stimulating comments on this research. In addition, I would like to thank Dao-Zhi Zeng, Ruixiang Xiong, Frédéric Robert-Nicoud, David McKenzie, Ben Zissimos, Kishore Gawande, Jason T. Kerwin, Sandra Poncet, Zhao Chen, Zhigang Li, Ching-mu Chen, and the participants at the GSIS Seminar (Tohoku University), GRIPS, Nagoya University, ABCDE 2014 (World Bank), the Third RUSE Workshop (Fudan University), the JEA Spring Meeting (Doshisha University), the European meeting of UEA (St. Petersburg) and ARSC meeting (Naha) for the beneficial discussions. I am grateful for financial support from the Japan Society for the Promotion of Science (DC2: 26-4705). All errors and shortcomings are my own.

[2] Except trade, Emran and Hou (2013) find the impact of domestic market access is significantly large on per capita consumption of households in rural China, and the probability that rural households participate in non-farm activities.

Figure 3.1: The geographic proximity of China and Central Asia
Note: Alashankou and kashi are the two main frontier ports for China-Central Asia trade

it is necessary to use long-distance, domestic land transport by either train or truck[3]. To reach these five countries, goods must go through the ancient Silk Road[4], that is, first reach Urumqi, the capital of Xinjiang, and then exit China through frontier ports. To go through the Silk Road, the transportation will pass through a large, unpopulated area (desert), because of which, this route is unique (for both rail and road).

For example, for goods shipped from Jiangsu (a province of Eastern China) and exported to Central Asia, the domestic transportation distance is more than 4,000 km (Jiangsu-Urumqi is 3,600 km and Urumqi-Alashankou (one of the main frontier ports to Central Asia) is 477 km). As a result, due to the long domestic distance, many goods are not profitable to export from Eastern China to Central Asia. Domestic transportation proves to be too expensive and time-consuming, although eastern regions in China have a much higher productivity when compared with that of western regions because of a huge market, rich labor supply, and agglomeration.

This situation was, to some extent, relieved in 2000. Along with the progress of China's railway speed-up project (1997-2007), in the third round of this national railway speed-up project on Oct. 21, 2000, the Longhai (Lianyungang-Lanzhou) and the Lanxin lines (Lanzhou-Urumqi) were sped-up simultaneously within a day. Travel time between Xinjiang and Eastern China was significantly lowered, and train capacity was greatly increased due to higher shift frequency. In addition, the freight cargo to the western regions was given priority in cargo loading, because of the policy-oriented China's Go

---

[3] There is no river for shipment, and airfreight exports are rare.

[4] Trade on the Silk Road was a significant factor in the development of the civilizations of China, the Indian subcontinent, Persia, Europe, and Arabia. It opened long-distance political and economic interactions between the civilizations (Bentley, 1993). The Chinese portion of the Silk Road is a long distance land route, from Xi'an to Alashankou, a straight valley that penetrates the Dzungarian Alatau mountain range, along the border between Kazakhstan and Xinjiang.

West Campaign (People's Daily, 2000). This means the efficiency of long-haul domestic transportation in China-Central Asia trade was greatly improved[5].

Thus, this research evaluates the impact of the railway speed-up on exports from China to Central Asia. I set this speed-up project as a quasi-experiment to identify the role of domestic market accessibility in firms' export performance. The railway speed-up project (1997-2007), as a long-term national development strategy mainly for promoting regional integration in the country, is not subject to the performance of China-Central Asia trade, since the related trade volume is small compared to China's overall national trade. This clears out the endogeneity concern of the speed-up being related to China-Central Asia exports. Depending on the different freight modes, speed-up improves the domestic accessibility to railway freight exports (the treatment group) while for others (the control group), such as roads, air, etc., it remains the same. By contrasting exports in both groups before and after the speed-up, and controlling for potential confounding factors, I can consistently estimate the impact of the railway speed-up project on firms' exports. Precisely, I perform difference-in-differences estimations on firm transaction-level data that contains information about the exact geographical origin of the exports, and the freight modes through which the products exit China.

There is a concern that the railway speed-up effects to exports are heterogeneous based on points of origin. Exporters enjoying more with sped-up mileage benefit more on cost and time saving. By restricting the railway freight export samples, I test the relation between sped-up mileage variation and export performance. That is, I set exports through rail and gain long-distance speed-up as the treatment group and exports through rail but with limited sped-up mileage as the control group. Then, I carefully address the robustness concern by excluding specific samples (with respect to demand fluctuation of specific products and macroeconomic shocks, like the "Sept. 11 Attacks") and conducting counter-factual analysis with data from other years.

I find that railway speed-up increases annual export values to Central-Asia by around 30% at firm-product-destination country level, compared with exports through other freight modes, and compared with rail freight exports, but enjoy limited speed-up benefits. A positive impact on exporters' shipment frequency is also significant but much smaller, suggesting that the speed-up helps exports by expanding the scale of export per shipment. Consistently, results show that the intensive margin effect, but not extensive margin, plays a major role in explaining the impact of the speed-up on exports.

Export creation effects from improving transport efficiency are the dominant source of the speed-up effect. In addition, I find two kinds of diversion effects are mixed into the

---

[5] Throughout this chapter, speed-up refers to the speed-up project (containing increased speed, expanded capacity, and upgraded operation modes, etc.), not only to the increase in speed.

export creation effect. To some extent, some exports once freighted by trucks tend to be transferred to trains because of improving efficiency in railways. And, Xinjiang exporters (Xinjiang is the gateway from China to Central Asia, being close to Central Asia. It enjoys limited speed-up in Oct. 2000) lose some of the Central Asian market to exporters in other regions (far from Central Asia; competitive in manufacturing productivity), which obtain better domestic accessibility to Central Asia by the speed-up. This is in line with the core-peripheral effects of trade integration founded in increasing returns to scale trade theory and economic geography (Helpman and Krugman, 1985; Fujita et al., 1999). That is, higher efficient transportation (e.g., railway speed-up, expressway construction) could be associated to a domestic process of integration in manufacturing location (e.g., Faber, 2014), also, the location of exporters. Customs data also shows exports of medium-value products are the most sensitive to speed-up. For low-value products, exports are not affected by the speed-up because they are suitable for leisurely shipping, not elastic to change in transportation efficiency.

Additionally, I test the export performance with respect to speed-up in another direction, exports from Xinjiang exporters to other markets, like Japan and US. In this case, products must experience long-haul freight from Xinjiang to the eastern coastal lines. If products are moved by rail, then they use almost the full length of both the Lanxin and the Longhai lines. By checking the export performance of Xinjiang in different markets, I find the locational advantage of Xinjiang becomes weaker in China-Central Asia export because of the railway speed-up, while Xinjiang gains better accessibility to other international markets. The role of speed-up in the regional development of Xinjiang is two-fold.

My analysis might contribute to two different bodies of literature. First, I provide evidence on the role of domestic accessibility in international trade participation, and investigate the channels of domestic transportation improvement effects on export from different angles, for which both the empirical and theoretical literature are far from adequate (see e.g., Ago et al., 2006; Albarrán et al., 2013; Martincus and Blyde, 2013; Faber, 2014). The most difficult aspect of empirics is that, to identify such effects requires the successful addressing of potential endogeneity problems affecting the relationship between domestic transport costs, as determined by infrastructure, and trade. In this research, I first combine micro firm transaction-level export data and exploit the railway speed-up project as an exogenous source of variation in transport efficiency. Second, I provide evidence on the externalities of large-scale transport infrastructure to specific regions. Xinjiang, as the gateway from China to Central Asia, suffers the loss of locational advantage because of improving domestic accessibility (similar theoretical results were suggested in Mori (2012)), and on the other hand, improved domestic accessibility to the coast promotes its export to countries like the US and Japan. There are two-sided externalities to the investment in transport infrastructure on the development of specific regions.

The rest of this chapter is organized as follows. Section 3.2 contains the background of the China railway speed-up project. In Sections 3.3 and 3.4, I present the empirical strategy and results, respectively. Section 3.5 discusses the channels and heterogeneity of railway speed-up effects on export, and Section 3.6 is a simplified evaluation of overall export impacts on Xinjiang. Section 3.7 concludes.

3.2 The China railway speed-up project

The China railway speed-up project refers to initiatives undertaken by the Ministry of Railways, from 1997 to 2007, to increase the speed and capacity of train travel in China by improving the nation's railways. The project was implemented in six rounds, and increased the average speed of passenger trains in China from 48.1 km/h to 70.2 km/h (there is no official statement regarding the degree of improvements in the average speed of freight trains). Although the railway speed-up does not directly decrease monetary transport costs, it decreases the time costs through higher speed and more frequent shifts. In addition, the speed-up project improves China's railway system by adopting new and advanced trains and modified operation modes.

The overall project was undertaken in six sub-rounds and only selected railway lines were sped-up. The treated railway lines are sped-up at specific dates. Detailed information of these six rounds is shown in Table 3.1. In the first two rounds, treated lines were the three capital lines from Beijing to Southern, Eastern, and Northeastern China, respectively. Then in 2000, in the third round, three major west-east lines were sped-up. Two of them are the concern of this research, connecting Xinjiang and Eastern China (the Lanxin and the Longhai lines). Another west-east line is the Zhegan line, which is much shorter than the Lanxin and the Longhai lines, and not directly related to Xinjiang. In addition, the Jingjiu line (Beijing-Kowloon), a capital line from Beijing to Southern China was sped-up in 2000. Then in 2001, 2004, and 2007, many other lines were sped-up successively, and some lines were sped-up once more.

Table 3.1: Schedule of the railway speed-up project

Date	Sped-up Lines
Apr. 1, 1997	Jingguang, Jinghu, Jingha
Oct. 1, 1998	Jingguang, Jinghu, Jingha
Oct. 21, 2000	Longhai, Lanxin, Zhegan, Jingjiu
Oct. 21, 2001	Jingguang, Jingjiu, Zhegan, Huhang, Hada
Apr. 18, 2004	Jinghu, Jingha
Apr. 18, 2007	Jingha, Jingguang, Jinghu, Jingjiu, Longhai, Lanxin, Zhegan

The achievement of the speed-up project (1997-2007) is significant, as shown in Table 3.2. In this decade, the highest rail speed increased from less than 140 km/h to 250 km/h, and more passengers can reach their destinations (if less than 1,500 km) within one night. For three kinds of freight trains (five-set freight trains[6], bulk cargo trains, and parcel trains), shift frequency was largely increased. For the speed-up in Oct. 2000, a considerable number of new five-set freight trains, as well as bulk cargo trains, were applied on sped-up lines, suggesting capacity was expansed. There is no exact speed data on freight trains, for the passenger trains, around 20% of the travel time was saved for trips from Urumqi to Lanzhou (from 33 to 27 hours), Beijing (from 63 to 50 hours), and Shanghai (from 66 to 51 hours) (Sources: China's Train Time Table (*Lieche Shike Biao*), 2000, 2001). For both capacity and speed, it is obvious that they were greatly improved because of the speed-up project.

TABLE 3.2: Achievements of the railway speed-up project

Panel A: Overall achievements of the speed-up project (1997-2007)

	Before 1997	1997	1998	2000	2001	2004	2007
Highest speed (km/hr)	n.a.	140	160	n.a.	n.a.	200	250
Five-set freight trains (no. of lines)	0	40	40	71	79	92	121
Bulk cargo trains (no. of lines)	n.a.	n.a.	119	138	157	180	406
Parcel trains (no. of paired lines)	0	0	8	14	15	19	n.a.
Average speed of passenger trains (km/hr)	48.1	54.9	55.2	60.3	61.6	65.7	70.2
Passenger lines starting at daybreak and arriving at sunset (no. of lines)	n.a.	78	228	266	n.a.	305	337

Panel B: Travel time changes between major cities due to the speed-up on Oct. 21, 2000

	pre-speed-up	post-speed-up	time saved
Urumqi-Lanzhou (1,892 km)	33 hrs	27 hrs	18 pct
Urumqi-Beijing (3,216 km)	63 hrs	50 hrs	21 pct
Urumqi-Shanghai (4,077 km)	66 hrs	51 hrs	23 pct

Sources: Various, collated by author.

As shown in Figure 3.2, the Longhai (Lianyungang-Lanzhou) and the Lanxin (Lanzhou-Urumqi) lines are the main routes from Eastern China to Xinjiang, as well as Central Asia. Total mileage of these two lines is 3,651 km, accounting for a large part of the overall China-Central Asia transportation network. Compared to the high density of China's railway network, most of the railways, except the Longhai and the Lanxin lines, were not to be sped-up in this round.

---

[6] Five-set freight trains: freight trains with scheduled stations, train numbers, routes, times, and prices, with simplified freight procedures, reasonable prices, and guaranteed time, capacity, and cargo security. It is a new freight service of the China Railway Group, available since 1997 in response to the decreasing market share. The China Railway Group operates China's railway freight under bureaucracy and complete monopoly, its passive market-oriented reform is argued to be a main reason for its low freight efficiency and decreasing freight market share. Compared to truck transportation, capacity availability and arrival times are not well set in the railway freight system. Train shippers do not know the exact train and date the goods are loaded, and the exact date of arrival, when they order a freight service. Cargo capacity for a specific train, at a specific station and specific date is not set, and dependent on the usage rate of cargo capacity at previous stations, cargo demands at current station, etc., that is, in freight season, the travel time and capacity cannot be guaranteed. As a result, trucks have become the premium mode for land freight in China.

Figure 3.2: China's railway network in 2009

Note: The Lanxin (Lanzhou-Urumqi) and the Longhai (Lianyungang-Lanzhou) lines, marked in bold lines, are sped-up on Oct. 2000.

3.3 Empirical strategy

Prior to empirical strategy, I provide direct evidences at the country level that the speed-up in specific railway lines indeed affected the related railway traffic, which suggests that the empirical studies in this research are advisable. Then, I address the exogeneity concern of the railway speed-up project to China-Central Asia export, and introduce data and quasi-experimental settings for the main estimations. Precisely, baseline estimations consist of two kinds of quasi-experimental approaches.

3.3.1 Speed-up effects on national railway traffic

To clearly tell the story that the railway speed-up projects affected exports, it is necessary to provide more direct quantitative evidence that the national railway speed-up project indeed affected the regional transportation in China. To do so, I conduct a panel data estimation proving this argument. Among all the national railway speed-up projects in 1997-2007, some railway lines got sped-up at least once, while others never got sped-up. Hence, I construct a panel data on the turnover/traffic growth of freight/passenger for China's main railway lines, and test whether or not the railway lines that were sped-up have a higher growth rate (in freight turnover, freight traffic, passenger turnover, and passenger traffic) in the corresponding year (if the speed-up occurred before April: 1997, 2004, 2007, I set that year as treated; if the speed-up occurred in October: 1998, 2000, 2001, I set the following year as treated). The sample is 26 railway lines in 10 years. If railway line i gets speed-up treatments in the year t, then

I set $Treated_{i,t}$=1. I test the speed-up effects with Equation 3.1. *Growth* represents the annual growth rates for the four indicators, and *Treated* is the binary variable, as mentioned above. η and μ are line and year fixed effects.

$$Growth_{i,t} = \alpha_0 + \alpha_1 Treated_{i,t} + \eta_i + \mu_t + \epsilon_{i,t}. \tag{3.1}$$

TABLE 3.3: Railway turnover and traffic, with respect to the speed-up project

	Annual growth rate of			
	freight turnover (1)	freight traffic (2)	passenger turnover (3)	passenger traffic (4)
Treated	0.0660**	0.1080*	0.1135*	-0.0149
	(0.0316)	(0.0577)	(0.0599)	(0.0808)
With two-way fixed effects				
Observations	235	236	189	186
R-squared	0.0139	0.0233	0.0128	0.0001

Notes: Treated is a binary variable, but not a typical dummy variable. It is set as 1 for the lines in the year of the sped-up. Some lines are exclusively devoted to transport of cargo, so the sample size for passenger turnover/traffic is relatively small. For column 1, the average annual growth rate for the full sample is 6.5%, suggesting growth rate doubled for treated lines in the sped-up year, reaching 13.1%. For freight traffic (column 2), average annual growth rate is 5.6%, and that for sped-up lines reaches 16.4%, on average.
Source: China Railway Yearbook 1999-2008.
Standard errors in parentheses
*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Results in Table 3.3 are suggestive. Considering freight turnover and traffic growth, both of them present significant promotion effects in the treated samples and years, achieving 6.6% and 10.8% higher than average annual growth rate, respectively, which are rather considerable. For column 1, the average annual growth rate of the full sample is 6.5%, suggesting the growth rate doubled for treated lines in sped-up years, reaching 13.1%. For freight traffic, average annual growth rate is 5.6%, and sped-up lines reached 16.4% on average. Passenger turnover was also promoted, although it is not related to export issues here. Passenger traffic was not affected significantly. Quantitatively, the railway speed-up effects on railway transportation are presented, which suggest that the following tests are reasonable.

3.3.2 Exogeneity concern

Prior to the quasi-experimental setting, I emphasize the exogenous concern of the speed-up projects on China-Central Asia exports. The speed-up in 2000 was the third round of the entire speed-up project, with six total rounds. The entire project was mainly considered as a national long-term development strategy to promote regional integration, population (goods) transportation between inland and coastal regions, and guaranteeing national defense. It is reasonable to assume that the China-Central Asia trade was not a determinant of the speed-up project. Actually, trade with Central Asia is limited, compared with China's total international trade. In 2000, China's total exports were 249.71 billion USD, while the export to Central Asia was 0.77 billion USD, accounting for only 0.31% of the national export (Source: China External Economics Statistical Yearbook 2001).

Then, speed-up promotes the railways, but not other freight modes like roads, shipping, and air. In 2000, the Lianhuo expressway (Eastern China-Xinjiang) had not yet been constructed. In addition, in this speed-up round, few railway lines were sped-up, except these two lines, which guaranteed that other railway routes were unchanged. Moreover, unlike the construction of roads, this railway speed-up was finished simultaneously within one day and not gradually, so the impact was more direct and observable. The public noticed the third round speed-up before Oct. 21, 2000; as a result, it induced an expectation, causing it not an absolute exogenous shock, like a natural disaster. Export, to some extent, is time-sensitive and subject to contract obligations, that is, exporters cannot put off the shipping to meet the speed-up project. So one can suppose that the overall bilateral trade (annual level) will not be much affected by the expectation of the speed-up project. Further, exporters participate in Central Asia market might not export only to Central Asia, in most of the cases, they will export also to other international markets. In this case the benefit of firm re-location to obtain access of sped-up railways is not that large, since the speed-up in 2000 only benefits four railway lines, for the freight to other regions, except Xinjiang, the freight efficiency is not necessary to be improved. Thus we suppose the endogenous problem of exports response to railway speed-up (through firm re-location) is minor.

In addition, unlike most international trade, China-Central Asia trade is more dependent on rail and road freight, since there is no river and sea for shipping, and the domestic transportation distance is rather long, representing the transportation efficiency (time and monetary costs, and business risks) is considerable in total trade cost.

Also, the macro economic environment in Central Asia will affect the demand of Chinese products in Central Asia market, which potentially affects the effectiveness of my quasi-experiment estimation. Table B.1 and B.2 give the descriptive summary of Central Asia economy, as well as the rail transport infrastructure improvements in Central Asia. First, Table B.1 suggests the the macro social economy of Central Asia is stable in the periods of railway speed-up (Oct., 2000) in China. In the periods of 1996-2005, GDP per capita in all of these fives countries grows rapidly. Population size of Kazakhstan decreases by 3% (0.5 million) in this period, while the rest four countries have stable increase. Imports of Kazakhstan, as the largest economy in Central Asia, increase greatly by 78%

in this decade. Turkmenistan and Tajikistan get doubled in their import scale, although their economic sizes are much smaller than Kazakhstan and Uzbekistan. However, in the much shorter period, 2000-2001, one year before and one year after the speed-up (Oct., 2000) in China, import value in all these five countries are stable. I also present the export value, and share of manufactures imports in total merchandise imports, both are stable in the study period, particularly in 2000-2001. The stable external economic environment ensures trade fluctuation from demand shock might be avoided. Second, Table B.2 gives the descriptive statistics in transport infrastructure. Obviously, there is no big push on the transport infrastructure in the period closing to 2000 (rail speed data in Central Asia is not available.). These natural settings allow a quasi-experiment approach for comparing the performance of China-Central Asia export between benefited and other exporters, before and after the project.

3.3.3 Quasi-experimental settings and data

I conduct two kinds of quasi-experimental approaches estimating the railway speed-up effects on export. The first approach is to compare the difference in export performance to Central Asia between 2000 and 2001, and between railway freight products and products with other freight modes[7]. I employ the firm transaction-level customs data from Chinese Longitudinal Firm Trade Transaction Database (CLFTTD), in which I use the firm transaction-level export data showing the unique firm ID code, address of the firm, transaction date, export products (eight-digit Harmonized Commodity Coding System, HS-8), place of manufacture (city level), destination country, freight mode[8], export value, unit value, etc. I extract the records referring to exports to Central-Asia in 2000-2001[9], and I employ only short period data (two years), because customs data is not available before 2000, and long-term export performance tends to be affected by noise, such as road construction.

---

[7] I focus exclusively on the export, but not import from Central Asia to China, because the bilateral trade between China and Central Asia is dominated by the export from China to Central Asia.

