Transport Capacity, Price Gaps, and the Economic Return:

A Unique Natural Experiment in China [*]

Zhigang Li

4 November 2003

Abstract

The goal of this study is to improve upon the existing literature in quantifying the economic return of transportation infrastructure. Specifically, I propose incorporating the changes of between-city price gaps to approximate economic benefits that have been omitted by the current literature. Identifying the causal impact of new infrastructure on price gaps is complicated. To circumvent this problem, I propose to employ empirical settings in which natural experiments can be constructed to eliminate the effects of confounding factors. In particular, I consider two cities in northwestern China; they are connected by a railroad, the capacity of which was doubled in my sample period (1993-1996). I find strong evidences suggesting that increasing the railroad capacity significantly decreased price gaps. The change in price gaps implies a real economic return of between 12% and 24% per year.

---

[*] I thank Roger Gordon for invaluable guidance. The draft is preliminary and all errors are mine. Contact: Department of Economics 0508,University of California, San Diego, 9500 Gilman Drive,La Jolla, CA 92093-0508,U.S.A., zli@econ.ucsd.edu.

1 Introduction

The economic return of transportation infrastructure is a factor of essential importance in public finance decision-making. Over the past five decades serious researches have been carried out to *quantify* this factor, and solid groundwork has been laid. The initial wave of attempts was made by a group of American economic historians — e.g. Albert Fishlow, Robert W. Fogel, and Harvey H. Segal — around the end of 1950's. This literature concerned the economic significance of the "new" transportation technologies in the nineteenth-century — canals and railroads in particular. An important concept of the literature is "social saving", typically measured by the difference between the *actual* shipping cost under an interested infrastructure and the *hypothetical* shipping cost with the best alternative infrastructure. See O'Brien [11] for a survey of this literature and see McClelland [10] for an informative criticism. This shipping-cost-based approach has also been adopted by practitioners to analyze the benefit of *future* infrastructure investment. In this type of analysis a difference is estimated typically between the total shipping cost under the *current* infrastructure and the *predicted* cost if a specific infrastructure investment is made. See, for example, Gramlich [9] for relevant discussions. Another main strain of the literature is based upon the "production-function" approach initiated by Aschauer [8]. This strain attempts to answer the question with macro-econometric estimations. As discussed in a Congressional Budget Office (CBO) report [6], this approach "generally estimates an aggregate 'production function,' ... (which) can (then) be used to compare the effects on output of added public and private investment."

Despite these research attempts, effective tools and reliable empirical evidences are still scarce, as acknowledged by the 1991 CBO report [6]. Specifically, the macro-approach has been heavily criticized for its intrinsic econometric problems and non-robust findings; moreover, since this approach measures a *reduced-form* relation between output and public capital, it has a weak predicting power for the economic effects of *future* infrastructure. In contrast, the shipping-cost approach is more practicable in terms of predictability, but it is also much more expensive due to its taller data requirement. As a result, "few cost-benefit analyses have been conducted for different infrastructure investment strategies. (CBO [6])" More seriously, even if the data problem can be conquered, the shipping-cost approach may still provide a biased estimate of the economic benefit. This bias results from the failure of the approach to consider *capacity* changes due to new investment, and the bias could be very large.

In this study I propose to correct for the bias by incorporating the between-economy price differentials of tradable goods. In particular, the key information I need is *how much* a specific infrastructure affects the price gaps between two economies; this effect on price gaps can then be translated into social welfare values with a simple formula. I show that this price-gap approach improves upon the existing techniques and may even eliminate the bias.

The difficulty with the price-gap approach, like that with the macro-econometric approach, lies on identifying the impact of the infrastructure on price gaps. The identification requires having good controls of confounding factors, and this imposes high (sometimes unattainable) data requirement on empirical studies. Failure to control the confounding factors may generate simultaneity bias and provide spurious relations; this is particularly relevant to the infrastructure literature since infrastructure investment is obviously endogenously determined. Added to this difficulty is the coexistence of more than one transportation facilities. Due to their substitutability, it is necessary to consider not only the infrastructure in interest, but also all other competing infrastructures. If the information on competing infrastructure is omitted, the estimate of economic return could be seriously misleading. To summarize, probably due to the data requirement of the price-gap approach, the shipping-cost approach has been dominant and the growth of the literature has been limited.

To circumvent this data problem, in this paper I propose to choose empirical settings in which "natural experiment" can be constructed to eliminate the confounding factors and to identify interested causal effects. Specifically, I consider a railroad in China that has the following features. First, the railroad has *asymmetric* flows — shipping is constrained by the rail capacity in one direction but not in the other. Second, the railroad underwent a sudden and pre-determined change in my sample period. Third, natural environment determines that there is little competition from other transportation facilities. I show that the identification is extraordinarily simplified in this setting. Strong evidences are provided showing that the between-city price gaps shrank about 40 percent by an capacity-doubling investment in the railroad.

The organization of this paper is as follows. In section one I discuss the price-gap approach and compare it with the shipping-cost approach that has been widely used in the literature. In section two I describe in details the micro-setting under which I can apply the price-gap approach conveniently. In section three I introduce my data. In section four I turn to econometrically estimating the effect of railroad capacity on price gaps. Specifically, I shall first analyze the price gap patterns of all 35 commodities; then I shall apply my difference-in-difference estimator to the data. Finally, I conclude the exercise and discuss various important extensions.

2 What is the Literature Missing?

Consider a world in which a good is produced *only* in economy O (the outside) and is consumed in both economies O and H (home). The good is produced at constant marginal cost c. The two economies are connected by a transportation infrastructure with two attributes: marginal operating cost r (assumed to be constant) and capacity K.