[8] The database records freight mode of each transaction when exiting China. I assume the domestic freight mode is equal to that when exiting China. It is reasonable, since the domestic and foreign routes are connected in Alashankou/Kashi, and since both are forms of land transportation, to switch the freight mode incurs a switching cost. Moreover, railway and road freight goods are not homogeneous. Roads are responsible for light industry goods while railways transport heavy industry goods and raw materials. Hence, it is reasonable to assume the freight mode when exiting China is the same as that for domestic freight.

[9] Donation samples (government donations) are excluded. In the transaction records of Central Asia export, the exporting firm and the manufacturer are the same or belong to the same province in more than 95% of the observations. For the rest, they are domestically imported from other regions (places of manufacture) to Xinjiang, and then exported by firms in Xinjiang. I exclude these samples (less than 5%) to ensure credibility in identification.

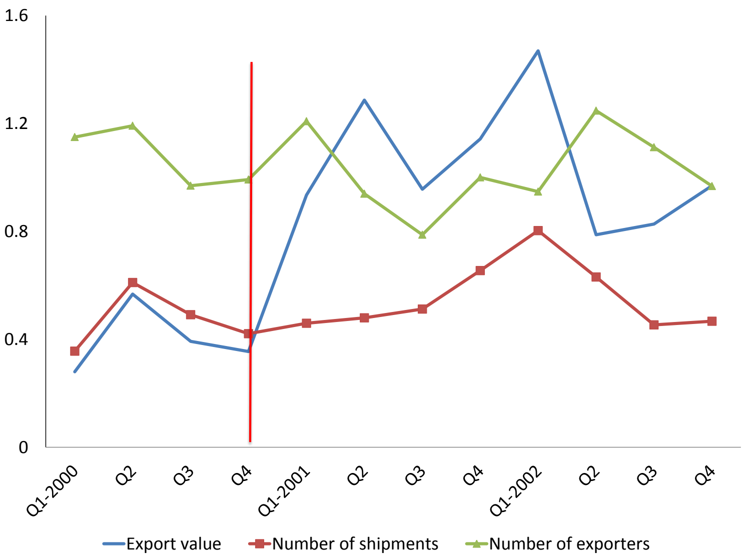


FIGURE 3.3: Ratio of rail/non-rail freight exports from China to Central Asia

Hence, depending on the different freight modes, I assume domestic transport efficiency improved for rail freight exports that enjoyed the speed-up, while that for others remained the same, thus inducing better export performance for rail freight products. In 2000-2001, within the total shipment records from China to Central Asia, 33% of the shipments were taken by railways and 57% were taken by roads, with the rest 10% by air or ocean shipments (indirect shipment that was transferred in Russia), suggesting railway is an important freight mode in this bilateral trade. Figure 3.3 shows that in the four quarters of 2000, the ratios of rail/non-rail freight export in export value, number of shipments, and the number of exporters are stable. However, from the first quarter of 2001, two months after the speed-up in Oct. 2000, the related ratio of export value increases, while for another two indicators, they keep a trend similar to that before the speed-up. Figure 3.4 shows that the export value per shipment and export value per exporter immediately increase after the speed-up for rail freight exports, compared with that of non-rail freight exports.

From Figures 3.3 and 3.4, it is shown that the potential speed-up treatment effect occurs in the first quarter of 2001, that is, 2000 is not affected by the speed-up project. Thus I use annual data, instead of quarterly data, to relieve the seasonal data noise. I aggregate the firm transaction-level data into firm-product (HS-8)-country (destination)-freight mode-year level. In short, I aggregate the transaction data with the same exporter, products (the same HS-8 code), destination country, freight mode (rail, road, air, and ocean shipping), and year. Data description is presented in Table B.3. For a part of aggregated annual data, there is no paired data in the firm-product-country-freight mode-level. New exporters in 2001 have no positive export value in 2000, and vice-versa, for exporters exiting this market. For the unpaired data, I set the corresponding value

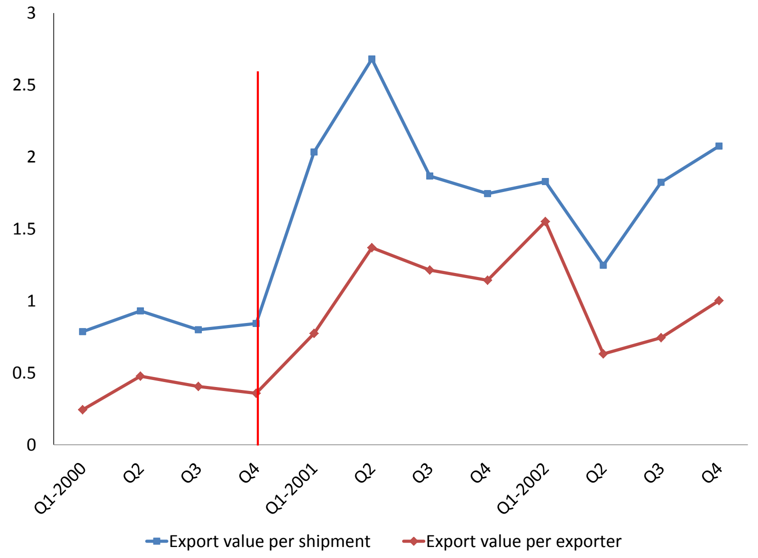


Figure 3.4: Ratio of rail/non-rail freight exports from China to Central Asia (2)

in the related year to 0, which induces a larger sample size in econometric estimation than in data summary. Baseline estimation equation is as follows:

$$\mathit{ExportValue}_{f,p,c,t,y} = \alpha_0 + \alpha_1 \mathit{Rail}_{f,p,c,y} + \alpha_2 \mathit{Rail}_{f,p,c,y} \times \mathit{y2001}_{f,p,c,t} + \alpha_3 \mathit{y2001}_{f,p,c,r} + \epsilon_{f,p,c,t,y}. \tag{3.2}$$

Rail is a dummy variable, identifying the freight mode, where *Rail*=1 if the export transaction is undertaken by railway freight and *Rail*=0 otherwise. *y2001* is a dummy variable for year 2001, with *y2001*=1 for the exports recorded in 2001 and 0 for those in 2000. Subscripts f, p, c, t, and y represent firm, product, country, freight mode, and year, respectively. To eliminate firm-product-country-freight mode (f-p-c-t) level heterogeneity, I use a within estimator. That is, I normalize the export value by subtracting the mean of annual export value (*Ann.Exp.*) in 2000-2001. I set $ExportValue_{f,p,c,t,y}$=$Ann.Exp._{f,p,c,t,y}$-$(Ann.Exp._{f,p,c,t,2000}$+$Ann.Exp._{f,p,c,t,2001})/2$, *y=2000, 2001.*

By controlling for the f-p-c-t level heterogeneity, year, and rail freight dummies, the interaction term $Rail \times y2001$ captures the potential railway speed-up effects. If the related exporting activities are promoted, it should show a positive coefficient.

For the second empirical approach, it is straightforward to suppose that the railway speed-up effects to exports are heterogeneous with respect to the origin. For exporters with long-haul domestic freight, time costs and business risks (unexpected delivery delay or loss) are higher compared to that of short-haul exporters. Speed-up in 2000 saves time for freight from Eastern China to Central Asia. And, thirty-one five-set freight trains were applied after speed-up in this round (see in Table 3.2), suggesting a higher capacity and higher efficiency in west-east transportation. Further, although freight rate is not directly decreased by the speed-up, the actual transportation cost might also decrease. Li and Chen (2013) find the price difference goes to smaller between connected regions in the case of relieved transportation congestion, even if the freight rate is unchanged[10]. These factors suggest that exporters to the Central Asian market, with long-haul domestic transportation, are supposed to benefit more through the speed-up. In the estimation of Equation 3.2, I drop this variation by uniformly setting a binary variable for rail freight exports, which inevitably causes information loss. I hence conduct an estimation fully using the variation on sped-up mileage.

Freight and passenger transportation in rail are operated separately in most cases. Coding systems for freight trains are not tractable, thus, I cannot get the exact sped-up mileage from the route for each freight train. I use data from passenger trains as a proxy to measure the sped-up mileage. Although the routes for freight and passenger trains from the same starting and terminal stations are not necessarily the same, it is reasonable to use as a proxy.

Instead of calculating the exact railway mileage from each firm to frontier ports, for simplicity, I choose the distance from the capital of the province the firms are located in, to Urumqi, to represent the total mileage for each firm to Alashankou and Kashi (all of the exports from China to Central Asia will go through the route from Urumqi to Alashankou or Kashi, and these routes were not sped-up, so I need not consider these routes). I calculate the total railway mileage and travel routes (with respect to the operating rail trips) based on the minimal transportation times (first screening principle) and minimal total mileage (second screening principle) (data is based on the official Train Time Table of 2000). Under this calculation strategy, for example, the total railway mileage from Fuzhou (Capital of the Fujian Province) to Urumqi is 4,555 km, and the trip will be divided into two parts: Fuzhou-Zhengzhou (1,476 km) and Zhengzhou-Urumqi (3,079 km). In this case, I set the sped-up mileage as 3,079 km, since the second part was sped-up in Oct. 2000 and the first part was unchanged. Cities within the Xinjiang Province also enjoy considerable sped-up mileage (up to 500 km) because Xinjiang has a rather large land area. In the calculation of the sped-up mileage, however, I set the location of firms at the provincial level for simplicity, that is, cities within Xinjiang are set as 0 km in sped-up mileage. The details of sped-up mileage are presented in Table B.4[11].

---

[10] When speaking of inefficient capacity and congestion, long-haul transportation is more likely to suffer from these than short-haul transportation. For example, in the peak season, long-haul tickets are always much tighter than short-haul tickets.

[11] When calculating the sped-up mileage, I consider another two lines sped-up in Oct. 21, 2000: the Zhegan line and the Jingjiu lines. The sped-up mileage on these two lines is also included in the calculation of speed-up mileage.

Estimation strategy for this approach is shown in Equation 3.3. I use only the samples of rail freight exports to avoid the disturbances from fluctuations in non-rail freight exports. I set *Mileage* as the logarithm of sped-up mileage (=ln(Sped-up_mileage+1)). *ExportValue* and *y2001* are the same variables as in Equation 3.2. To fit my assumption, the estimated coefficient of the interaction term between *Mileage* and *y2001* is supposed to be significantly positive.

$$ExportValue_{f,p,c,y} = \alpha_0 + \alpha_1 Mileage_f + \alpha_2 Mileage_f \times y2001_{f,p,c} + \alpha_3 y2001_{f,p,c} + \epsilon_{f,p,c,y}. \tag{3.3}$$

3.4 Baseline results and robustness tests

3.4.1 Baseline results

In column 1 of Table 3.4, I present the estimation results for Equation 3.2. The interaction term is found to be significantly positive. The concerned interaction coefficient is 0.0957, which means that on average, the railway speed-up project promotes the related firm-product-country level rail freight exports by 95.7 thousand USD per year. Since the mean of annual export value for rail freight products in firm-product-country-level is 137.8 thousand USD in 2000 (see Table B.1), I might infer that the speed-up promotes a 69% increase in related exports[12]. This result sounds extremely large but is still reasonable. First, the base for China-Central Asia export is small, thus a large growth rate is easy to achieve, and export shocks on non-rail freighted products lead to an overestimation on this result (will be shown in robustness tests). Further, in columns 2-4, I present the estimated results with data aggregated in other levels. Column 2 refers to an estimation with firm-product-freight mode-year level data, that is, I do not distinguish the destination countries of Central Asia. Columns 3 and 4 refer to estimations with firm-country-freight mode-year level data and product-country-freight mode-year level data. Results are all significant, and suggest positive export promotion effects from the railway speed-up project.

---

[12] In estimations, I use the absolute value of export as the dependent variable instead of the logarithm of the export value because there are many unpaired data. For the unpaired data, export value in a year without positive export records is naturally set as 0 after taking logarithms (assume the export value is 1 USD), the export performance change will thus be noise in the statistical estimation. For example, when comparing the performance between two exporters: exporter A (export value is 0 USD in 2000, and 1,000 USD in 2001) and exporter B (export value is 10,000 USD in 2000, and 100,000 USD in 2001), the actual export growth is obviously larger for exporter B. However, the growth is eliminated by taking logarithms: For exporter A, the growth is ln1000-ln1=6.9, and for exporter B, it is ln100000-ln10000=2.3. When I restrict samples to the records containing positive pair data, the baseline result remains consistent. The estimated interaction coefficient is 0.4208 at a 10% significance level, suggesting the speed-up increased related export value by 52.3% (Results are available upon request).

Furthermore, I present the estimation of Equation 3.2 with data of other years as a placebo test in columns 5-8 of Table 3.4. I set the placebo test with data of 2001-2002 (2001 is set as the "control" year and 2002 is set as the "treated" year), both of these two years being after the speed-up in 2000[13]. If the baseline results on the speed-up effects are unbiased and exclusively originated from the railway speed-up project, the placebo test result is supposed to be insignificant[14]. As shown in columns 5-8, I find that the results are not consistent with baseline estimations, neither of the interaction terms between *Rail* and *y2002* is significant and all the coefficients have much smaller absolute values than the baseline estimations. This implies that the estimated speed-up effect is not systematically biased.

Table 3.4: Baseline and placebo regressions

	Export Value							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	2000-2001				2001-2002			
Rail×y2001	0.0957***	0.1027***	0.6326***	0.1950***				
	(0.0176)	(0.0202)	(0.2348)	(0.0383)				
Rail×y2002					0.0075	0.0070	-0.0980	-0.0158
					(0.0161)	(0.0160)	(0.1323)	(0.0286)
Sample type	f-p-c	f-p	f-c	p-c	f-p-c	f-p	f-c	p-c
Observations	12916	12058	2216	6990	15390	15098	2836	7996
R-squared	0.0038	0.0036	0.0049	0.0057	0.0032	0.0034	0.0075	0.0069

Notes: Coefficients of single and constant terms are not reported. This applies for all the following estimations.
f-p-c: firm-product-country-freight mode-year level data;
f-p: firm-product-freight mode-year level data;
f-c: firm-country-freight mode-year level data;
p-c: product-country-freight mode-year level data, these apply for all specified tables.
Standard errors in parentheses
*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Then, I further consider the within rail freight exports variation of the speed-up project. Before conducting an estimation of Equation 3.3, in Figure 3.5, I present the kernel density of the export transaction distribution, with respect to sped-up mileage for the rail freight exports in 2000 and 2001. Descriptively, for the post speed-up year, 2001 (dashed line), a higher share of the transactions to Central Asia is taken by exporters located in regions more than 3,000 km away from Urumqi, and the share of transactions by exporters located within Xinjiang decreased compared with that in 2000. This suggests

---

[13] To ensure the pre-treatment parallel trend in rail freight and non-rail freight exports, the preferred placebo test is to compare export performance difference between 1999 and 2000. Unfortunately, customs data in 1999 is not available. In Oct. 2001, there was another railway speed-up (fourth round of the national railway speed-up project), however, this round focused on other railway lines, and does not affect China-Central Asia trade.

[14] Based on customs data, it is impossible to clearly identify whether or not manufacturing exporters (road freight exports) in western China benefited from the railway speed-up by any alternative, such as a decrease in the transportation cost of imported raw materials and intermediates (rail freight imports) from coastal regions or other countries. The placebo test, using data of 2001-2002, systematically relieves this concern.

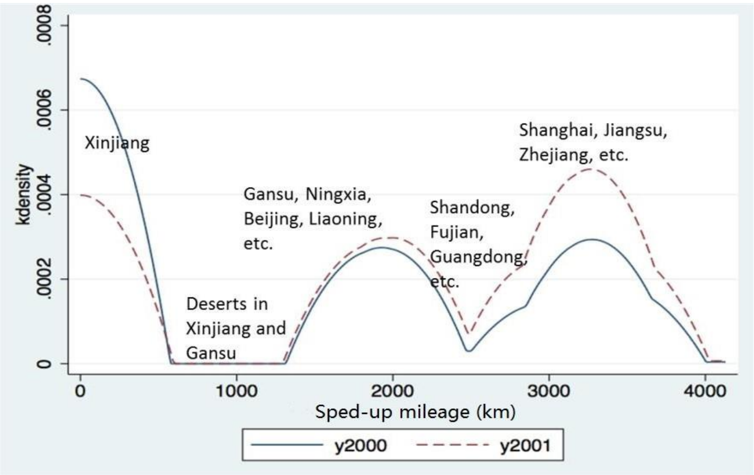


FIGURE 3.5: Kernel density of export transaction distribution with respect to sped-up mileage

Note: Cities within Xinjiang also enjoys considerably sped-up mileage (up to 500 km) because Xinjiang is of a rather large size. In the calculation of the sped-up mileage, I set the location of firms to the provincial level for simplicity. As a consequence, all exports originating from Xinjiang, no matter whether from Urumqi (capital of Xinjiang), or other cities, are seen as identical, with sped-up mileage being 0 km. Figures are based on the data of railway freight exports to Central Asia in 2000 and 2001.

that Central Asia exports are more likely to be made by coastal exporters instead of exporters in Xinjiang, because of the improving domestic transport efficiency. Figure 3.6 shows that before-and-after the speed-up, the ratio of export (Xinjiang/other regions) noticeably decreases.

TABLE 3.5: Effects of the speed-up on mileage

	Export Value					
	(1)	(2)	(3)	(4)	(5)	(6)
		2000-2001			2001-2002	
Mileage×y2001	0.0053*	0.0057*	0.0202			
	(0.0028)	(0.0030)	(0.0158)			
Mileage×y2002				-0.0010	-0.0009	-0.0355
				(0.0042)	(0.0041)	(0.0308)
Sample type	f-p-c	f-p	f-c	f-p-c	f-p	f-c
Observations	5102	4788	1120	6428	6344	1508
R-squared	0.0013	0.0014	0.0039	0.0022	0.0024	0.0053

Notes: Same as Table 3.4.
Standard errors in parentheses
*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Column 1 of Table 3.5 tells the related econometric results of equation (3). By restricting the samples in rail freight exports consistently, the interaction term is significantly positive, which suggests the west-east railway speed-up helps the long-distance exporters more. Average sped-up mileage in 2001 in firm-product-country-freight mode level is 1,772 km, and when compared to the case without speed-up, the promotion effect is 0.040 million USD (=0.0053×ln1773) (I set the *Mileage* as ln(Sped-up_mileage+1) to

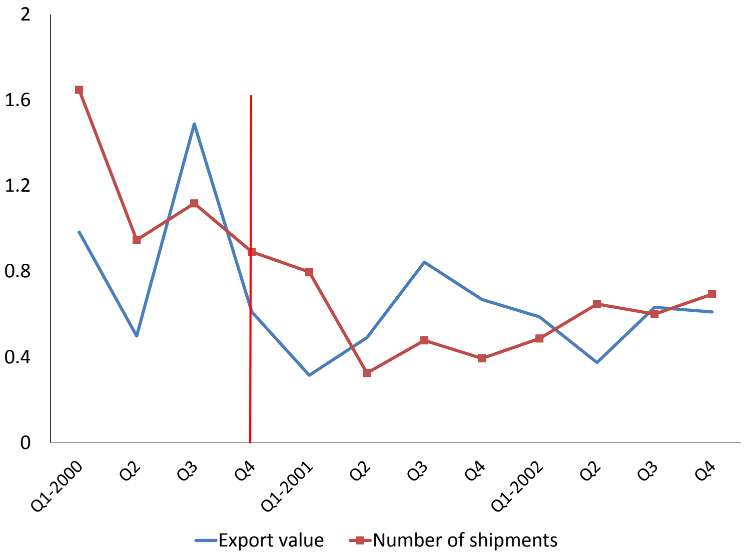


FIGURE 3.6: Ratio of railway freight exports from China (Xinjiang/other regions) to Central Asia

avoid a zero value for the logarithm). Mean of rail freight exports in firm-product-country level in 2000, with positive sped-up mileage, is 0.142 million USD, suggesting the average promotion to exports by the speed-up to be 28%. This is much smaller than the estimate in Equation 3.2.

I also test using other aggregated level data and the results are presented in columns 2 and 3. At a firm-country level, results are insignificant. Since I need information on location to obtain the sped-up mileage, I do not conduct the estimation with product-country level data, as in Table 3.4. Then, same as the baseline estimations, the counter-factual analysis in columns 4-6 of Table 3.5 do not overthrow the baseline conclusions. In 2001-2002, there is no positive difference-in-difference result, since there is no railway speed-up project in this period benefiting locations on the Silk Road.

As discussed in Section 3.3, only the export value of rail freight exports increases due to the speed-up in 2001, but shipment frequency does not increase proportionally. This is a natural result, for when facing a demand increase, exporters will choose to increase export value per shipment but not increase shipment frequency to save on freight costs (each shipment has a fixed cost, and the unit cost in freight has an increasing return feature). This is shown in Panel B of Table B.1, as rail freight exports increase the value per shipment compared to that of non-rail freight exports in 2001. I use the number of shipments to Central-Asia as the dependent variable, testing Equations 3.2 and 3.3. In Table 3.6, I present the results with 2000-2001 data in the first two columns, and then put the placebo test in columns 3 and 4. Column 1 shows that the railway speed-up promotes the benefited exporters by 0.22 times per year, in terms of number of shipments. Noting that the related average number of shipments in 2000 is 2.55

times, of which 0.22 (9%) is a considerable, however, much smaller growth rate than that of export value. This suggests that the promotion effect plays a role by increasing the export value per shipment but not through shipment frequency. Columns 2-4 refer to estimations of Equation 3.3 and placebo tests, where results are consistent with my assumptions, although in column 4, the related coefficient is negative.

Table 3.6: Effects of the speed-up on the number of shipments

	Number of Shipments			
	(1)	(2)	(3)	(4)
	2000-2001		2001-2002	
Rail×y2001	0.2213***			
	(0.0482)			
Rail×y2002			-0.0595	
			(0.0370)	
Mileage×y2001		0.1076***		
		(0.0076)		
Mileage×y2002				-0.0379***
				(0.0063)
Samples	all	rail	all	rail
Observations	12916	5102	15390	6428
R-squared	0.0183	0.0478	0.0654	0.0868

Notes: Estimation results are based on firm-product-country-freight mode-year level data. This applies for the following tables without specification.
Standard errors in parentheses
*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

3.4.2 Robustness tests

Robustness tests are mainly concerned with the demand shock of specific product exports and macroeconomic shocks. It is obvious from Table B.1 that in 2001, non-rail freight exports experience a sharp decrease, although it might be partially caused by the diversion effect of the freight mode, that is, from non-rail freight to rail freight because of the improvement in efficiency in rail freight from the railway speed-up. It is necessary to take into account the potential disturbance from export demand shocks. The main products exported to Central Asia through railways are machines, chemical raw materials, and building materials, while through non-rail modes are clothes, shoes, textiles, and household appliances. Products of heavy industries are mainly freighted through rail and light industries through non-rail modes, mainly roads (see the details of the exporting categories at the bottom of Table 3.7).

Since the main products are heterogeneous between rail and non-rail freight, it raises a potential disturbance to my baseline identification. Rise or fall of short-term demand for specific products that are freighted exclusively by rail or non-rail modes will affect the baseline results. For example, even if the export of rail freight products are actually not affected by the speed-up project, the baseline estimation will still be positively significant, as presented in the baseline result, if there is an unexpected fall in the export of clothes, shoes, textiles, household appliances, etc. (dominated by non-rail freight) in 2001, compared with that of 2000. In addition, if there were a rise in the demand of heavy products like machines, chemical raw materials, and building materials, it would also induce the same results.

To check this concern, firstly, in column 2 of Table 3.7, I exclude five product categories that have the highest export values in rail freight exports (Top 5 in rail) from the full sample, and estimate Equation 3.2. Interaction term coefficient becomes slightly smaller, to 0.0804, and remains significant at a 1% level. When I only include these five categories, the result is positively significant. Then in column 4, I exclude five categories which have the highest export values in non-rail freight exports (Top 5 in non-rail) from the full sample, the concerned coefficient becomes much smaller, although it is still significant at a 1% level. It could be explained by a fall in the demand for exports of non-rail dominated exports, which caused the overestimation of baseline results in column 1. After excluding the specific non-rail exports, the coefficient decreases to 0.0377, signifying a 37.7 thousand USD increase in related annual export value (increased by 27%)[15]. Similarly, in columns 5-8, I present the placebo tests with data from 2001-2002, in which neither of the concerned interaction terms is significant, suggesting that the baseline estimation results are not biased by specific product exports.

Secondly, on macroeconomic shocks, 2001 is a special year in international trade, for both China and the World. In Dec. 2001, China joined the WTO as its 143rd member, which promoted China's international trade significantly in the following years. However, this variation does not affect my estimations, since I set the treated and control samples within the year 2000-2001, namely, these two years' export performance is actually not affected by the WTO agreement, issued by the end of 2001. Then, another economic shock is the "Sept. 11 Terrorist Attack," as a consequence of which world trade shows an unexpected negative growth. Since the 1980s, the only year with negative growth in world trade is 2001. World trade value increases by 13% in 2000, but drops by 4% in 2001 because of the global economic and political instability. Then in 2002, it recovers by increasing 4% (Source: International Trade Statistics, WTO). Overall, Xinjiang is sensitive to terrorism because of its geographic location as a border area, and historical

---

[15] In the case of excluding top 10 product categories in non-rail freight exports, this figure further decreases to 23% and is still statistically significant at a 1% level. Actually, the top 5 product categories in non-rail freight exports account for 67% of the total exports with non-rail freight. For the top 10, the share reaches 82%, therefore, excluding the top 10 non-rail freight exports almost removes all the control samples. Hence, exclusion of the top 5 product categories is supposed to exclude the disturbance from non-rail export demand shocks. In addition, if the decrease in exports under non-rail freight is not caused by intra-industry demand shocks but by overall economic fluctuations, the estimated promotion effect is underestimated if excluding these products exports. Thus, I can suppose that the actual treatment effect should be around 30%. For the number of shipments, the related result decreases to around 13% when excluding top 5 product categories in non-rail.

Table 3.7: Demand fluctuations from specific products

	Export Value							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	2000-2001				2001-2002			
	Full	ex Top 5 rail	only Top 5 rail	ex Top 5 non-rail	Full	ex Top 5 rail	only Top 5 rail	ex Top 5 non-rail
Rail×y2001	0.0957***	0.0804***	0.0690**	0.0377***				
	(0.0176)	(0.0210)	(0.0341)	(0.0121)				
Rail×y2002					0.0075	-0.0001	(0.0143)	0.0163
					(0.0161)	(0.0143)	(0.0445)	(0.0167)
Observations	12916	9822	3094	9656	15390	11050	4340	12142
R-squared	0.0036	0.0062	0.0046	0.0007	0.0030	0.0080	0.0001	0.0014

Notes: Same as Tables 3.4 and 3.6.
Customs database records the eight-digit customs classification code (HS-8) for each transaction. For simplicity, I cut the first two digits of the HS-8 (HS-2), and classify the cargo of China-Central Asia export into 89 categories. The details can be checked at: http://www.allmyinfo.com/eng/services/index1-3.asp. Top 5 in rail: 84,27,28,85, and 73 (ranked); Top 5 in non-rail: 64,62,61,58, and 85 (ranked). Products that the HS-2 refers to:
27: Coal, Asphalt, Lubricating oils, Vaseline, etc.;
28: Chemical products (Iodine, Calcium, Phosphorus, etc.);
58: Various textiles;
61, 62: Suits, Coats, Skirts, Sportswear, Shirts, Trousers, Underwear, etc.;
64: Shoes, Accessories of shoes;
73: Pipes, Screws, Steel building materials, etc.;
84: Engines, Air conditioning, Refrigerators, Machine tools, etc.;
85: Telephones, Digital cameras, Televisions, Lamps, Microwaves, Cookers, etc.
Standard errors in parentheses
*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

ethnic conflict, so the “Sept. 11 Attacks” greatly affect the China-Central Asia export. The ratio of export value to Central Asia (Sept.-Dec./annual) from 2000-2003, fell in 2001 due to the “Sept. 11 Attacks,” dropping from higher than 0.4 (in 2000, 2002, and 2003) to less than 0.3 (Sources: China’s customs database, 2000-2003). In the period of Sept.-Dec. 2001, it is obvious that China-Central Asia export experienced abnormal decline, as did world trade, because of the terrorist attacks. I test the stability of my baseline results concerning this macroeconomic and political fluctuation. I exclude the exports in Sept.-Dec. in both years, since exports are seasonally heterogeneous. Then I estimate the placebo test with the data excluding the Sept.-Dec. period.

In column 1 of Table 3.8, the interaction term between *Rail* and *y2001* remains significant and valued at 0.0769, slightly smaller than that of the baseline estimation of 0.0957. Similarly, in column 2, I keep the rail freight export samples only, and insert the sped-up mileage into estimation, with the result remaining unchanged. Columns 3 and 4 give a placebo test with consistent results.

Table 3.8: Effect of the "Sept. 11 Attacks"

	Export Value			
	(1)	(2)	(3)	(4)
	Jan.-Aug., 2000-2001		Jan.-Aug., 2001-2002	
Rail×y2001	0.0769*** (0.0155)			
Rail×y2002			0.0029 (0.0135)	
Mileage×y2001		0.0069** (0.0033)		
Mileage×y2002				-0.0000 (0.0025)
Samples	all	rail	all	rail
Observations	9264	3660	10622	4520
R-squared	0.0031	0.0015	0.0007	0.0000

Notes: Same as Tables 3.4 and 3.6.
Standard errors in parentheses
*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

3.5 Mechanisms

In this section, I investigate the channels of the speed-up effect on China-Central Asia export. Precisely, I try to distinguish between the diversion and creation effects. I am concerned about the diversion of Central Asia exports between different freight modes: rail, road, etc., and between different locations: Xinjiang and other regions of China. In addition, I check time- and capacity-related causes to export performance stemming from the speed-up, as well as intensive and extensive margin effects. Then, I test the heterogeneity on export promotion, with respect to unit price and firm size.

3.5.1 Channels of the speed-up effects on exports

Speed-up effect on exports is observed in the results of baseline estimations. Except the export creation effect, diversion effects also might contribute to the baseline estimations if exports once taken by trucks, are switched to trains because of increased speed and capacity of railways. In this case, the two effects are mixed. I divide exports by freight mode to distinguish export diversion and creation effects. If a product (HS-8 level) is freighted by both trains and trucks within the same firm-product-country level export, I set it as having multiple modes of freight. Otherwise, it is identified with a single freight mode. For the first case, it is supposed to show a diversion effect from non-rail to rail responding to railway speed-up.

Results of columns 1 and 2 of Table 3.9 suggest that for multiple freight mode products, the coefficient of the interaction term is much higher than baseline estimations and that of single freight mode products. Based on the estimated coefficients and mean value in related aggregated level, the export promotion effect of the railway speed-up in multiple freight mode products is 167%, and for the rest, it is 64%. When I test using the number of shipments instead of export value, the difference persists, the extent being 42% (multiple) to 12% (single) (results are available upon request). Obviously, multiple freight mode products enjoy a greater promotion effect because of a latent diversion effect from non-rail freight modes to rail. However, multiple freight mode samples account for only 18% of rail freight exports to Central Asia in 2000-2001, suggesting the diversion effect from non-rail to rail freight mode cannot fully explain the effects of the railway speed-up on export. Results of column 2 suggest that the export creation effect dominates the speed-up effect, from the evidence of the other 82% of rail freight exports (use single freight mode). Then in columns 3 and 4, I identify the exports with a multiple and single freight mode, at product (p) level instead of the firm-product-country (f-p-c) level. The estimated diversion effects become weak, with both the magnitude and the extent being similar in these two groups. This shows that the export diversion effect exists but is not the main driving force, as the creation effect matters more.

Next, I am concerned about another kind of diversion effect, the diversion of Central Asia exports between Xinjiang and other regions of China. As shown in Table 3.5, exporters with longer sped-up mileage will benefit more from the speed-up. It is straightforward to suppose that exports in Xinjiang (enjoying limited sped-up mileage) will be transferred to other regions, such as Jiangsu and Zhejiang (coastal regions in Eastern China), which have a relatively higher manufacturing productivity and competitive advantage from Xinjiang. Column 5 presents a simple estimation similar to Table 3.5. $Xinjiang$ is a dummy variable equal to 1 for Xinjiang exporters and 0 for others. Within the rail freight exports, Xinjiang exporters have weaker performance than their competitors in other regions after the speed-up. Referring to the estimated coefficient, Xinjiang exporters with rail freight saw a decrease of 31% in export values in 2001, compared with that in 2000 and that for exporters in other regions.

Additionally, in column 6, I conduct an alternative estimation to evaluate this diversion effect. For a product that is exported to Central Asia both by Xinjiang exporters and exporters in other regions in 2000, I set it as a “both” product with a value equal to 1, and for other product exports taken by either Xinjiang exporters or exporters in other regions, the variable is equal to 0. Within Xinjiang exports, I find that the export value of “both” products decline compared to other products in Xinjiang, by 39% in 2001 compared to their 2000 level. Columns 5 and 6 suggest that Xinjiang exporters are negatively affected by the diversion effect from the speed-up.

Table 3.9: Two kinds of diversion effects

	Export Value					
	(1)	(2)	(3)	(4)	(5)	(6)
	Freight mode (rail&non-rail)				Origin (Xinjiang&others)	
	Multiple: f-p-c	Single: f-p-c	Multiple: p	Single: p	Full	Within Xinjiang
Rail×y2001	0.4057***	0.0768***	0.0903***	0.1054***		
	(0.1346)	(0.0144)	(0.0237)	(0.0262)		
Xinjiang×y2001					-0.0418*	
					(0.0222)	
Both×y2001						-0.0656**
						(0.0319)
Samples	all				rail	
Observations	1120	11796	7650	5266	5102	2092
R-squared	0.0244	0.0029	0.0053	0.0025	0.0013	0.0021

Notes: Same as Tables 3.4 and 3.6.
Standard errors in parentheses
*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

In Table 3.10, I test the differential domestic access improvement effects with respect to time saving and capacity expansion. The speed-up improves both railway speed and cargo capacity. Perishable products are most sensitive to travel time, while machinery are generally capacity sensitive. I test using these two products to distinguish between the role of time and capacity effects. In column 1, I conduct an estimation of equation (2) within food exports, and in column 3, I restrict samples within rail freight exports and set a dummy variable for food (equal to 1 for food exports and 0 for others). Although in the first estimation I chose non-rail freight food exports as the control, and in column 3 I set rail freight non-food exports as the control, interaction terms of both estimations capture the time effect. In columns 2 and 4, I use machinery exports instead of food and run the same estimations with columns 1 and 3, interaction terms thus capturing the capacity effect. Obviously, the time effect is not statistically significant in columns 1 and 3, while machinery exports benefit from the speed-up significantly, by 31% (column 1) and 45% (column 3). Results suggest that capacity expansion is more important than time saving for railways[16].

Finally, I test the speed-up effect on exports from the viewpoint of extensive/intensive margin effects. For the extensive margin, I estimate the variation in the number of exporters to the speed-up project, followed by Equation 3.4. The dependent variable is the number of exporters in product (HS-8)-country-freight mode-year level. *Rail* and the $y2001$ dummy and equation settings are identical to Equation 3.2.

$$Number.Exporters_{p,c,t,y} = \alpha_0 + \alpha_1 Rail_{p,c,y} + \alpha_2 Rail_{p,c,y} \times y2001_{p,c,t} + \alpha_3 y2001_{p,c,t} + \epsilon_{p,c,t,y}. \tag{3.4}$$

---

[16] However, the insignificant time effect might be latently caused by two facts: food exported to Central Asia is not always perishable. For example, wheat flour and rice are not so sensitive to time in trade, while perishable products like fruits are rarely exported to Central Asia. The travel time saved by the speed-up in freight trains is not as much as in passenger trains.

Table 3.10: Distinguishing between time and capacity effects

	Export Value			
	(1) Time-sensitive	(2) Capacity-sensitive	(3) Time-sensitive	(4) Capacity-sensitive
Rail×y2001	0.0068	0.0659*		
	(0.0134)	(0.0392)		
Food×y2001			0.0022	
			(0.0412)	
Machinery×y2001				0.0966***
				(0.0250)
Samples	Food	Machinery	Rail	Rail
Observations	1032	2586	5102	5102
R-squared	0.0069	0.0046	0.0006	0.0035

Notes: Same as Tables 3.4 and 3.6.
Time-sensitive: Food refers to products with HS-2 (01-24).
Capacity-sensitive: Machinery and equipment refer to products with HS-2 (84-89).
Standard errors in parentheses
*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

In column 1 of Table 3.11, results show that the number of exporters increased by 8% (mean of the related indicator is 1.40 in 2000) because of the railway speed-up. Magnitude of the extensive margin effect is not large, in spite of positive and significant coefficients. This is latently caused by the diversion effect, as shown in columns 5 and 6 of Table 3.9. Due to the railway speed-up and improvement in domestic accessibility to Xinjiang, on the one hand, exporters in other regions tended to participate in the Central Asian market. On the other, Xinjiang exporters tended to exit this market due to weak competitive power. In this case, extensive margin is not significant, if the overall exporters do not increase, even if there is a switch between exporters in Xinjiang and other regions. In columns 2 and 3, I separately test the extensive margin for Xinjiang exporters and exporters from other regions to avoid such an offset effect. Results suggest that this is not the case, as the extensive margin is limited, even if taking into account the offset effect. For Xinjiang exporters, there is no significant decrease in the number of exporters, and for other regions, the extensive margin is significantly positive but the extent is still less than 10%. In column 4, I estimate at the product (HS-4)-country level instead of the product (HS-8)-country level, with result being unchanged, in both significance and magnitude of the extensive margin (around 8%).

On the intensive margin effect, I keep only the existing exporters at the firm-product-country level that have export records in 2000 and run estimations of Equation 3.2. The promotion effect is 58% in export value (column 5) and 23% in the number of shipments (column 7), which is only slightly smaller than the results with the full sample. In columns 6 and 8, I present the results of Equation 3.3 when keeping only the samples from existing exporters and rail freight exports. Results are similar to that of the full samples (containing the new exporters in 2001). Based on these estimations, results suggest that intensive margins play a major role, with the speed-up effect on exports originating from extensive margins being limited.

TABLE 3.11: Distinguishing between extensive and intensive margins

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Extensive Margin				Intensive Margin			
	Number of Exporters				Export Value		Number of Shipments	
Rail×y2001	0.1048***	-0.0415	0.0727***	0.1447***	0.1068***		0.4651***	
	(0.0365)	(0.0455)	(0.0241)	(0.0523)	(0.0256)		(0.0585)	
Mileage×y2001						0.0051***		0.0807***
						(0.0028)		(0.0086)
Samples	all	only Xinjiang	exclude Xinjiang	all	existing	existing &rail	existing	existing &rail
Sample type	p(HS8)-c	p(HS8)-c	p(HS8)-c	p(HS4)-c	f-p-c	f-p-c	f-p-c	f-p-c
Observations	6990	3530	4256	3918	8052	3138	8052	3138
R-squared	0.0221	0.0303	0.0026	0.0223	0.0139	0.0099	0.2693	0.3214

Notes: Same as Tables 3.4 and 3.6.
Standard errors in parentheses
*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Based on a series of estimations, the big picture of the channels of the speed-up effect on export becomes clear. In the total speed-up effect (increase in related exports by 69%, as in baseline estimations (Table 3.4), and at least 27% in robustness checks, after removing various disturbances (Tables 3.7 and 3.8)), the export creation effect plays the main role (column 2 of Table 3.9), with a limited diversion effect from non-rail to rail freight exports (column 1 of Table 3.9). Within rail freight exports, to some extent, it is transferred from Xinjiang exporters to exporters from other regions of China (mainly the coastal regions), because of improvement in accessibility to Xinjiang (columns 5 and 6 of Table 3.9). Capacity expansion is the more direct factor promoting export rather than the increased speed (Table 3.10). I use only a short period (two years) data, as exporters might not choose to enter/exit markets immediately in response to the railway speed-up, but existing exporters increase/decrease the related exports, as a result of which, intensive margin becomes a major source of the speed-up effect (Table 3.11).

3.5.2 Heterogeneous effects with respect to unit price of exporting products and firm size

As per the opinion of the World Bank (2006), interior cities in China are more likely to be able to compete in the production of bulk goods (e.g., coal) suitable for leisurely shipment or high-value goods (e.g., computer chips) suitable for air cargo. Transport will tend to have the most deleterious impact on medium-value goods for inland cities of China, and hence, they are most sensitive to improvement in transport. To consider this, I divide the exported products into three categories, with an equal number of items in each classification, based on the unit price: low-, medium-, and high-value products. Table 3.12 gives results consistent with World Bank (2006). Both the estimations with the dependent variable as the export value and the number of shipments present significant promotion effects for medium-value products. For low-value products, neither of the results is significant. That is, low value products tend to enjoy leisurely shipment, thus are not sensitive to the increased speed and capacity in rail. Actually, when I restrict the samples in raw material exports (that enjoy leisurely shipment), there is no significant speed-up effect observed (results are available upon request). For high-value products, results are not consistent, with the test with respect to export value being significantly positive while the test with respect to the number of shipments being insignificant, suggesting that the speed-up project raised the export value per shipment for high-value products. It is a reasonable assumption, because exports to Central Asia are dominated by low- and medium-value products. Even for high-value products in this classification (not less than 4 USD in unit price), most of them cannot afford air cargo. As shown in Section 4, the number of shipments is not increased as much as the export value from the speed-up project, ensuring that the insignificant result on column 6 is not contradictory to that on column 3. Overall, the export of medium-value products is promoted by 135%, for high-value products it is 86%, while for low-value products it is insignificant[17].

Then, on the firm sizes of exporters, there is no data directly corresponding to the size. I use the average value per shipment to be representative of the firm size. Specifically, I calculate the mean of export value per shipment for each firm, based on the data of 2000-2001, and set the medium value as the threshold. Firms whose average export value per shipment is larger than the medium value are viewed as a large firm and as a small firm otherwise. Results in Table 3.13 are significant for two kinds of exporters, and for two indicators: export value and the number of shipments. However, the extent is very different. For large exporters, rail freight export in 2001 increased by 85% compared to that in 2000, and that for small exporters is relatively small, by 25%. On the number of shipments, the difference becomes much smaller, with 16% and 13%, respectively. Large exporters benefit more than small exporters in the speed-up. This is reasonable, as larger

---

[17] I test the variation on the unit price itself. I use the unit price instead of export value as the dependent variable for the samples that have records in both the years. Results are not significant, as shown in Table B.5. The unit price is the F.O.B. price, containing the producer's price and domestic freight costs. Within the f-p-c level, there is no actual price change, since the speed-up does not directly decrease monetary freight costs. Although there might be an indirect decrease in freight costs because of capacity expansion and relieved congestion (as suggested by Li and Chen (2013)), and the increasing returns in transportation (as suggested by Mori (2012)), if exporters increase their export volume, it is reasonable that the unit price does not react immediately to cost change. Lu et al. (2013) estimate the unit price effect of China's export response to US anti-dumping investigations, although export volume greatly decreased and though the anti-dumping duty is high, up from 44%-300% of the export value, the unit price does not change significantly.

Table 3.12: Effect of products with different unit prices

	Export Value			Number of Shipment		
	(1)	(2)	(3)	(4)	(5)	(6)
	Low	Medium	High	Low	Medium	High
Rail×y2001	0.0199	0.1040***	0.1394***	0.0550	0.3127***	0.0691
	(0.0168)	(0.0355)	(0.0347)	(0.1014)	(0.0851)	(0.0651)
Observations	4026	4804	4086	4026	4804	4086
R-squared	0.0032	0.0050	0.0033	0.0127	0.0458	0.0006

Notes: Same as Tables 3.4 and 3.6.
Low: F.O.B. price not-higher-than 0.8 USD; Medium: F.O.B. price higher than 0.8 and lower than 4 USD; High: F.O.B. price not-lower-than 4 USD. Price is based on the average unit price at firm-product-country-freight mode level in 2000.
Top 5 categories (based on export value) in low value products, HS-2 in rail: 27, 9, 28, 73, and 55; HS-2 in non-rail: 58, 61, 54, 95, and 63.
27: Coal, Asphalt, Lubricating oils, Vaseline, etc.;
9: Tea, Coffee beans, Paprika, etc.;
28: Chemical products (Iodine, Calcium, Phosphorus, etc.);
73: Pipes, Screws, Steel building materials, etc.;
55: Fabrics;
58: Various textiles;
61: Suits, Coats, Skirts, Sportswear, Shirts, Trousers, Underwear, etc.;
54: Cloth and Gauze;
95: Toys, Pokers, Accessories of sporting goods;
63: Bedspreads, Bath towels, Rags, etc.
Standard errors in parentheses
*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

exporters are more likely to be restricted by railway cargo capacity, and by increasing the cargo capacity, they benefited more from the increasing returns of transportation.

Table 3.13: Size of the exporters

	Export Value		Number of Shipment	
	(1)	(2)	(3)	(4)
	Large	Small	Large	Small
Rail×y2001	0.1537***	0.0037***	0.3093***	0.1801***
	(0.0276)	(0.0012)	(0.0697)	(0.0501)
Observations	8150	4766	8150	4766
R-squared	0.0059	0.0014	0.0322	0.0021

Notes: Same as Tables 3.4 and 3.6.
Standard errors in parentheses
*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

3.6 Exports of Xinjiang

In this section, I investigate the overall speed-up effect on the export performance of the Xinjiang, the gateway of China to Central Asia. As shown in previous results, there is a diversion of exports from Xinjiang to other regions. In the case of high transportation costs, Xinjiang enjoys a locational advantage for Central Asian export, and it thus becomes the main exporter to Central Asia, although it does not have comprehensive comparative advantages in productivity. However, due to increasing market accessibility for other regions to Xinjiang, the importance of the productivity advantage matters more, which negatively affects Xinjiang's comparative advantages. As mentioned in Mori (2012), when the travel time between Tokyo and Osaka (the two largest cities, also the two largest transportation hubs in Japan) decreased from 20 hours (1889), to 8 hours (1935), and further to 2.5 hours (2005), the population size in Tokyo was 1.1 (1889), 1.6 (1935), and 2.7 times (2005) as large as in Osaka. This is similar to the ratio of exports to Central Asia provided by the Xinjiang exporters. In 2010, 70% of exports that exited China through Urumqi customs (most to Central Asia, few to Middle East, etc.) are from non-Xinjiang exporters (Wu, 2011). While in 2000, non-Xinjiang exporters accounted for only 29% of the exports to Central Asia (Source: Customs database). This huge transformation happens consistently throughout the decade, as China experienced sharp expansion of transportation infrastructure.

On the other hand, except the Central Asian market, the speed-up also increased the market accessibility to other international markets for exporters in Xinjiang. Xinjiang exporters might gain higher accessibility to China's coastal lines and ports, as well as markets in Japan, US, etc. Although Xinjiang does not have comparative advantages in export compared to coastal regions because of its remoteness to the coast, it has comparative advantages in specific heavy industry production that is heavily dependent on energy and raw materials, since it has a rich endowment of mineral resources. Previous quantitative evidence has proven that domestic accessibility is important to export participation, and hence, I estimate the speed-up effect on export performance of Xinjiang exporters to other international markets except Central Asia.

Similar to the exports from Eastern China to Central Asia, exports from Xinjiang have to experience the land route through the Silk Road to reach the coastal lines in the case of ocean shipping to destination countries. I extract the export records from Xinjiang exporters, and go in two directions: countries that are located to the west of Xinjiang, namely, Central Asia and the Middle East, for which goods exit China through Urumqi customs. This direction is less affected by the railway speed-up and the domestic freight distance is short. Another direction is countries for which goods are freighted through ocean shipping, like Japan, US, etc., that is, these exports have to experience long-haul domestic freight from Xinjiang to coastal ports, like Tianjin, Shanghai, or Guangzhou ports. Among these ports, Shanghai is the most frequently used, based on export records. If products are moved by railways, then to reach the Shanghai port, they have to use almost the full length of both the Lanxin and the Longhai lines, which were sped-up in Oct. 2000. This provides a good opportunity to observe the railway speed-up effects when I limit the export samples of Xinjiang exporters, and go for two exit customs: Urumqi and Shanghai[18].

I identify the domestic accessibility change, based on the exiting customs. Rail freight exports from Xinjiang exporters exiting through Urumqi customs are set as the control group, and rail freight exports from Xinjiang exporters exiting through Shanghai customs are set as the treatment group. Estimation equation is as follows, $ShanghaiCustom$ is a dummy variable, equal to 1 if the transaction is finished in Shanghai customs and 0 if it exits China through Urumqi. Other variables and equation settings are identical with Equation 3.2:

$$\begin{split} ExportValue_{f,p,c,y} = \alpha_0 + \alpha_1 ShanghaiCustom_{f,p,c,y} + \alpha_2 ShanghaiCustom_{f,p,c,y} \times y2001_{f,p,c} \\ + \alpha_3 y2001_{f,p,c} + \epsilon_{f,p,c,y} \end{split} . \tag{3.5}$$

An identification problem arises from this estimation strategy. For the exports from Xinjiang exiting through Urumqi (control group), it is easy to identify the domestic freight mode from the customs database (the freight mode when exiting China). For the treated group, exports from Xinjiang to countries like Japan and the US, the domestic freight mode might be roads or rail, but the exiting mode is ocean shipping. Customs data records only the exiting freight mode in ocean shipping, but does not present the domestic modes for each transaction. In such cases, I divide them based on the export records from Eastern China to Central Asia (the opposite direction), since they need to experience the same domestic route. Specifically, for one kind of export product (HS-2 level), if exports from Eastern China to Central Asia are mainly freighted by rail/road, then I assume in exports from Xinjiang to the coast, it is also freighted by rail/road.

---

[18] Although for some export cases, it is possible to, firstly, use rail or roads to reach the Yangtze River (from Xinjiang to Chongqing, or Wuhan), and then transfer on the Yangtze riverside, that is, they transfer to ship freight, and finally exit China through the Yangtze River. Such cases are rare, since the Yangtze River is far from the Silk Road, so it is not advisable to use the Yangtze River for transport from Xinjiang to Shanghai.

This is reasonable, as the products freighted by rail and road are naturally different[19][20].

TABLE 3.14: Export performance of Xinjiang

	Export Value			Number of Shipment		
	(1)	(2)	(3)	(4)	(5)	(6)
ShanghaiCustom	0.0754***	0.1238***	0.0552	1.6801***	1.4831***	0.2454
×y2001	(0.0320)	(0.0322)	(0.0572)	(0.1260)	(0.1233)	(0.2137)
Samples						
#treated (Shanghai)	rail	rail	road	rail	rail	road
#control (Urumqi)	rail	road	road	rail	road	road
Observations	1976	4852	4510	1976	4852	4510
R-squared	0.003	0.008	0.007	0.091	0.042	0.030

Notes: Same as Tables 3.4 and 3.6.
I extract only the export records of the first eight months (Jan.-Aug.) of the related years to relieve the "Sept. 11 Attack" effects, and exclude the postal and air freight exports, and government donation exports.
Standard errors in parentheses
*, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

In columns 1 and 4 of Table 3.14, for both treated and controlled samples, I restrict only for rail freight goods. Results suggest that after the railway speed-up, Xinjiang exporters increased their exports through Shanghai customs, as compared to that through Urumqi customs, by 70% in export value and 144% in number of shipments. For the latter, it is obviously larger than in the opposite direction and suggests that the speed-up benefits the small-sized exporters in Xinjiang. Then, I conduct estimations with the road

---

[19] I identify the freight mode from Xinjiang to the coast based on the two-digit customs classification code (HS-2) that mostly exports in the opposite direction. The eight-digit customs classification code (HS-8), with specific goods and the related freight mode, is a possible exercise in this case, but I find in the HS-8 level, export products moving from Xinjiang to the coast, and export products moving from Eastern China to Central Asia via Xinjiang, are totally different. In the HS-2 level, I find that trains convey most of the heavy industry products, while trucks mainly convey light industry products. Precisely, I divide the export records based on HS-2, and then choose the top ten categories in export value. In these top 10 categories, I identify each category with its related freight mode: railway or road, based on the main freight modes for the same category in China-Central Asia export (opposite direction). The top ten categories in HS-2 (based on export value) are: 85, 52, 29, 84, 60, 61, 82, 62, 55, and 63. I keep only 85, 52, 29, 84, and 82 as railway freighted, since in China-Central Asia export, such exports are mainly taken by rail, and they are heavy industry products (82,84,85: machine tools, 52: chemical raw materials) and agricultural products (55: cotton). For 60-63 and 55, they are light industry products (textiles, fabrics, clothes, etc.) and are mainly road freighted. These top five heavy industry products account for 39.3% of the total exports from Xinjiang (22.8% in the number of shipments). In columns 1 and 4, I choose the treated samples (Shanghai customs) as exports of these five categories, and the control samples (Urumqi customs) as rail freighted exports from Xinjiang.

[20] I am also concerned about the fact that if exports from the same firm with the same destination country and product might exit China through different customs, my identification might be affected by this endogeneity. That is, because of the speed-up, for the same firm, product, and destination country, the exit port might change to meet the higher freight efficiency, such as transfers from Dalian to Shanghai ports. This will also increase the exports through the Shanghai port. I evaluate such cases as follows: if in the f-p-c level, there is at least one transaction that occurred in Shanghai customs in 2000-2001, I assume that for this f-p-c level data, all the transactions in 2000-2001 are seen as having occurred in Shanghai, even if some transactions actually occurred at other customs. The fact is that such cases are rather rare in exports from Xinjiang, less than 200 cases in two years. For the same f-p-c level export, if firms choose to exit from the Shanghai customs, in most cases, they will not use other ports of exit. Thus, the concern about switching ports could be omitted.

freight exports as control samples. As the rail freight exports from Xinjiang through Urumqi customs tend to be transferred to coastal exporters based on the results of Table 3.9, thus the control group in estimations of columns 1 and 4 are also affected by treatment. However, road freight exports of Xinjiang (exit through Urumqi customs) are not affected by this kind of diversion effect. Results in columns 2 and 5 are consistent. Significant export promotion effects still exist, and the placebo tests in columns 3 and 6 are insignificant when I include road freight exports in both treated and control samples, that is, neither of the groups is affected by the speed-up project.

For the export flows from Xinjiang to China's coast, I do not conduct a series of estimations investigating the channels, because identification on freight mode is not feasible due to data restrictions. However, the limited results are still informative. The overall effects to Xinjiang from this sudden exogenous variation are two-fold. A higher domestic market access induces the declining locational advantage of Xinjiang as the gateway from China to Central Asia. However, better domestic market access also improves the accessibility of Xinjiang to international markets through the coast. When talking about the overall export performance of Xinjiang, although there is a temporary decline in export value in 2001 (partially because of the "Sept. 11 Attack"), from 2002, it continues to rapidly grow in terms of international trade. In the Tenth Five-Year Plan period (2001-2005), its annual trade growth was 28.5% (33.2% in exports and 22.3% in imports), much higher than its GDP growth of 10.1% and total investment in fixed asset growth of 17.2%, in these five years. This figure was also higher than the national annual trade growth of 24.6% (Source: Statistical Communiquá of the National Economic and Social Development, 2005).

3.7 Conclusion

I investigate the effect of domestic accessibility improvement on export performance, using the variation from China's railway speed-up in west-east lines. I find that exports from China to Central Asia through railways increased by (at least) around 30% after the railway speed-up, compared with exports through other freight modes. Domestic accessibility proves to be important to international trade participation, and it has heterogeneous effects on products and exporters (medium-value products and large exporters benefit more). The speed-up caused re-organization of exports, in terms of location of exporters and freight modes. Capacity expansion is the more direct factor promoting export, rather than the increased speed. Intensive margin proves to be the major source of the speed-up effect, not the extensive margin.

For a specific region, the net welfare gains of Xinjiang, the gateway of China to Central Asia, from the speed-up project are mixed. Xinjiang exporters have fewer exports to Central Asia, but more to other international markets, like the US and Japan. By increasing transportation efficiency, the locational advantage as the gateway is changing, and Xinjiang's economic development is thus affected in its export structure.

Although China-Central Asia export is not crucial, in the sense of export value, to the Chinese economy, the implication of domestic accessibility could be extended. First, social returns from railway investments prove to be considerable and prospective. Railways hold a considerable share of freight for raw materials in China, but for manufacturing goods, its share is less than 15% (China Transportation Yearbook, 2006). Trucks are still the dominant freight mode, undertaking 72% of the total freight in China (Annual Report of Ministry of Transport of P. R. China, 2006). This is highly confounding, since trains have a lower per unit cost in long-haul freight than trucks. Effective investment on improving railway freight efficiency is supposed to be promising.

Second, transportation infrastructure investment is always beneficial to overall national welfare, but not necessarily to specific regions. China pushes forward to reach balanced regional development between coastal regions and inland, landlocked regions, and puts a priority on investment in inland regions, such as the "Go-West Campaign." Transport infrastructure in these regions is heavily supported. On the one hand, higher domestic accessibility helps productive inland exporters to participate in exports, with the exported products of inland regions enjoying comparative advantages. On the other, China tries to extend domestic demand instead of an export-oriented economy for future economic growth, which is seen as an opportunity for the development of inland regions because of huge local populations, and potentially strong local demand. However, with high domestic accessibility, coastal regions tend to dominate both the international and the domestic inland markets (manufacturing sector), if the coastal regions retain their competitive power. It is necessary for policymakers to consider this trade-off.

Chapter 4

The Role of Coal Mine Regulation in Regional Development[1]

4.1 Introduction

Coal is the dominant primary energy in China, accounting for approximately 80% of China's energy structure and 70% of its primary energy consumption. In addition, over 90% of China's coal consumption is domestically provided. China's coal mining and coal industry are therefore crucial for the social economy. Consistent with high coal production, coal mine accidents are frequent in China and have been considerably more highlighted. China produces 35% of global coal. However, because of insufficient safety facilities and poor mining productivity, coal mine deaths in China account for 80% of the global coal industry mortality (2003). The overall coalmine labor productivity is also low, reaching approximately 2.2% of that of the United States in 2003. One of the main reasons for this is the small-size town and village enterprise (TVE) coal mines once permitted in China.

These poor conditions have prompted the Chinese government to promote regulation in the coal industry. China has regulated the coal industry since the 1990s by expanding fixed-asset investments and by closing small-sized mines with low productivity and high

---

[1] This is a joint-work with Kentaro Nakajima. We thank Ching-mu Chen, Dao-Zhi Zeng, Huihui Deng, Jae Hong Kim, Zhao Chen, Yuan Zhang, Ruixiang Xiong, Zhikuo Liu, Jianfeng Wu, and participants of the Third Asian Seminar in Regional Science (Hualien), seminars in Tohoku University (GSIS), and Fudan University (CCES) for the valuable comments. We gratefully acknowledge the financial support from the Japan Society for the Promotion of Science (JSPS) (#22223003) and from the JSPS Research Fellowship for Young Scientists (DC2). All errors are ours.

mortality rates. However, as shown previously, China's productivity rates and mine accidents are still far from satisfactory. In 2008, mortality accounted for 70% of the world total; however, this number has improved since the 1990s. From 1997 to 2009, the mortality rate (death worker/ton) declined by 80%, labor productivity (ton/worker) improved by 205%, and national coal output increased by 123% (sources: Statistical Communiqué of the National Economic and Social Development and the China Coal Industry Yearbook).

In this study, regulation is evaluated in terms of regional economic growth. According to the existing arguments on the resource curse and its effect mechanism on economic performance, the crowding-out effect (which suggests that resource abundance crowds out investment, human capital, and innovation in non-resource sectors and thus hinders economic growth) is a main cause of the resource curses in many regions globally (see, e.g., Sachs and Warner, 1995; 2001; Frankel, 2010). In addition, strong institutions can help avoid the resource curse, whereas weak institutions deteriorate it (Hodler, 2006; Andersen and Aslaksen, 2007; Bhattacharyya and Hodler, 2010). By upgrading the regulation and management of coal mines, coal mining productivity is improved and the crowding-out effect is expected to be relieved. Under this background, by investigating the relief of crowding-out effects, we investigate whether China's coal mine regulation has positively affected the economic growth of regions with high coal industry dependence since the 1990s.

Our analysis is as follows. To guide our empirical investigations, we develop a simple framework that shows the relationship between the resource crowding-out effects of entrepreneurs and regional economic growth. Without reasonable regulation, entrepreneurs tend to flock to the coal mining industry because of high profit; thus, they lower the overall productivity of both the coal mining industry (with economies of scale) and other industries (with economies of agglomeration).

Regarding the empirical strategy estimating the effects of regulation on regional development, first, we apply two difference-in-difference (DID) approaches. The first is a comparison between the pre- and post-regulation periods as well as between regions with and without rich small-size coal mines, and we find that regions with rich small-size coal mines are more affected by the regulation. The second is a comparison between the pre- and post-regulation periods as well as between regions rich in coal mines and those rich in other mines (except coal mines). We find positive economic spillover effects of regulation on both the whole regional economy and the non-coal based economy. Then, using a robustness check of the regulation effects, the key and most difficult procedure is finding a good measurement for coal industry regulation. Because regulation is a complex and comprehensive policy, a computable official indicator is not available. We present arguments showing that mortality rate in coal mining is a suitable proxy for the quality of coal mine regulation. Using provincial-level panel data from 1995–2009, we confirm that coal mine regulation positively improved the economic performance of related regions. The results remain stable under various settings and are strengthened using indicators for wage instead of for per capita GDP. The positive effects are observed not only in the coal industry but also in other economic activities, and we address these spillover effects to the moderation of natural resource crowding-out effects.

This study contributes to recent studies on the relationship between institutions and the resource curse. Although most studies associated with the resource curse and institutions address overall social institutions, this study is concerned with the coal industry; thus, controlling endogeneity is easier, and the institution effect can be more directly observed. Then, although empirical works such as Angrist and Kugler (2008) and Sala-i-Martin and Subramanian (2003) confirm that a relationship exists between poor institutions and poor economic performance in some resource-dependent countries, they did not confirm this using the dynamic improvements of the institutions. In this study, we attempt to empirically show that the improving management of natural resources can positively affect the social economy of related regions.

The remainder of this chapter is organized as follows. Section 4.2 contains a brief review of issues related to the relationship between the resource curse and institutions and those related to the within-country resource curse. Section 4.3 presents an introduction to China's coal industry regulation, and Section 4.4 shows the theoretical framework. Sections 4.5 and 4.6 provide the empirical results on regulation effects. Section 4.7 presents the mechanisms of regulation quality affecting regional development, and Section 4.8 concludes.

4.2 Resource curse

4.2.1 Resource curse and institutions

Natural resource endowments generate economic gains in related regions but may provide few sources of income to local people because of poor resource policies. Accounts of both the relationship between the resource curse and poor social institutions and resource policies are available in many cross- and within-country studies.

Natural resources cause fighting between rivaling groups, which reduces productive activities and weakens property rights, making productive activities less attractive. Natural resources lower incomes in fractionalized countries but increase incomes in homogenous countries (Hodler, 2006). Andersen and Aslaksen (2007) indicate that the so-called resource curse is present in democratic presidential countries but not in democratic parliamentary ones. Bhattacharyya and Hodler (2010) predict that resource rents cause an increase in corruption only if the quality of the democratic institutions is relatively poor. The relationship between resource rents and corruption depends on the quality of the democratic institutions.

Empirically, a series of studies find the positive relationship between poor institution and poor economic performance around the natural resource. An unexpected boom in the demands for coca leaf in Columbia increases its employment opportunity, however, due to the weak institution and civilian conflict, rural regions in Columbia with high coca production become more violent but not go to prosperity (Angrist and Kugler, 2008). Caselli and Michaels (2013) find that living standards have not been improved although revenue and reported government expenditures have greatly increased because of oil windfall in some regions of Brazil. The "missing money" is accounted for by embezzlement surrounding the oil windfall. Also, oil revenues per capita in Nigeria increased by around ten times in the period of 1965-2000, but its income level still drops in the poorest country group in the world. Violence, waste and corruption from oil rather than the Dutch disease has been responsible for the poor long-run economic performance. Improving the quality of public institutions is a method for transforming economics and politics in such regions (Sala-i-Martin and Subramanian, 2003).

By contrast, some researchers discuss some positive experiences. As a large oil exporter, Norway reaches lasting economic growth and least corruption by utilizing well-developed institution (Larsen, 2006; van der Ploeg, 2011). In Botswana, 40% of the GDP stems from diamonds, but it has the second highest public expenditure share on education and enjoys high economic growth (Sarraf and Jiwanji, 2001).

4.2.2 Within-country resource curse

Although the resource curse has been mainly studied using cross-country samples (e.g., Sachs and Warner, 1995, 2001; Gylfason, 2001; Mehlum, Moene, and Torvik, 2006), a substantial amount of literature claims a within-country resource curse. James and Aadland (2011) and Papyrakis and Gerlagh (2007) find a natural resource curse using county-level and state-level data showing that mining is negatively related to economic outcomes across the United States. Deaton and Niman (2012) and James and Aadland (2011) argue that in the United States, an area dependent on coal mining is likely to have deep poverty because of weaker local governance, entrepreneurship, and educational attainment, as well as environmental degradation, poor health outcomes, and limited other economic opportunities. Michaels (2011) finds positive evidence for this point. After oil discovery in the southern United States, related regions reached higher employment density, higher population growth, and higher per capita income.

For the regional level in China, results of resource curse-related works vary across measurement and estimation strategies. Some studies have argued that within China, the resource curse exists at the provincial-level through crowding-out of human resource accumulation and innovation, causing rent-seeking and corruption (Xu and Han, 2005; Hu and Xiao, 2007; Shao and Qi, 2008). Ji, Magnus, and Wang (2012) analyze a cross-province sample of China by focusing on the interplay between resource abundance, institutional quality, and economic growth. They find that resource abundance had a positive effect on economic growth at the provincial-level between 1990 and 2008 that depended non-linearly on institutional quality. Ding, Wang, and Deng (2007), Fang, Ji, and Zhao (2008), and Fan, Fang, and Park (2012) show that there is no significant evidence to support the existence of a resource curse phenomenon in China at either the city level or the prefecture level.

4.3 China's coal mines and coal mine regulation

In the 1980s, to meet the increasing demands of coal in society, the central government of China began encouraging coal mining wherever possible and by whatever means. A coal mining license, once not permitted for non-stated-owned enterprises (non-SOEs), became available to town and village enterprises (TVEs). Thus, TVE coal mines, generally small and with poor production equipment, began flourishing countrywide. China had 84,000 operating coal mines by the end of 1996, 81,000 of which were small-size TVE mines with an annual coal output of less than 30,000 tons. In contrast, the United States, whose total coal production was similar to China's in 1997, had only 2,196 coal mines at that time (Pan, Pu, and Xiang, 2002). In the 1990s, almost 60% of the national coal output in China was generated by small-size TVE mines.

Although TVE coal mines greatly contributed to the energy supply at the time, their disadvantages were obvious. Most TVE mines had less than 10 workers, but coal mining is an industry with economies of scale. Serious problems concerning workers' safety and their long-run health as well as environmental damage and low productivity (i.e., coal resource waste) were highlighted throughout the coal resource-rich regions. The typical exploitation rate of TVE mines was approximately 10%–15%, but some were under 10%. However, the exploitation rate of state-owned key coal mines has generally been approximately 50%. Thus, in TVE mines, each ton of coal mined wastes approximately eight tons of coal and discarded resources that cannot be mined repeatedly (Wang, 2006).

In addition, coal mining licenses have created significant rent-seeking and corruption opportunities for the local economy. Because of the large number of small-size mines, management of the coal industry has become very difficult.

Further, the prosperity of small-size TVE mines indirectly reduces the competitiveness of large-size SOE mines. Coal is a product with no diversity preferences, that is, there is no preference as to whether it comes from a large-size SOE mine with developed facilities and skilled employees or from a small-size TVE mine with rough facilities and unskilled employees: It can be sold to the same customer at the same price. However, although SOE mines have higher labor productivity, they face a higher sunk cost and heavier social burdens (e.g., issues such as recovery of destroyed environment, safety and medical insurance for employees, and education of employees' children). By contrast, TVE mines, particularly illegal mines, can hire enough employees with a relatively low wage and without other social costs. Compared to the high coal price for TVE mines, production costs are relatively low.

As a result, TVE mines flourish but cause serious social problems; however, SOE mines are unprofitable and uncompetitive (non-profitability of SOEs was found across many industries in China in the 1990s. Since then, China has experienced large-scale reform of SOEs countrywide). China's coal industry has become disordered, coal mine accidents are increasingly frequent, and the regional economy has been negatively affected. In the mid-1990s, the central government decided to regulate the coal industry, because most accidents happened in TVE mines where they lacked investments in production safety and where employees mainly comprised the surrounding agricultural population (i.e., they lack knowledge and training in coal mining). Thus, most small-size TVE mines have been gradually closed, and coal mining licenses have been restricted to small-size mining entrants. The remainder of the small-size mines have been required to improve their technology and facilities or to merge with large-size SOE mines.

Policy outcomes have been satisfactory. As discussed in Section 1, the coal mining mortality rate was reduced by 80% and productivity improved by 205% during the period 1997–2009. Although small-size mines have been largely closed, total national coal production has not been reduced; alternatively, it has increased gradually because the exploitation rate in large-size mines has increased. Hence, the degree of industry concentration has increased rapidly. The production share of the largest four or eight enterprises within an industry (C4 and C8, respectively) is a general indicator of industry concentration in a country. The degree of coal industry concentration is high in almost all main coal-producing countries whether they are developed or developing economies: The United States (C4 is 45%), South Africa (C4 is 87%), Germany (C4 is 65%), and Australia (C4 is 50%). In India, more than 90% of total coal is produced by one enterprise (source: Authors' own collection from various sources). Although still low, the degree of industry concentration (C4) in China increased from 6.9% in 1996 to 20% in 2009 (Pan, Pu, and Xiang, 2002). Because coal is crucial to China's energy supply and coal mines in China are only located in specific regions, the economies of these regions are particularly affected by coal. Therefore, strong coal mine regulation and significant changes in the coal industry should be associated with their recent economic performance, making this research interesting and informative.

4.4 Theoretical framework

To provide guidance for our empirical analysis, this section presents a simple framework that formalizes the effects of coal mine regulation on regional development. Following the framework by Torvik (2002), two sectors exist in the region: a coal mining sector (c) and other sectors (o). The total number of entrepreneurs (G) is provided exogenously. Entrepreneurs can choose to engage in one of the two sectors such that $G = G_c + G_o$. G_c is the number of entrepreneurs engaging in the mining sector, and G_o is that engaging in the other sectors.

The mineral reserves (R) are provided exogenously, and the total amount is divided equally to all mining entrepreneurs such that each mining entrepreneur holds the same amount of resources. The total output in the coal mining sector is F_c, which is a function of G_c and R (which is provided), and $\frac{\partial F_c(G_c)}{\partial G_c} < 0$ (in the coal mining sector, a high degree of concentration indicates high industrial productivity (i.e., economies of scale) because of huge sunk costs). The profit for each mining entrepreneur is set to be homogeneous, that is, $\pi_c = \frac{F_c}{G_c}$.

$f_o(G_o)$ is the profit for each entrepreneur engaging in other sectors, and it is a function of G_o. We assume $\frac{\partial f_o(G_o)}{\partial G_o} > 0$, that is, agglomeration economies exist in other sectors, and high firm density indicates high productivity.

The equilibrium is reached when $\pi_c = \pi_o$, that is, for each entrepreneur, no profit difference is seen between the two sectors, $\frac{F_c}{G_c} = f_o$.

We set the total social profit function (i.e., the total profit of both sectors) as

$$\text{Total Profit (TP)} = F_c(G_c) + f_o(G_o) \times G_o.$$

The equilibrium is thus

$$\text{TP} = f_o(G_o) \times G_c + f_o(G_o) \times G_o = f_o(G_o) \times G.$$

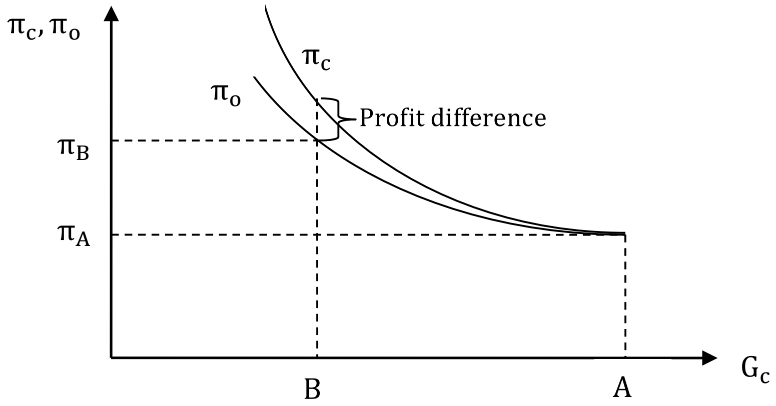


FIGURE 4.1: The relationship between entrepreneurs and productivity

The first-order partial differential to G_c is

$$\frac{\partial \mathrm{TP}}{\partial G_c} = \frac{\partial f_o(G_o)}{\partial G_c} = -f_o^{'}(G_o) < 0.$$

Proposition (relief of crowding-out effects): When the central government chooses to close small-size coal mines, coal mining entrepreneurs move from the coal mining industry to other sectors (i.e., G_o increases), the total profit of both sectors increases, and the relief of crowding-out effects improves the overall social output.

As shown in Figure 4.1, with no restrictions on G_c, entrepreneurs naturally move from other sectors to coal mining because π_c is always larger than π_o until point A, $\pi_o = \pi_c = \pi_A$. The social welfare is $\pi_A \times G$.

When the government places restrictions on G_c (point B in Figure 4.1), G_o is fixed as well. Because the profit of coal mining is higher than other sectors at point B, profit differences exist attributed to tax, markup, or corruption. In sum, the social welfare becomes $\pi_B \times G > \pi_A \times G$. We can observe an increase in social welfare by closing small-size mines (although we assume that size is homogeneous in the model, real coal mining firms are of different sizes and small-size mines are closed first when G_c is restricted. This does not affect the conclusion: Crowding-out effects are relieved when the industrial concentration degree of the coal mining industry is increased). Ideally, coal mining should be undertaken by a single firm to reach the highest productivity. However, because of geographical features, the coal mining industry may not be highly concentrated in China.

This setting is reasonable and may be explained by Figure 4.2. After the 1980s, easing the policies on TVE coal mines caused the number of TVE mines to increase rapidly in China, peaking at approximately 100,000 TVE mines in 1991. In 1997, the central government began rectifying TVE coal mines that failed to meet basic safety standards,

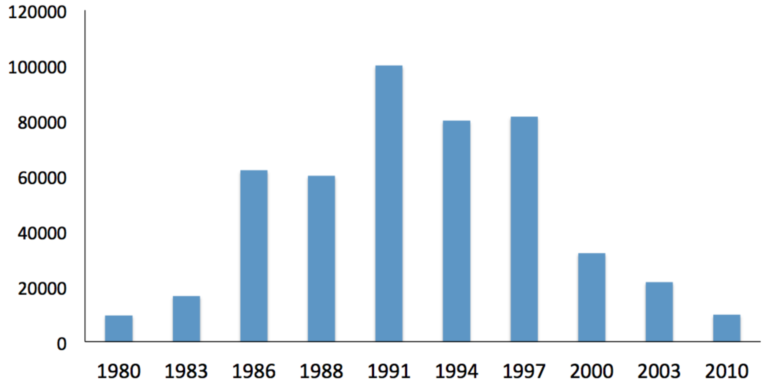


Figure 4.2: The number of TVE coal mines in China

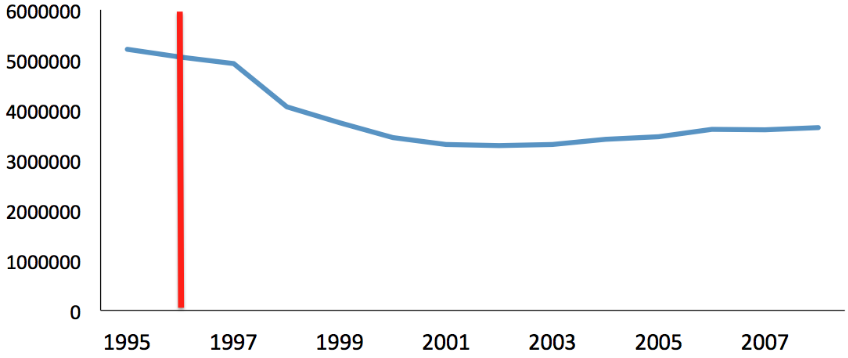


Figure 4.3: Employment in China's coal mining industry
Note: For most small-size coal mines, employment is not counted in the statistical yearbook; thus, after regulation, the actual reduction in coal mining employment should be much greater.

and over 10,000 mines were closed that year. The government has increased its efforts since then, requiring all unregistered illegal mines to be shut down. By the end of 2010, the number of TVE mines had been reduced to approximately 10,000. Thus, over 85% of small-size coal mines have been closed since 1997. Furthermore, employment in the coal mining industry decreased sharply in the regulated period (Figure 4.3). Therefore, the setting of G_c is acceptable.

Another important issue concerning the model is the firm size and economies of scale in coal mining. In the model, we set the productivity in coal mining to be positively correlated with firm size. Coal mining is an industry with economies of scale, and, intuitively, small-size mine owners should have the incentive to merge and reach higher productivity and profits. However, typical cases in China do not follow this logic, possibly for four reasons.

First, there is no incentive for mine owners to put considerably investments in coal mining because of an unstable coal mining policy. The mining license might be withdrawn by the government, and mine owners are therefore more concerned with short-term profits rather than long-term ones. Thus, they try to maximize exploitation while the mines are still in operation. Second, coal mining requires little knowledge and few skilled workers. Therefore, employees are pulled from the agricultural population, that is, mining wages may be rather low[2], and labor substitutes for advanced technology. Third, unlike other industries, the investment threshold for mining is too high for private capital: Even if two or three small-size mines are merged, the investment threshold cannot typically be met. Generally, only the large-size SOE mines can afford the total technology investment. Furthermore, after a large investment, mines require advanced management. Most mine owners cannot manage this effectively because of insufficient management knowledge. Additionally, returns on coal mine investment are very slow and accompanied by high risk. Fourth, there is no difference between the coal products produced in SOE mines and those in TVE mines, and thus, small-size TVE mines are already competitive in the market. Therefore, small-size mine owners tend to exploit the coal resource-extensively but not resource-intensively and labor-intensively but not capital-intensively.

4.5 Empirical results

4.5.1 Empirical strategy

In the empirical section, we use two DID approaches to estimate the regulation effects on regional development. For the first DID approach, we are concerned with the regulation effects on the whole regional economy. We compare economic performance before and after regulation to estimate the effects of policy changes in the coal mining industry between regions with many or few small-size coal mines. The role of regulation may vary in these regions: Regions with many small-size coal mines (determined by coal endowment and local geology) are affected by regulation more than other regions are because regulation is mainly targeted at small-size mines. For the second DID approach, we focus on the regulation effects of the crowding-out of the coal mining industry on the non-coal based economy. we compare the tertiary industry (non-coal based) performance before and after regulation between regions with rich coal mine endowments and those with other rich natural resource endowments. Both regions are highly dependent on natural resources and have a similar regional industrial structure; however, only coal mines were strongly regulated during the study period.

---

[2] One picture taken in 2007 shows a typical case of coal mining in China. A 17-year-old miner carries over 50 kg of coal for 1 km. For each trip, he is paid 1 yuan (approximately 0.16 USD), which shows that despite heavy work, the wage is low. (http://blog.sciencenet.cn/home.php?mod=space&uid=4699&do=blog&id=33987)

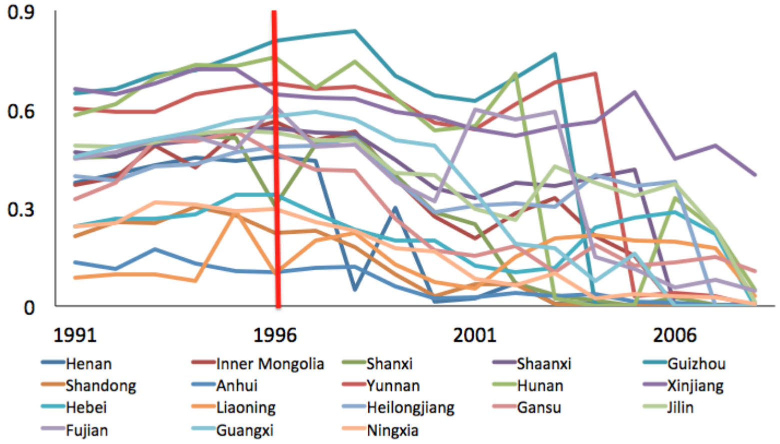


Figure 4.4: The product share of small-size coal mines in total coal production

4.5.2 Coal mine regulation and the whole regional economy

We cannot observe the exact number of existing small-size coal mines, and the heterogeneity between mines is large. Alternatively, we use the output of small-size coal mines to reflect the approximate number of small-size coal mines. Because the central government decided to close small-size mines nationwide in 1997, we set this as the starting point of the regulation policy.

We test the coal mine regulation effects using the DID approach. According to the official classification, small-size coal mines mainly refer to TVE mines with an annual output of less than 30,000 tons. Most coal mine accidents occurred in small-size mines, and, thus, the central government decided to gradually close these mines or require them to integrate into large-size mines. Because we use a provincial-level dataset and because the geological heterogeneity among provinces is large, some regions are able to develop large-size mines whereas others cannot because of natural reasons. That is, the output amount of small-size mines can be viewed as exogenous before regulation (i.e., before 1997), but it has been greatly affected by regulation (i.e., since 1997). The total production of small-size coal mines has greatly decreased in the past decade. In 1996, the national output was 59,000 tons. By 2008, the output of small-size coal mines decreased to 6,000 tons. Figure 4.4 shows that before 1997, the total coal production product share of small-size mines was stable among provinces. After 1996, when the regulatory policy on coal mines was enacted, most provinces closed, upgraded, or merged their small-size mines, and the product share for small-size mines was substantially reduced. Thus, stable output of small-size mines before the regulation and reduced output after regulation provide an exogenous source for checking coal mine regulation effects on economic growth. That is, regions with high output of small-size mines before the regulation are expected to be more heavily affected by the regulation.

Governments in regions with high output of small-size mines prioritize regulation, and thus, these regions are expected to benefit more from this policy change. That is, we estimate the following equation. The annual GDP per capita growth (i.e., the deflated annual GDP per capita index) in province i and year t is a function of the dummy variable featuring the output of small-size coal mines (=1 if a province had high per capita small-size coal mine output in 1991–1996 (before regulation)), the interaction term between the small-size coal mines dummy and the post-regulation dummy variable (=1 if year >1996), and controls such as coal dependence in the related province and year. η_i and μ_t are the province and year dummies, respectively. If the regulation policy improved economic performance, there should be a difference before and after regulation, and α_2 is expected to be positive.

$$\begin{aligned} GDP_{i,t} = \alpha_0 + \alpha_1 \mathit{Small\text{-}size}_{i,t} + \alpha_2 \mathit{Small\text{-}size}_{i,t} \times \mathit{Post}_t + \alpha_3 \mathit{CoalDependence}_{i,t} \\ + \alpha_4 \mathit{Controls}_{i,t} + \eta_i + \mu_t + \epsilon_{i,t} \end{aligned}. \quad (4.1)$$

We use annual data from 1986–2008 in 18 sample provinces, and we include all main coal producers in China. We choose nine provinces with high per capita small-size coal mine output before 1997 as the treated samples (Inner Mongolia, Shanxi, Shaanxi, Guizhou, Hunan, Xinjiang, Heilongjiang, Jilin, and Ningxia) and the other nine provinces with low per capita small-size coal mine output as the control (Henan, Shandong, Anhui, Yunnan, Hebei, Liaoning, Gansu, Fujian, and Guangxi). Under this setting, we guarantee that all of our samples are heavily dependent on the coal economy but only the treated samples are heavily affected by coal mine regulation. Figure 4.4 shows that during 1997–2008, most provinces almost shut down all of their small-size mines. However, the product share of small-size mines in total coal production differs among provinces: Some provinces immediately started to regulate small-size mines in 1997, whereas others started later, for example, after 2000. This suggests that a long period (1986–2008) is required to observe the regulation effects. Columns 1–3 in Table 4.1 show that under different fixed-effect settings, small-size mine dummies are not significantly associated with economic performance; however, the interaction terms with the post-regulation dummy are positively significant. Evidently, there is a difference before and after regulation. As shown in column 1, on average, regulated provinces have 2.67% higher real GDP per capita growth. Compared to the average annual real GDP per capita growth in our samples (10.22% before regulation and 11.04% post regulation), regulation considerably increased overall economic performance.

However, during the regulation period, macroeconomic policy for inland China underwent significant changes. During the high growth period of the 1990s, China's coastal regions gained significantly more than inland regions because of the "Reform and Opening Up" policy. To relieve the widening regional disparities, three regional development plans, the "Go West Campaign," "Revitalizing Northeast China," and "the Rise of Central China" were gradually issued and enacted after 1999. These plans aimed to stimulate development of inland regions and to moderate the large income gap between coastal and inland China. These strategies proved to be efficient, and the overall GDP growth of inland regions began overtaking that of coastal regions after 2004.

Because most of our sample provinces are located in inland regions, it is necessary to pay attention to this exogenous variation. We conduct two types of estimations. First, in columns 4–6, we insert three regional development policy dummies for specific provinces (for West, Central, and Northeast; classification is presented in the Appendix) and years: the "Go-West Campaign" from 2000 (post2000 = 1 if year >2000), "Revitalizing Northeast China" from 2003 (post2003 = 1 if year >2003), and "the Rise of Central China" from 2006 (post2006 = 1 if year >2006). Although Table 4.1 shows that the concerned coefficients became smaller, the regulation effect on annual per capita GDP growth shown in column 4 is 2.50% (slightly smaller than 2.67% in column 1); thus, it is still significant. Then, in columns 7–8, we use the annual real wage growth index (deflated by the provincial consumer price index) instead of the annual GDP per capita growth index as the indicator and run estimations with two-way fixed effects. This is because the regional development policies took effect mainly using investments, for which GDP per capita growth is more sensitive than real wage growth is. We find that the wage effect is smaller but still significant. Regulated provinces have 1.76% higher annual real wage growth (column 7), a figure that increases to 1.82% after controlling for the three regional development policies (column 8).

Finally, common trend between treatment and control groups, which is required for an effective difference-in-difference approach, is guaranteed by placebo tests. We are concerned on the common trend before the coal mine regulation in 1996, that is, in the period of 1986-1996. We assume the economic performance between treatment and control regions should not be systematically different, columns 1-2 of Table C.1 give the related evidences. We set a counterfactual treatment cut point in 1990, that is, we assume a placebo coal mine regulation was issued in 1990. Thus the time period of 1986-1990 (1991-1996) is seen as the pre-treatment (post-treatment). By apply the identical estimation strategy, our results are compatible to the baseline estimates. In Column 1 of Table C.1, the coefficient of intersection term shows to be negative and statistically significant, which is opposite with the sign in baseline estimates. This suggests the treatment group actually has poor economic performance before the regulation treatment. This is reasonable, as we argued in the first section, regions with rich coal endowment tended to have poor economic performance, partially caused by the weakly ordered coal industry in the 1980s and early 1990s. Hence the coal mine regulation was issued since 1996. Then, column 2 gives the estimates in real wage, the interested coefficient comes to be insignificant now.

TABLE 4.1: DID approach with respect to whole regional economy

	(1) GDP	(2) GDP	(3) GDP	(4) GDP	(5) GDP	(6) GDP	(7) Wage	(8) Wage
Small-size × Post	2.669***	2.312***	2.780***	2.499***	2.077***	2.640***	1.760*	1.816*
	(0.693)	(0.806)	(0.739)	(0.621)	(0.661)	(0.675)	(1.007)	(1.000)
Small-size			-2.935***			-2.868***		
			(0.954)			(0.858)		
Post	-1.507**	-1.302*	1.969***	-1.424**	-1.619**	1.572**	-11.42***	-11.46***
	(0.582)	(0.727)	(0.727)	(0.563)	(0.628)	(0.729)	(1.555)	(1.588)
CoalDependence	1.048	3.574***	0.264	1.120	3.068***	0.195	-0.133	-0.146
	(1.049)	(0.681)	(0.558)	(1.081)	(0.763)	(0.489)	(0.711)	(0.821)
West × Post2000				0.526	0.874***	0.636		-0.441
				(0.662)	(0.406)	(0.643)		(0.978)
Northeast × Post2003				1.780**	3.828***	1.195*		1.109*
				(0.836)	(0.611)	(0.690)		(0.626)
Central × Post2006				0.285	1.242	0.222		1.366*
				(0.719)	(0.726)	(0.705)		(0.677)
Constant	99.48***	77.17***	108.0***	98.83***	81.84***	108.6***	122.4***	122.5***
	(9.726)	(6.307)	(4.972)	(9.937)	(7.073)	(4.414)	(7.059)	(8.023)
Year FE	yes	no	yes	yes	no	yes	yes	yes
Province FE	yes	yes	no	yes	yes	no	yes	yes
N	414	414	414	414	414	414	360	360
R^2	0.577	0.175	0.574	0.583	0.209	0.579	0.615	0.615

Note: Heteroskedasticity-robust standard errors are reported in parentheses, and standard errors are clustered at the province level. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. Years sampled: 1986–2008.

4.5.3 Coal mine regulation and non-coal based economy

Based on the theoretical framework, we show that the regulation on the coal mine sector relieves the crowding-out effects of entrepreneurs in other sectors. Similar to the number of small-size mines, the exact number of entrepreneurs in each sector is not available, and we cannot distinguish between coal-based and non-coal-based economies. Alternatively, we use the tertiary industry to represent non-coal based industries (in the theoretical model, this is related to the “other sectors”) and determine the crowding-out effects.

Specifically, we apply this strategy using data of the performance of the tertiary industry on the prefecture and city levels. Three considerations exist: First, ordinary cities significantly differ from mine cities and cannot perform as a suitable control in DID estimates. Therefore, other mine cities with rich oil, metal, or non-metallic minerals are suitable candidates for controls because they have a similar industrial structure to coal mine cities, that is, high dependency on natural resource (coal based or other natural resource-based industries). Second, other mining industries were not regulated during our study period, providing a chance to observe the policy change effects using the DID approach. Identification of mine cities is based on China’s mining city database (http://www.chinamining.com.cn/city/city.asp), and we selectively exclude mine cities in which mined resources were almost completely exploited by 1999. Based on the database, we choose 24 prefecture-level coal mine cities as the treated samples and the other 36 mine cities as the control samples. Third, at the prefecture level, the economic performance of prefectures rich in natural resources are highly dependent on resource-related industries. Thus, a direct observation of GDP growth is not suitable for checking for coal mine regulation effects (e.g., the crowding-out effects), because GDP in these prefectures include substantial mining and mine-related secondary industries. We conduct estimations on the tertiary industry performance with respect to regulation effects (i.e., the relief of crowding-out effects). Table 4.2 shows the DID results with annual data for 1988–2007 and the following estimated equations.

$$Tertiary_{i,t} = \alpha_0 + \alpha_1 CoalCity_i \times Post_t + \alpha_2 CoalCity_i + \mu_t + \epsilon_{i,t}. \tag{4.2}$$

$\text{Tertiary}_{i,t}$ is the value-added growth of the tertiary industry in city i and year t, Post_t is a dummy variable indicating pre- and post-1996, and CoalCity_i is a dummy variable for whether the related city is a coal mine city. In the interaction term between Post_t and CoalCity_i, the coefficients presented in Table 4.2 are consistently positive and significant. This provides evidence that the non-resource intensive industries benefited from coal mine regulation in related regions. In columns 1–3, we conduct estimations with different fixed effects, and in columns 4–6 and 7–9 we insert a coastal dummy (=1 if the sample prefecture is located in coastal regions) and three regional development policy dummies, respectively, to control for variation in macroeconomic policy changes (equal to the dummy setting in Table 4.1). After controlling for the coastal dummy and for the three regional development policy dummies, we find that the concerned coefficients, both their absolute value and significance, are almost unchanged. Similarly, we present a placebo test in column 3 of Table C.1 as estimates at the provincial level. By setting a placebo treatment cut point in 1990, and restricting the sampled periods no later than 1996, we do not find significant difference in the performance of tertiary industry between treatment and control cities.

4.6 Robustness checks using mortality rate

The DID approaches may provide straightforward evidence of regulation effects. However, these approaches cannot offer evidence on the effects of the degree of regulation on economic performance. We present another estimate in this section as further evidence. One problem arises. There is no comprehensive indicators representing regulation, because regulation is a complex policy mainly (but not limited to) addressing the shuttering of small-size mines. Alternatively, we use the mortality rate in coal mining as a proxy of overall regulation quality.

Table 4.2: DID approach with respect to the non-coal based economy

	(1) Tertiary	(2) Tertiary	(3) Tertiary	(4) Tertiary	(5) Tertiary	(6) Tertiary	(7) Tertiary	(8) Tertiary	(9) Tertiary
CoalCity × Post	0.028* (0.016)	0.028* (0.016)	0.032** (0.016)	0.029* (0.016)	0.028* (0.016)	0.032** (0.016)	0.028* (0.016)	0.028* (0.016)	0.032* (0.016)
CoalCity	-0.011 (0.012)			-0.010 (0.012)			-0.011 (0.012)		
Post	-0.041* (0.022)	-0.034 (0.029)	-0.103*** (0.013)	-0.044* (0.023)	-0.036 (0.029)	-0.104*** (0.014)	-0.030 (0.021)	-0.025 (0.030)	-0.104*** (0.012)
Coastal × Post				0.008 (0.019)	0.007 (0.019)	0.005 (0.019)			
Coastal				0.007 (0.014)					
West × Post2000							-0.025 (0.018)	-0.019 (0.023)	-0.001 (0.020)
Northeast × Post2003							-0.059 (0.046)	-0.050 (0.063)	-0.031 (0.068)
Central × Post2006							-0.002 (0.028)	-0.002 (0.028)	0.079*** (0.024)
constant	0.229*** (0.015)	0.218*** (0.020)	0.215*** (0.006)	0.228*** (0.016)	0.218*** (0.020)	0.215*** (0.006)	0.229*** (0.015)	0.218*** (0.020)	0.214*** (0.006)
Year FE	yes	yes	no	yes	yes	no	yes	yes	no
City FE	no	yes	yes	no	yes	yes	no	yes	yes
N	939	939	939	939	939	939	939	939	939
R^2	0.278	0.278	0.102	0.278	0.278	0.103	0.278	0.278	0.110

Note: Heteroskedasticity-robust standard errors are reported in parentheses, and standard errors are clustered at the prefecture level. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. Years sampled are 1988–1991&1995–2007. In 1992–1994, the statistical yearbook did not include the related data. We exclude the samples with missing data

4.6.1 Coal mine regulation measurement: why mortality rate is a suitable proxy

Coal mine regulation is a comprehensive policy that includes cancellation or rejection of coal mining licenses to small-size mines with insufficient investment and low annual output, closing illegal mines and most existing small-size mines with low productivity, improving mining technology, and enhancing the standard of compensation for mining accidents. Because a computable official indicator is not available, we present the argument that mortality rate is a potential suitable indicator for the quality of regulation because it is highly correlated with the main regulation events. Mortality rate is highly associated with investment improvements in the coal mine industry, an increase in standards of compensation for mining accidents, and the closure of small-size mines.

Improved mining technology and investments not only enhance productivity but also better guarantee safety in coal mining. High compensation standards for mining accidents force mine owners to invest more in safety facilities and technical training of employees as well as to further relieve the frequency of mining accidents. Before 2000, the owners of TVE mines paid limited reparations for mining accidents, generally less than US$ 5,000 to each of the victim's families. Compared to the revenue gained from mining, reparations were very affordable, that is, there was no incentive for mine owners to improve production technology and safety or to provide vocational training for mine employees. Furthermore, because of the low education levels, most employees in TVE mines could not effectively protect their rights after accidents using the judiciary system. Post-regulation (though still far from sufficient compared to developed countries), compensation and safety supervision have significantly increased: Compensation for mining accidents to each victim's families was approximately US$ 30,000 in 2004, US$ 40,000–60,000 in 2009, and US$ 90,000–100,000 in 2011[3]. Consistently, China's coal mining mortality rate declined rapidly, for a reduction of 80% from 2000–2009. Thus, compensation standard has a negative relationship with mortality rate.

To regulate the coal industry, the Chinese government has withdrawn coal mining licenses for many small-size mines, particularly small-size mines with annual coal output of less than 30,000 tons. Furthermore, the government stopped issuing licenses to new small-size mining entrants in the mid-1990s. During the period 1995–2010, over 85% of TVE mines were shut down (see Figure 4.2), most of which were small-size mines. In 1995, small-size mines accounted for 43% of national coal output, but this figure decreased by 2009 to less than 3%. The process of shutting down small-size mines is a recent main characteristic of the Chinese coal industry. Consistently, in this period, mortality rate was highly associated with the shuttering policy, decreasing from 4.89 to 0.89 person/million tons coal (see Figure 4.5). Mortality rates have varied greatly across types of mines, and most accidents have happened in small-size mines. Whereas the safety record for large-size SOE mines is as good as their counterparts in advanced countries, small-size mines are potential death traps (Wang, 2006). In 1995, 10,572 people in China died in mining accidents, 70% of which occurred in TVE mines.

4.6.2 OLS relationship between mortality rate and GDP growth

We estimate the regulation effects using the mortality rate in coal mining as a proxy of regulation quality. Following the general setup of the estimate on economic growth, we use the control variables of social investment, labor resources, and technology. Coal dependence varies among provinces, so we add controls for the region's coal dependence. Specifically, we model the real GDP per capita growth rate on the mortality rate (i.e., the logarithm of the mortality rate in coal mining), coal dependence (i.e., the logarithm

---

[3] Data is from XinHua Net for different accidents:
http://news.xinhuanet.com/newscenter/2005-01/30/content_2525350.htm;
http://www.yn.xinhuanet.com/newscenter/2009-12/30/content_18633383.htm;
http://www.yn.xinhuanet.com/newscenter/2011-04/08/content_22472427.htm.
Although there is no official standard or detailed record for compensation standards of mining accidents and they differs between regions, the type of the mine, and social consequences, the compensation standards increase based on coal mine regulation

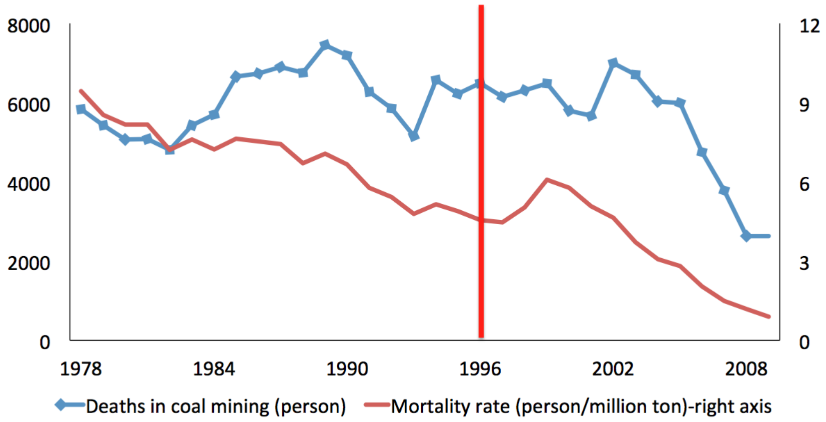


FIGURE 4.5: The mortality rate of coal mining in China

of coal production per capita), and other controls: investment (i.e., the share of total fixed-asset investments in the GDP), human resources (i.e., the share of the population with an educational level higher than senior high school in the provincial workforce), technology (i.e., using the logarithm of foreign direct investment per capita as a proxy). Furthermore, the logarithm of initial real GDP per capita to the control converges across provinces.

We employ provincial panel data for the 15 years 1995–2009. Because not all provinces have coal mines, we choose 20 provinces and provincial-level cities (18 provinces mentioned in Section 5.1 plus Sichuan and Chongqing) of mainland China[4] based on the coal mine locations and data availability. In total, these 20 locations account for 93.2% (98.2%) of total coal production in 1995 (2009). To mitigate the noise caused by the volatility of the year-after-year short-term economic growth and short-term mortality rate, we set the panel data by using five time series (i.e., 1995–1997, 1998–2000, 2001–2003, 2004–2006, and 2007–2009) and by employing the mean of the indicators into the regression. Details of the dataset are presented in the Appendix. The panel regressions include both provincial and time dummies to control for heterogeneity denoted by η_i and μ_t, respectively. The baseline setup is as follows:

$$\Delta \ln(GDP)_{i,t} = \alpha_0 + \alpha_1 Mortality_{i,t} + \alpha_2 CoalDependence_{i,t} + \alpha_3 Controls_{i,t} + \eta_i + \mu_t + \epsilon_{i,t}. \tag{4.3}$$

The benchmark result is presented in column 1 of Table 4.3. The mortality rate proves to be negatively associated with GDP growth. The degree of coal dependence varies widely among provinces. To control for geographical heterogeneity and to determine whether the benchmark result is driven by specific province features, we exclude Shanxi, Inner Mongolia, Ningxia, Xinjiang, and Heilongjiang, which were the five provinces highest

---

[4] We exclude Sichuan and Chongqing in the DID approaches (Section 5.1), because Chongqing was part of Sichuan before late 1990s; thus, separate data on these two regions before 1995 is not available.

in coal dependence during 1995–1997. Mongolia and Shanxi have most large opencast coal mines in China. An opencast coal mine is naturally safer than underground mines, and mine accidents are easier to mitigate. Unsurprisingly, the estimated coefficient (column 2 of Table 4.3) is reduced by 40% because these five excluded provinces are China's main coal producers, accounting for 40.0% (50.4%) of national coal output in 1995 (2009). However, the coefficient is still significant at the 1% level. Next, we must consider that the coal industry mortality rate differs naturally among provinces because of the geological features of the mines. Mining is considerably more dangerous in coal mines deep under the surface of the earth than in opencast coal mines.

TABLE 4.3: Regressions with mortality rate

	(1) $\Delta\ln$(GDP) baseline	(2) $\Delta\ln$(GDP) ex top 5 in coal output	(3) $\Delta\ln$(GDP) ex top 5 in mortality rate	(4) $\Delta\ln$(GDP) with interaction term	(5) $\Delta\ln$(GDP) with regional policy dummy	(6) $\Delta\ln$(GDP) initial value	(7) $\Delta\ln$(Wage)
Mortality	-0.016***	-0.009***	-0.019***	0.010	-0.015***	-0.014**	-0.016***
	(0.005)	(0.002)	(0.005)	(0.013)	(0.005)	(0.006)	(0.005)
CoalDependence	-0.000	-0.004	-0.003	0.012	0.001	0.003	-0.015
	(0.008)	(0.006)	(0.008)	(0.009)	(0.008)	(0.008)	(0.009)
Mortality × CoalDependence				-0.0025*			
				(0.0014)			
InitialGDP	-0.122***	-0.190***	-0.109***	-0.147***	-0.126***	-0.107***	-0.070
	(0.035)	(0.061)	(0.035)	(0.033)	(0.037)	(0.034)	(0.080)
Educ	0.000	0.001	-0.000	0.001	0.000	-0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Invest	0.072***	0.030	0.084***	0.069***	0.075***	0.082***	0.064*
	(0.026)	(0.030)	(0.032)	(0.027)	(0.027)	(0.028)	(0.037)
FDI	0.008	0.007	0.004	0.007	0.009	-0.000	0.008
	(0.006)	(0.008)	(0.008)	(0.006)	(0.006)	(0.003)	(0.007)
West × Post2000					-0.001		
					(0.006)		
Northeast × Post2003					0.002		
					(0.005)		
Central × Post2006					-0.013*		
					(0.007)		
Constant	1.109***	1.646***	1.085***	1.179***	1.114***	1.108***	0.747
	(0.267)	(0.476)	(0.282)	(0.249)	(0.279)	(0.345)	(0.607)
Year sample: 1995–2009 (three years per period); With two-way fixed effects							
N	100	75	75	100	100	80	100
R^2	0.809	0.849	0.835	0.817	0.819	0.836	0.792

Note: Heteroskedasticity-robust standard errors are reported in parentheses, and standard errors are clustered at the province level. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. Column 6 refers to the period 1998–2009 (four periods) because of data restrictions.

To control for the heterogeneity in the mortality rate, we exclude Guizhou, Hunan, Sichuan, Chongqing, and Xinjiang, the five provinces and provincial-level cities with the highest mortality rate during 1995–1997. The results presented in column 3 are almost unchanged, the coefficient of mortality rate is still significant at the 1% level, and the magnitude becomes slightly larger. Furthermore, we insert an intersection of mortality rates and coal dependence into the regression in column 4. Thus, the coefficient of the mortality rate becomes insignificant, and the coefficient of the intersection becomes significant. Therefore, reasonably, the effects of coal mine mortality rates on GDP growth heavily depend on a region's coal dependence, which we can interpret as additional evidence that the benchmark result is not endogenously caused by disturbances not associated with the coal industry.

In column 5, we address the robustness with respect to the newly issued regional development policies. Results in column 5 remain consistent. Then, in column 6, we use the initial value of all control variables instead of the mean values to check these results, and they are almost unchanged. Finally, in column 7, we use the real wage growth instead of GDP per capita growth as a robustness check, and the results are consistent.

The potential endogeneity of the mortality rate to GDP growth must be controlled. First, GDP growth rate is an important measuring index for China's local government performance; therefore, if the economic performance is poor, the government may overlook safety inspections in social production, whereas good economic performance encourages the local government to more closely monitor production safety. This criticism has always existed in developing economies, that is, governments tend to rapidly develop their economy at the cost of high environmental pollution and frequent production accidents. In this case, the inference previously mentioned might be biased.

We use another indicator of mortality rate to verify this concern. If a mortality rate drop in the coal industry is attributed to the improving social economic performance, then the safety of all industries improves but is not limited to coal industry. That is, in other industries, similar empirical results between the mortality rate and real GDP per capita growth may be obtained. We employ the total mortality rate in non-agricultural industries (i.e., industry, mining, construction, commerce, and trade) as a substitute for mortality rates in the coal industry, and we run the same regression. Column 1 of Table 4.4 shows that the result is significant, which potentially suggests that the benchmark result is biased. However, the coal industry accounts for a high share of total mortality accidents among the non-agricultural industries. In the mid-1990s, victims in coal mines accounted for over 50% of total mortality accidents in social production, and coal mining became the highest-risk occupation. The significant result in column 1 may be caused by the large proportion of the coal industry for this indicator. In column 2, we insert both mortality rate indicators into the regression; thus, the mortality rate in all non-agricultural industries becomes insignificant, whereas the mortality rate for coal mining is still significant. In column 3, we exclude the coal industry mortality from the total non-agricultural industry mortality, and the modified mortality rate becomes insignificant. That is, the significant result found in column 1 can be mainly attributed to the coal industry. For other industries, the mortality rate has no significant effects on provincial

GDP per capita growth. Finally, in column 4, we test with the real wage as a robustness check. The result with respect to the mortality rate (excluding the coal mining industry) is also insignificant.

TABLE 4.4: Regressions with additional mortality rates

	(1) $\Delta\ln$(GDP)	(2) $\Delta\ln$(GDP)	(3) $\Delta\ln$(GDP)	(4) $\Delta\ln$(Wage)
MortalityCoal		-0.016**		
		(0.007)		
MortalityTotal	-0.0198*	-0.004		
	(0.0104)	(0.012)		
MortalityTotalExcoal			-0.003	0.006
			(0.014)	(0.012)
InitialGDP	-0.069	-0.137***	-0.080**	0.012
	(0.044)	(0.053)	(0.041)	(0.034)
Invest	0.052*	0.054*	0.049	0.043
	(0.031)	(0.028)	(0.031)	(0.029)
Educ	0.001	0.001	0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.002)
FDI	0.006	0.009	0.006	-0.009
	(0.009)	(0.009)	(0.010)	(0.011)
Constant	0.073*	1.301***	0.777*	0.028
	(0.042)	(0.466)	(0.403)	(0.275)
Year sample: 2001–2009 (three years per period); With two-way fixed effects				
N	60	60	60	60
R^2	0.734	0.773	0.719	0.201

Note: Heteroskedasticity-robust standard errors are reported in parentheses, and standard errors are clustered at the province level. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. Data refers to the period 2001–2009 (three periods) because of data restrictions.

4.7 Mechanism

Based on the previous sections, we can suppose that coal industry regulation has a positive effect on the economic growth of regions with high coal dependence. Two mechanisms can be used to interpret this result. First, productivity of the coal mine industry is improved because of increased investment in fixed assets and economies of scale originating from industrial concentration. Second, the crowding-out effect is potentially relieved, because coal mining licenses in TVE mines are, to some extent, prohibited. Even after obtaining the license, environmental recovery and technological investment are strongly required, and production safety is inspected more seriously. In this case, private capital in the coal industry may transfer to other industries that can stimulate local labor resource accumulation and strengthen agglomeration economies.

Primary mining is an industry with low labor resource accumulation, and the social total factor productivity (TFP) growth is heavily dependent on labor resource accumulation. In this study, we are concerned more with the second mechanism, which might improve overall economic performance.

For the first mechanism, Table 4.5 shows that after controlling for the initial labor productivity of the coal industry, the product share of small-size mines (annual output: less than 30,000 tons) and middle-size mines (annual output: 30,000–300,000 tons), the investment in the coal industry, and the skill of employees (here, we take the relative wage as the proxy of skill: average wage of coal mining/average wage of all industries), the mortality rate is still highly associated with coal mine labor productivity whether the independent variable is based on the mean value (column 1) or the initial value (column 2).

Table 4.5: Mechanism (intra-industry)

	(1) LabProd	(2) LabProd
Mortality	-0.248***	
	(0.073)	
MortalityInitial		-0.132*
		(0.073)
SmallSizeMines	0.211	0.211
(less than 30,000 tons)	(0.192)	(0.182)
MiddleSizeMines	0.128	0.206
(30,000-300,000 tons)	(0.210)	(0.234)
LabProdInitial	0.295***	0.310***
	(0.082)	(0.097)
InvestCoal	0.205***	0.247***
	(0.045)	(0.040)
Skill	0.166	-0.006
	(0.259)	(0.200)
Constant	4.145***	4.132***
	(0.912)	(1.041)
Year sample: 1995–2009 (three years per period);		
With two-way fixed effects		
N	100	80
R^2	0.892	0.866

Note: Heteroskedasticity-robust standard errors are reported in parentheses, and standard errors are clustered at the province level. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. Column 2 refers to the period 1998–2009 (four periods) because of data restrictions.

Thus, we try to distinguish the effects of regulation on the relief of crowding-out effects between intra-industry productivity improvements. Regarding the estimation of the crowding-out effects of coal resources, it is impossible to find a suitable measurement directly. In the implementation, we first insert the labor productivity of the coal industry

Table 4.6: Mechanism (whole economy)

	(1) Δ ln(GDP)	(2) Δ ln(GDP)Excoal	(3) Δ ln(GDP)
Mortality	-0.014***	-0.017***	-0.019***
	(0.004)	(0.006)	(0.006)
LabProd	0.014**		
	(0.007)		
CoalDependence	-0.008	-0.006	-0.004
	(0.008)	(0.005)	(0.006)
InitialGDP	-0.125***	-0.120***	-0.127**
	(0.032)	(0.047)	(0.057)
Educ	0.000	-0.001	-0.000
	(0.001)	(0.001)	(0.001)
Invest	0.068***	0.078**	0.073**
	(0.026)	(0.033)	(0.034)
FDI	0.007	0.004	0.008
	(0.006)	(0.006)	(0.007)
Constant	1.132***	1.188***	1.209***
	(0.250)	(0.343)	(0.406)
Year sample: 1995-2009 (three years per period);			
With two-way fixed effects			
N	100	80	80
R^2	0.817	0.789	0.834

Note: Heteroskedasticity-robust standard errors are reported in parentheses, and standard errors are clustered at the province level. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively. Columns 2 and 3 refer to the period 1995–2006 (four periods) because of data restrictions.

into the baseline setup. In column 1 of Table 4.6, the mortality rate coefficient remains significant and the magnitude only slightly decreases after controlling for coal mining labor productivity. Thus, intra-industry productivity improvement may not fully explain the coal mine regulation effects on regional economic growth, because otherwise, the coefficient of the mortality rate should become much smaller. The results of column 2 further confirm this. We calculate the modified real GDP per capita growth rate by excluding the coal industry value-added in the total GDP. The estimated result supports this assumption. Even without considering the coal industry, the mortality rate is still highly associated with GDP per capita growth. Compared to the benchmark estimates using the original GDP per capita growth rate in column 3, the magnitude is reduced from -0.019 to -0.017, and the t-test doesn't present significant differences between these two GDP growth rate indicators. Therefore, the regulation effects are not limited only to the coal industry, and the spillover effect originating from the relief of crowding-out effects may play a role here.

4.8 Conclusion

Although rich resource endowment is not necessarily negatively associated with regional economic growth, the resource curse is observed in many countries worldwide, and institutions are seen as a main factor determining the role of resource abundance. Using the ongoing regulation of the coal industry in China, empirical evidence indicates that after regulation, labor productivity in the coal industry improved and the crowding-out effect of coal mining on other economic activities was relieved. Both consequences positively improved regional economic performance. Our findings present evidence that the relationship between resource abundance and regional economic performance is subject to resource management and institutions. In addition, we find that coal mine regulation has gradually reduced frequent mine accidents and deaths in China. Finally, we admit a shortcoming of this study that both our theoretical model and empirical strategies neglect the cross-regional mobility of entrepreneur and capital, which is common in real world.

Appendix A

Appendix of Chapter 2

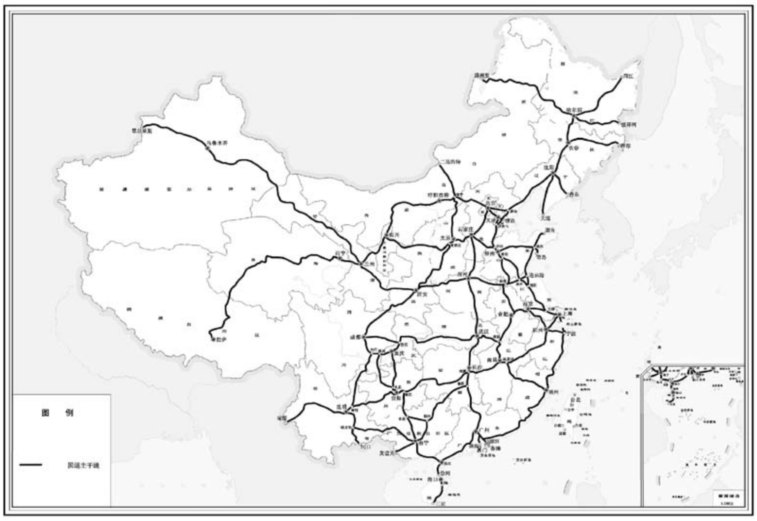


FIGURE A.1: National trunk highway system
Source: Ministry of Transport of P. R. China.

FIGURE A.2: National highway system planned 7918 highway network
Source: Xinhua Net.

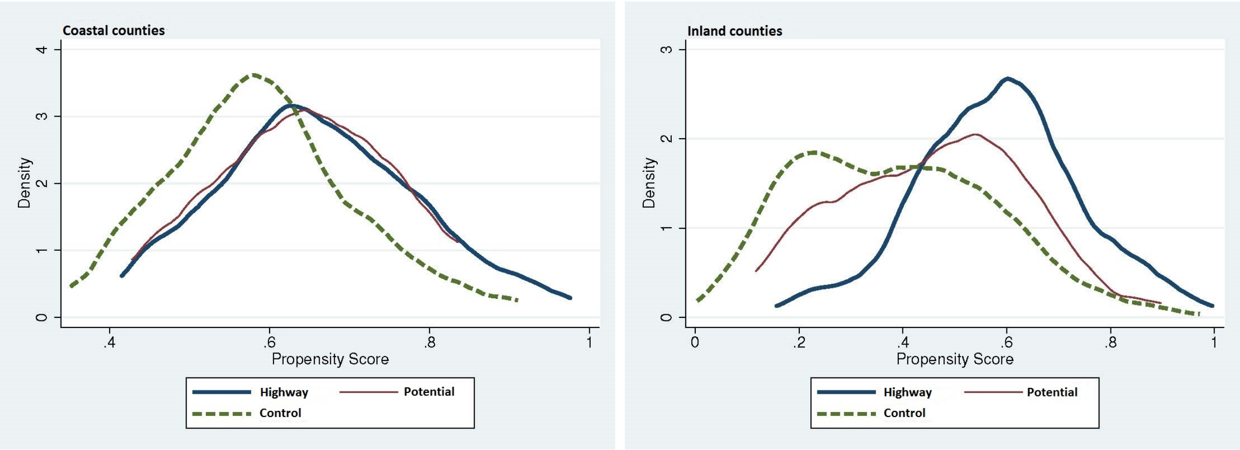


FIGURE A.3: Kernel density of propensity score (1)

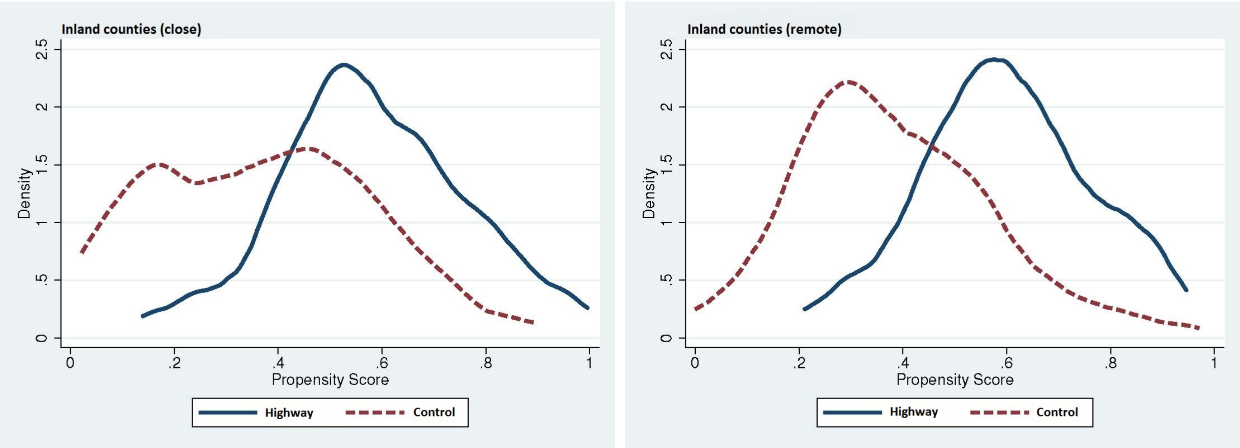


FIGURE A.4: Kernel density of propensity score (2)

Table A.1: Mean of covariates used in the matching procedure

Variate	(1)	(2)	(3)	(4)	(5)	(6)
area size	1766	1988	2393	2057	3286	2287
	(1022)	(1569)	(1076)	(1511)	(4898)	(1235)
minimum distance to	170	185	221	312	336	306
large cities	(71)	(102)	(68)	(154)	(164)	(111)
minimum distance to	271	278	390	646	813	713
large ports	(133)	(151)	(115)	(281)	(380)	(308)
employment share of	0.17	0.14	0.12	0.09	0.06	0.06
secondary sector	(0.13)	(0.10)	(0.05)	(0.07)	(0.05)	(0.05)
employment share of	0.16	0.17	0.16	0.13	0.12	0.11
tertiary sector	(0.06)	(0.07)	(0.05)	(0.07)	(0.06)	(0.05)
firm density	0.07	0.04	0.03	0.03	0.01	0.02
	(0.08)	(0.05)	(0.03)	(0.04)	(0.02)	(0.02)
population density	447	321	298	344	205	258
	(278)	(270)	(253)	(239)	(211)	(192)
population share	0.11	0.11	0.12	0.11	0.09	0.09
(education$\geqslant$high school)	(0.03)	(0.03)	(0.02)	(0.05)	(0.04)	(0.03)
initial per capita GDP	8.68	8.56	8.44	8.20	7.82	7.90
	(0.51)	(0.41)	(0.39)	(0.40)	(0.49)	(0.53)
Region	Coastal	Coastal	Coastal	Inland	Inland	Inland
Highway access	Yes	No	Potential	Yes	No	Potential
No. obs.	155	101	20	243	379	85

Note: Owing to data restrictions, population, employment, and education data come from the 5th national population census in 2000. In PSM, it is required that we measure the covariates before participation or fix them over time, and as such, we assume that these indicators are relatively fixed from 1998 to 2000. Owing to China's "Hukou" system, population mobility in peripheral regions is fairly limited. S.D. in parentheses. The description of the covariates is as follows:
area size: land area (km^2);
minimum distance to large cities: minimum spherical distance to three 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, and Xi'an) in 1998;
minimum distance to large ports: minimum spherical distance to Dalian, Tianjin, Shanghai, and Guangzhou, the largest port cities of the Northeast, North, East, and South regions of mainland China, respectively (i.e., of Dongbei, Huabei, Huadong, and Hua'nan, respectively);
employment share of secondary sector: employment share of the secondary sector in total employment;
employment share of tertiary sector: employment share of the tertiary sector in total employment;
firm density: number of industrial firms above the designated size per km^2;
population density: resident population per km^2;
population share (education$\geqslant$high school): share of population with a high school or higher education;
initial per capita GDP: log of per capita GDP (Yuan) in 1998.

Table A.2: Mean of outcome indicators

Variate	(1)	(2)	(3)	(4)	(5)	(6)
GDP per capita 1998	6666	5690	4963	3930	2811	3069
	(3463)	(2579)	(1689)	(1648)	(1480)	(1493)
GDP per capita 2007	16729	14445	11881	10280	7496	8161
	(9904)	(9663)	(3545)	(5785)	(5468)	(5750)
Firm density 1998	0.07	0.04	0.03	0.03	0.01	0.02
	(0.08)	(0.05)	(0.03)	(0.04)	(0.02)	(0.02)
Firm density 2007	0.21	0.11	0.08	0.05	0.02	0.02
	(0.35)	(0.27)	(0.07)	(0.07)	(0.03)	(0.03)
Investment 1998	9.33	6.50	6.96	5.09	2.66	4.29
	(9.50)	(10.83)	(5.79)	(6.17)	(7.44)	(10.37)
Investment 2007	32.32	17.40	14.97	15.22	6.66	12.22
	(39.97)	(24.56)	(12.28)	(17.92)	(10.55)	(27.11)
Output 1998	5.28	3.09	4.02	2.50	1.11	1.59
	(5.84)	(4.01)	(4.30)	(3.03)	(2.39)	(2.25)
Output 2007	35.91	19.20	19.53	15.37	6.32	8.62
	(36.80)	(29.89)	(17.17)	(16.41)	(9.48)	(10.73)
Region	Coastal	Coastal	Coastal	Inland	Inland	Inland
Highway access	Yes	No	Potential	Yes	No	Potential
No. obs.	155	101	20	243	379	85

Note: S.D. in parentheses. Description of the indicators is as follows:
GDP per capita: per capita GDP (Yuan) in 1998 & 2007;
Firm density: number of industrial firms above the designated size per km^2 in 1998 & 2007;
Investment: net value of fixed assets of industrial firms above the designated size (100 million Yuan) in 1998 & 2007;
Output: value-added of industrial firms above the designated size (100 million Yuan) in 1998 & 2007.

Table A.3: Post-matching differences in covariates (bias %)

Covariate	2.1-2.4	2.5-2.8	3.1-3.4	3.5-3.8	7.1-7.4	7.5-7.8, 8.1-8.6	9.1-9.2	9.3-9.4
area size	-12.9	5.9*	29.2	-4.5	4.0	2.8	-14.0	-2.9
min. distance to large cities	-7.5	4.0	50.3	-8.7	5.1	-19.5	-3.7	-19.5
min. distance to large ports	-11.5	9.6	74.7*	0.2	2.1	16.6	-9.7	16.6
empl. share of secondary sector	9.7	5.5	-27.9	-5.2	-5.1	2.0	5.4	2.2
empl. share of tertiary sector	1.2	5.0	10.6	-11.0	5.0	-2.4	-2.8	-2.5
firm density	11.3	15.2	-47.0	12.7	10.8	-9.1	8.2	-9.1
pop. density	7.9	0.9	-32.1	7.2	-8.8	4.0	11.2	4.0
pop. share (educ.$\geqslant$high school)	-6.0	-1.5	42.9	-10.9	0.7	-16.2	-11.6	-16.3
initial p.c. GDP	10.4	8.8	-38.0	-5.2	8.2	2.9	8.7	2.9

Note: The numbers in the first line imply the related tables and columns in Tables 2.2–2.3, 2.7–2.9. For example, 3.1–3.4 means data related to columns 1–4 of Table 2.3.
* Significant at the 10% level.

Appendix B

Appendix of Chapter 3

TABLE B.1: Descriptive summary of Central Asia economy

Indicator	Country	1996	1998	2000	2001	2003	2005
Population	Kazakhstan	15.6	15.1	14.9	14.9	14.9	15.1
(in million)	Tajikistan	5.9	6.0	6.2	6.3	6.5	6.8
	Uzbekistan	23.2	24.1	24.7	25.0	25.6	26.2
	Kyrgyzstan	4.6	4.8	4.9	4.9	5.0	5.2
	Turkmenistan	4.3	4.4	4.5	4.6	4.6	4.7
GDP per capita	Kazakhstan	1990	2052	2344	2664	3187	3771
(constant 2005 US$)	Tajikistan	206	215	234	254	301	340
	Uzbekistan	399	422	446	459	486	547
	Kyrgyzstan	360	392	417	435	457	476
	Turkmenistan	1270	1170	1404	1449	1469	1707
Exports	Kazakhstan	17.4	15.5	22.1	21.7	27.2	30.6
(constant 2005 US$, in billion)	Tajikistan	0.4	0.4	0.4	0.3	0.5	0.6
	Uzbekistan	3.9	3.4	3.4	3.2	3.8	5.4
	Kyrgyzstan	0.8	0.9	0.9	0.8	0.9	1.0
	Turkmenistan	1.8	0.9	2.8	2.9	3.7	5.3
Imports	Kazakhstan	20.2	20.1	21.1	20.7	19.8	25.5
(constant 2005 US$, in billion)	Tajikistan	0.6	0.6	0.7	0.6	0.8	1.2
	Uzbekistan	4.8	3.4	3.0	3.2	3.1	4.1
	Kyrgyzstan	1.3	1.1	1.0	0.9	1.1	1.4
	Turkmenistan	1.8	1.8	2.4	2.7	3.4	3.9
Manufactures imports	Kazakhstan	64	71	74	75	78	78
(% of merchandise imports)	Tajikistan	-	-	-	-	-	-
	Uzbekistan	-	-	-	-	-	-
	Kyrgyzstan	48	57	59	56	57	52
	Turkmenistan	-	79	80	-	-	-

Note: Data source: World Development Indicators (WDI), World Bank.
-: related data is not available.

Table B.2: Rail transport infrastructure improvements in Central Asia

Country	Indicator	1996	1998	2000	2001	2003	2005
Kazakhstan	Railways, goods transported (billion ton-km)	-	-	125.0	-	147.7	171.9
	Rail lines (total route-thousand km)	-	-	13.5	-	13.8	14.2
Tajikistan	Railways, goods transported (billion ton-km)	1.7	1.5	1.3	1.3	1.1	-
	Rail lines (total route-thousand km)	-	-	-	-	-	0.6
Uzbekistan	Railways, goods transported (billion ton-km)	19.7	15.7	15.4	-	-	18.0
	Rail lines (total route-thousand km)	-	3.6	3.6	-	-	4.0

Note: Data source: World Development Indicators (WDI), World Bank.
Data on roads is not presented because most of them is not available for the related years in WDI.
Data on Kyrgrzstan and Turkmenistan is not available.
-: related data is not available.

Table B.3: Descriptive summary of dataset

Panel A: Summary of variables (2000-2001)						
	Variable	Obs.	Mean	Std. Dev.	Min	Max
Rail	Export value (USD)	2851	169021	810416	152	2.10E+07
	Number of shipments	2851	1.80	2.00	1	23
	Number of exporters	2060	1.38	1.06	1	17
	Unit price (USD)	6879	16341	258353	0.002	1.07E+07
Non-rail	Export value (USD)	4421	181593	1204678	220	4.00E+07
	Number of shipments	4421	2.37	2.90	1	31
	Number of exporters	2253	1.96	2.19	1	25
	Unit price (USD)	12798	1471	32173	0.003	2.90E+06

Panel B: Statistical summary of main indicators				
	2000	2001	2000	2001
	Rail		Non-rail	
Export value (total, USD 1,000,000)	216.4	266.14	553.62	251.9
Number of shipments (total)	2849	2293	6103	4387
Number of export firms (total)	248	331	241	319
Value per shipment (medium, USD 1,000)	17.82	23.33	14.4	15.84
Value per shipment (mean, USD 1,000)	75.96	116.07	90.22	57.48
Annual value per firm-product-country (medium, USD 1,000)	17.5	19.8	10.7	12
Annual value per firm-product-country (mean, USD 1,000)	137.8	207.2	224.1	128.4
Unit price (medium, USD)	1.3	1.6	1.2	1.1
Unit price (mean, USD)	62.2	71.6	16.9	29.5

Note: Panel A: Export value and number of shipments are based on aggregated firm-product (HS-8)-country-freight mode-year level data, number of exporters is based on aggregated product (HS-8)-country-freight mode-year level data, and the unit price is based on transaction level data.
Panel B: Unit price (mean, USD) is obtained after winsorizing 95% (one way cut-off) of the total observations to avoid disturbance from outliers (extremely large unit prices).

Table B.4: Sped-up mileage to Urumqi (Xinjiang)

Starting point (Province)	Speed-up point (Station)	Sped-up mileage (km)	Total mileage (km)
Xinjiang	Urumqi	0	0
Gansu	Lanzhou	1892	1892
Ningxia	Lanzhou	1892	2008
Qinghai	Lanzhou	1892	2120
Beijing	Lanzhou	1892	3216
Tianjin	Lanzhou	1892	3240
Inner Mongolia	Lanzhou	1892	3000
Liaoning	Lanzhou	1892	4007
Tibet	Lanzhou	1892	4080
Hebei	Lanzhou	1892	2822
Chongqing	Baoji	2395	3345
Guizhou	Baoji	2395	3840
Sichuan	Baoji	2395	3026
Hainan	Baoji	2395	5214
Yunnan	Baoji	2395	5564
Shaanxi	Xi'an	2568	2568
Shanxi	Xi'an	2568	2597
Henan	Zhengzhou	3079	3079
Shandong	Zhengzhou	3079	3134
Fujian	Zhengzhou	3079	4555
Guangdong	Zhengzhou	3079	4684
Anhui	Zhengzhou	3079	3723
Guangxi	Zhengzhou	3079	5493
Hunan	Zhengzhou	3079	3977
Hubei	Zhengzhou	3079	4342
Shanghai	Xuzhou	3428	4077
Jiangsu	Xuzhou	3428	3776
Zhejiang	Xuzhou	3428	4168
Heilongjiang	Xuzhou	3428	4490
Jilin	Xuzhou	3428	4244
Jiangxi	Nanchang	4129	4129

Table B.5: Effect of unit price changes

	ΔUnitprice	
	(1)	(2)
Xinjiang	0.0195	0.0954
	(0.0552)	(0.0767)
Samples	rail	rail
Sample type	f-p-c	p-c
Observations	257	148
R-squared	0.0005	0.0105

Note: $\Delta Unitprice = Unitprice_{2001}/Unitprice_{2000}$.
Regression equation is: $\Delta Unitprice_{f,p,c} = \alpha_0 + \alpha_1 Xinjiang_f + \epsilon_{f,p,c}$. $Xinjiang$ is a dummy variable (=1 for exports originating from Xinjiang, and, =0 for others).
I compare the unit price change between rail freight products originating from Xinjiang and other regions. If there is a decrease in transportation costs, there should be a decrease in the F.O.B. unit price, since it contains domestic transportation costs.
I drop the samples if the unit price difference is abnormally large. I keep only the samples where $\Delta Unitprice$ is larger than 1/3, and less than 3. I use the mean unit price in related sample types for a specific year. For column 1, I keep only paired data at the f-p-c level. For column 2, I identify two regions (Xinjiang and other regions) at the p-c level. I keep only paired data at the specified p-c level, if both Xinjiang and the other region have paired data at this p-c level (four records (twoyears*tworegions) should be matched at each p-c.).

Appendix C

Appendix of Chapter 4

Table C.1: Placebo tests

	(1) GDP	(2) Wage		(3) Tertiary
Small-size × post1990	-2.006*	-0.653	CoalCity × post1990	-0.020
	(1.094)	(1.117)		(0.032)
post1990	3.985***	3.967***	post1990	-0.032
	(1.248)	(1.676)		(0.034)
CoalDependence	-0.132	-7.306		
	(2.649)	(4.986)		
Constant	110.2***	86.54*	Constant	2.207***
	(24.01)	(46.11)		(0.018)
Year FE	yes	yes	Year FE	yes
Province FE	yes	yes	City FE	yes
N	198	144	N	258
R^2	0.546	0.748	R^2	0.236

Note: Heteroskedasticity-robust standard errors are reported in parentheses, and standard errors are clustered at the province level. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

C.1 Data source

Data used in this chapter were obtained from many sources.

GDP: deflated annual GDP per capita index, from the "China Statistical Yearbook."

Wage: annual real wage growth index deflated by consumer price index, from the "China Labor Statistical Yearbook" and the "China Statistical Yearbook."

CoalDependence: logarithm of (coal output/total population; ton/10,000 people), from the "China Coal Industry Yearbook."

West: dummy for Shaanxi, Gansu, Ningxia, Xinjiang, Sichuan, Chongqing, Yunnan, Guizhou, Guangxi, and Inner Mongolia.

Northeast: dummy for Heilongjiang, Jilin, and Liaoning.

Central: dummy for Henan, Hunan, Anhui, and Shanxi.

Tertiary: prefecture-level value-added growth of tertiary industries, from the "China City Statistical Yearbook."

$\Delta \ln(\text{GDP})$: Real GDP per capita annual growth rate is calculated as:

$$\Delta \ln(\text{GDP})=\frac{\ln \text{GDPp.c.}_{i+3}-\ln \text{GDPp.c.}_{i}}{3}$$

i refers to year. GDPp.c.: real GDP per capita, from the "China Statistical Yearbook."

$\Delta \ln(\text{Wage})$: Real wage annual growth rate in Section 6 is calculated as

$$\Delta \ln(\text{Wage})=\frac{\ln \text{RealWage}_{i+3}-\ln \text{RealWage}_{i}}{3}$$

RealWage: real wage (deflated by the consumer price index), from the "China Labor Statistical Yearbook" and the "China Statistical Yearbook."

Mortality: logarithm of mortality rate in the coal industry (person/100 million tons), from the "China Coal Industry Yearbook" and the "Statistical Communiqué of the National Economic and Social Development" at the provincial level.

InitialGDP: logarithm of the initial real GDP per capita, from the "China Statistical Yearbook."

Invest: the total social investment in fixed assets/GDP, from the "China Statistical Yearbook."

Educ: share of workers with an education level higher than senior high school of total employment (%), from the "China Labor Statistical Yearbook."

FDI: logarithm of (FDI/total population; US$/person), from the "China Statistical Yearbook."

MortalityTotal: logarithm of the mortality rate in non-agricultural industries (persons/million workers), from the "China's Work Safety Yearbook" and the "Statistical Communiqué of the National Economic and Social Development" at the provincial level.

MortalityTotalExcoal: logarithm of the mortality rate in non-agricultural industries after excluding the coal industry (persons/million workers), from the "China's Work Safety Yearbook" and the "Statistical Communiqué of the National Economic and Social Development" at the provincial level.

SmallSizeMines (less than 30,000 tons): product share of small-size mines (i.e., annual output less than 30,000 tons) to total coal output, from the "China Coal Industry Yearbook."

MiddleSizeMines (30,000–300,000 tons): product share of mines (annual output less than 300,000 tons and larger than 30,000 tons) to total coal output, from the "China Coal Industry Yearbook."

LabProd: logarithm of (coal output/coal industry employment; ton/person), from the "China Coal Industry Yearbook" and the "China Labor Statistical Yearbook."

InvestCoal: logarithm of [investment of coal industry in fixed assets (including only SOE mines)/employment in the coal industry; 10,000 yuan/person], from the "China Energy Statistical Yearbook."

Skill: average wage of coal industry/average wage of all industries, from the "China Labor Statistical Yearbook".

$\Delta\ln(\mathrm{GDP})$Excoal: modified $\Delta\ln(\mathrm{GDP})$ after excluding the value-added of the coal industry to total GDP, from the "China Statistical Yearbook," the "China Industry Economy Statistical Yearbook," and the "Statistical Yearbook (Provincial-level)."

Bibliography

[1] Acemoglu, D., S. Johnson, J.A. Robinson (2001) The Colonial Origins of Comparative Development: An Empirical Investigation, *American Economic Review* 91(5), 1369–1401.

[2] Ago, T., I. Isono, T. Tabuchi (2006) Locational Disadvantage of the Hub, *Annals of Regional Science* 40, 819–848.

[3] Albarrán, P., R. Carrasco, A. Holl (2013) Domestic Transport Infrastructure and Firms' Export Market Participation, *Small Business Economics* 40(4), 879–898.

[4] Andersen, J.J., S. Aslaksen (2007) Constitutions and the Resource Curse, *Journal of Development Economics* 87, 227–246.

[5] Angrist, J.D., A.D. Kugler (2008) Rural Windfall or a New Resource Curse? Coca, Income, and Civil Conflict in Columbia, *Review of Economics and Statistics* 90(2), 191–215.

[6] Baland, J-M., P. Francois (2000) Rent-Seeking and Resource Booms,*Journal of Development Economics* 61(2), 527–542.

[7] Banerjee, A., E. Duflo, N. Qian (2012) On the Road: Transportation Infrastructure and Economic Growth in China, NBER Working Paper 17897.

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

[9] Basu, A., D. Polsky, W. Manning (2008) Use of Propensity Scores in Non-Linear Response Models: The Case for Health Care Expenditures, NBER Working Paper 14086.

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

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

[12] Baum-Snow, N., M. Turner (2012) Roads, Transportation and the Decentralization of Chinese counties, *Mimeo*, Brown University.

[13] Bentley, J. (1993) *Old World Encounters: Cross-Cultural Contacts and Exchanges in Pre-Modern Times*. Oxford University Press, New York.

[14] Bhattacharyya, S., R. Hodler (2010) Natural Resources, Democracy, and Corruption, *European Economic Review* 54, 608–621.

[15] Boarnet, M. (1998) Spillovers and the Locational Effects of Public Infrastructure, *Journal of Regional Science* 38, 381-400.

[16] Caselli, F., G. Michaels (2013) Do Oil Windfalls Improve Living Standards? Evidence from Brazil, *American Economic Journal: Applied Economics* 5(1), 208–238.

[17] Chandra, A., E. Thompson (2000) Does Public Infrastructure affect Economic Activity? Evidence from the Rural Interstate Highway System, *Regional Science and Urban Economics* 30, 457-490.

[18] Costa-Campi, M.T., E. Viladecans-Marsal (1999) The District Effect and the Competitiveness of Manufacturing Companies in Local Productive Systems, *Urban Studies* 36(12), 2085–2098.

[19] Davis, G.A. (1995) Learning to Love the Dutch Disease: Evidence from the Mineral Economies, *World Development* 23(10), 1765–1779.

[20] Deaton, B.J., E. Niman (2012) An Empirical Examination of the Relationship between Mining Employment and Poverty in the Appalachian Region, *Applied Economics* 44, 303–312.

[21] Ding, J., Y. Wang, K. Deng (2007) Is There a “Resource Curse” in China’s Economic Development? (in Chinese), *The Journal of World Economy* 30(9), 38–46.

[22] Donaldson, D. (2014) Railroads of the Raj: Estimation the Impact of Transportation Infrastructure, *American Economic Review*, forthcoming.

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

[24] Emran, M.S., Z. Hou (2013) Access to Markets and Rural Poverty: Evidence from Household Consumption in China, *Review of Economics and Statistics* 95(2), 682–697.

[25] Faber, B. (2014) Trade Integration, Market Size, and Industrialization: Evidence from China’s National Trunk Highway System, *Review of Economic Studies*, forthcoming.

[26] Fan, R., Y. Fang, S.Y. Park (2012) Resource Abundance and Economic Growth in China, *China Economic Review* 23(3), 704–719.

[27] Fang, Y., K. Ji, Y. Zhao (2008) Curse or Blessing? A Revisitation of the “Resource Curse” in China (in Chinese), *Mimeo*, Xiamen University.

[28] Frankel, J.A. (2010) The Natural Resource Curse: A Survey, NBER Working Paper No. 15836.

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

[30] Gylfason, T. (2001) Natural Resources, Education, and Economic Development, *European Economic Review* 45, 847–859.

[31] Hu, Y., D. Xiao (2007) The Threshold of Economics Growth and the Natural Resource Curse (in Chinese), *Management World* 4, 15–23.

[32] Heckman, J., H. Ichimura, J. Smith, P. Todd (1998) Characterizing Selection Bias using Experimental Data, *Econometrica* 66, 1017-1098.

[33] Helpman, E., P. Krugman (1985) *Market Structure and International Trade*. Cambridge, MA.

[34] Hodler, R. (2006) The Curse of Natural Resources in Fractionalized Countries, *European Economic Review* 50(6), 1367–1386.

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

[36] —— (2004b) Transport Infrastructure, Agglomeration Economies, and Firm Birth: Empirical Evidence from Portugal, *Journal of Regional Science* 44, 693-712.

[37] Indurkhya, A., N. Mitra, D. Schrag (2006) Using Propensity Score to Estimate The Cost-Effectiveness of Medical Therapies, *Statistics in Medicine* 25, 1561-1576.

[38] James, A., D. Aadland (2011) The Curse of Natural Resources: An Empirical Investigation of U.S. Counties, *Resource and Energy Economics* 33, 440–453.

[39] Ji, K., J.R. Magnus, W. Wang (2012) Natural Resources, Institutional Quality, and Economic Growth in China, *Mimeo*, Tilburg University.

[40] Khandker, S.R., G.B. Koolwal, H.A. Samad (2010) *Handbook on Impact Evaluation: Quantitative Methods and Practices*. World Bank, Washington, DC.

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

[42] Larsen, E.R. (2006) Escaping the Resource Curse and the Dutch Disease? When and Why Norway Caught Up and Forged Ahead of Its Neighbours, *American Journal of Economics and Sociology* 65(3), 605–640.

[43] Li, Z., Y. Chen (2013) Estimating the Social Return to Transport Infrastructure: A Price-Difference Approach Applied to a Quasi-Experiment, *Journal of Comparative Economics* 41(3), 669–683.

[44] Lu, Y., Z. Tao, Y. Zhang (2013) How Do Exporters Respond to Antidumping Investigations?, *Journal of International Economics* 91, 290–300.

[45] Martincus, C.V., J. Blyde (2013) Shaky Roads and Trembling Exports: Assessing the Trade Effects of Domestic Infrastructure Using a Natural Experiment,*Journal of International Economics* 90, 148–161.

[46] Mehlum, H., K. Moene, R. Torvik (2006) Institutions and the Resource Curse, *The Economic Journal* 116, 1–20.

[47] Michaels, G. (2011) The Long-term Consequences of Resource-Based Specialisation, *The Economic Journal* 121, 31–57.

[48] Mori, T. (2012) Increasing Returns in Transportation and the Formation of Hubs, *Journal of Economic Geography* 12, 877–897.

[49] Nicolini, R. (2003) On the Determinants of Regional Trade Flows, *International Regional Science Review* 26(4), 447–465.

[50] Nie, H., M. Jiang (2011) Coal Mine Accidents and Collusion between Local Governments and Firms: Evidence from Provincial-Level Panel Data in China (in Chinese), *Economic Research Journal* 6, 146–156.

[51] Pan, K., J. Pu, T. Xiang (2002) Comparative Research of the Coal Market Concentration Degree between China and America (in Chinese), *Management World* 12, 77–88.

[52] Papyrakis, E., R. Gerlagh (2007) Resource Abundance and Economic Growth in the United States, *European Economic Review* 51, 1011–1039.

[53] People's Daily (2000) Interpretation of Train Diagram 2000 (in Chinese), October 23, p12.

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

[55] Qin, Yu (2014) No County Left Behind? The Distributional Impact of High-Speed Rail Upgrade in China, *Mimeo*, Cornell University.

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

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

[58] Roberts, M., U. Deichmann, B. Fingleton, Tuo Shi (2012) Evaluating China's Road to Prosperity: A New Economic Geography Approach, *Regional Science and Urban Economics* 42, 580-594.

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

[60] Rousslang, D., T. To (1993) Domestic Trade and Transport Costs as Barriers to International Trade, *Canadian Journal of Economics* 26, 208–221.

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

[62] Sachs, J.D., A.M. Warner (1995) Natural Resource Abundance and Economic Growth, *Mimeo*, Harvard University.

[63] ——, —— (2001) The Curse of Natural Resources,*European Economic Review* 45, 827–838.

[64] Sala-i-Martin, X., A. Subramanian (2003) Addressing the Natural Resource Curse: An Illustration from Nigeria, NBER Working Paper No. 9804.

[65] Sanchis-Guarner, R., T. Lyytikäinen (2011) Driving Up Wages: The Effects of Road Improvements in Great Britain, *Mimeo*, London School of Economics.

[66] Sarraf, M., M. Jiwanji (2001) Beating the Resource Curse: The Case of Botswana, Environment Department. World Bank, Washington DC.

[67] Shao, S., Z. Qi (2008) Energy Development and Economic Growth in Western China: An Empirical Analysis Based on the Resource Curse Hypothesis (in Chinese), *Economic Research Journal* 4, 147–160.

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

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

[70] Torvik, R. (2002) Natural Resource, Rent Seeking, and Welfare, *Journal of Development Economics* 67, 455–470.

[71] van der Ploeg, F. (2011) Natural Resources: Curse or Blessing? *Journal of Economic Literature* 49(2), 366–420.

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

[73] Wang, F. (2006) Development Problems and Policy Analysis of Small Coal Mines in China (in Chinese), *Journal of China University of Geosciences (Social Science Edition)* 6(6), 61–67.

[74] Wang, Q. (2003) Safety of Coal Mines: International Experience and Suggestions (in Chinese), *China Coal* 29(4), 44–53.

[75] Wang, S. (2006) Regulating Death at Coalmines: Changing Mode of Governance in China, *Journal of Contemporary China* 15(46), 1–30.

[76] World Bank (2006) Governance, Investment Climate, and Harmonious Society: Competitiveness Enhancements for 120 Cities in China, Report No. 37759-CN. (http://siteresources.worldbank.org/INTCHINA/Resources/318862-1121421293578/120cities_en.pdf)

[77] —— (2008) *World Development Report 2009—Reshaping Economic Geography.* World Bank, Washington, DC.

[78] Wu, H. (2011) The Trade Relationship between China and Central Asia (in Chinese), *Russian, Central Asian & East European Market.* (http://euroasia.cass.cn/news/494692.htm)

[79] Xu, H., H. Zhou (2014) Expressways and Manufacturing Employment in China's Counties, *Mimeo*, Tohoku University.

[80] Xu, K., J. Han (2005) The "Resource Curse" Effect on Regional Economy in China: Another Explanation for Regional Discrepancies (in Chinese), *The Economist* 6, 96–102.

[81] 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.

[82] 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.