Figure 1: Illustration of the Price-gap Approach

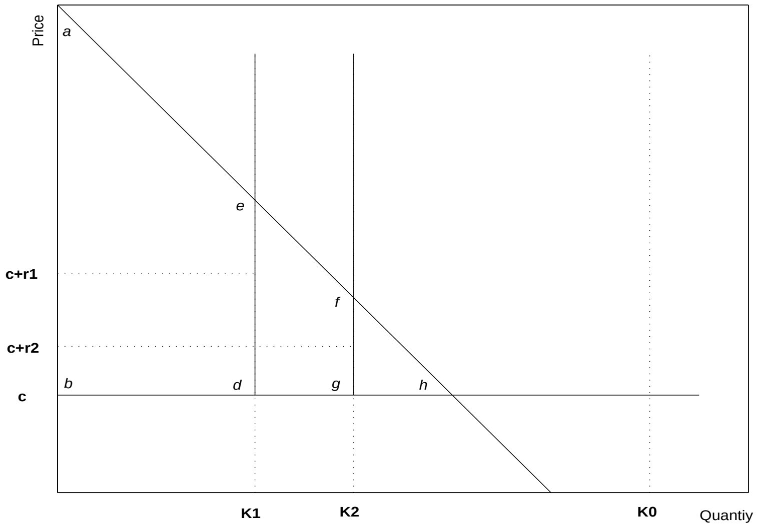

Figure 1 represents the market of economy H. If $r = 0$ and $K = K_0$, then the maximal social surplus of the world is represented by the triangle abh. If $r = 0$ and $K = K_1$, then the supply curve for economy H becomes perfectly inelastic at K_1. (For simplicity, I have assumed that there is no congestion before K_1 is reached.) In this case, the social surplus of the world shrinks to $abde$. Furthermore, if $r = r_1 > 0$, the social surplus is given by $abde$ minus the rectangle with area $r_1 \times K_1$. Finally, suppose that the infrastructure is upgraded such that $r = r_2$ and $K = K_2$, then the incremental benefit of this investment can be measured by $defg$ plus the rectangle with area $r_1 \times K_1$ minus the rectangle with area $r_2 \times K_2$. This suggests the following formula under linear-demand-curve and competitive-pricing assumptions:

$$Incremental\ Benefit = [(P_1^h - c) + (P_2^h - c)](K_2 - K_1)/2 + K_1 r_1 - K_2 r_2 \qquad (1)$$

Here P_1^h is the home price when $K = K_1$ and P_2^h is the home price when $K = K_2$. In practice, c is rarely observed. To deal with this problem, I approximate c with P^o, the price in economy O. This gives the following formula, which I shall call a "price gap approach" since $P_1^h - P^o$ and $P_2^h - P^o$ are the between-economy price gaps.[1]

$$Incremental\ Benefit = [(P_1^h - P^o) + (P_2^h - P^o))](K_2 - K_1)/2 + K_1 r_1 - K_2 r_2 \qquad (2)$$

Despite its simplicity, this price-gap approach has not been seen implemented in practice. Instead, the literature is dominated by the shipping-cost approach, which considers a formula of the following type:

$$Incremental\ Benefit = K'(r - r') \qquad (3)$$

---

[1] Here I have assumed that only one transportation facility is available between the economies. If there are more than one facilities, formula (1) provides a biased estimate due to a substitution effect. To illustrate, suppose that the demand curve for the good in economy H is $Q = a - bP$; suppose that there are two transportation facilities between economies H and O. One facility has zero marginal cost and a limited capacity K; the other facility has a marginal cost of r and an unlimited capacity. Suppose that K is so small that some goods are shipped through the expensive transportation facility. In this case the equilibrium price gap can be shown to be r. As K increases, some shippers switch from the expensive facility to the cheaper one, but the equilibrium price gap does not change until all shippers have switched. Without accounting for this substitution, formula (1) is biased. In this example, supposed that the new capacity if K', the bias can be shown to be $max[0, \frac{1}{2b}[a - b(c + r) - K][K' - a + b(c + r)]]$.

The two approaches provide the same estimate when $K = K'$, i.e. when transport capacity is not affected by the upgrading. However, if $K \neq K'$, then the shipping-cost approach provides biased estimates.[2]

Why the literature has overseen the price information for the past 50 years? One key reason may be the identification difficulties. As shown above, formula (2) requires information on changes in price gaps *caused* by changes in infrastructure. This task is difficult in practice since price gaps may change as a result of many other factors, e.g. inflation or competeting infrastructures. To address these problems, detailed information on confounding factors is necessary. This greatly raises the data requirement of the price-gap approach and may have made it too expensive to implement.

Nevertheless, the price-gap approach is not hopeless. Although we can not do controlled experiments, natural experiments could well be available given the large number of transportation facilities and varieties of geographical conditions. If a setting is available where the confounding factors either are absent or can be easily purged away, the identification will be much cheaper and the estimation results will be cleaner. The empirical study below makes exactly such a case.

3 The Empirical Setting

The empirical setting of this study involves three cities in northwestern China, two railroads connecting them, and the foreign countries with which the cities may trade.

3.1 Three Cities

Wulumuqi (abbreviated as WU below) is the economic center of Xinjiang province, the largest province in China and located in its northwest corner. Twenty-two percent of the province is covered by deserts. The deserts, together with mountains, seal the residential areas of the

---

[2] In this discussion I am focusing only on the short-term *direct* benefit — lower product prices — of an infrastructure investment. I have no intent to study the indirect benefit or long-term benefits and the externality of a new infrastructure. Furthermore, I am going to focus only on the benefit related to freight-shipping.

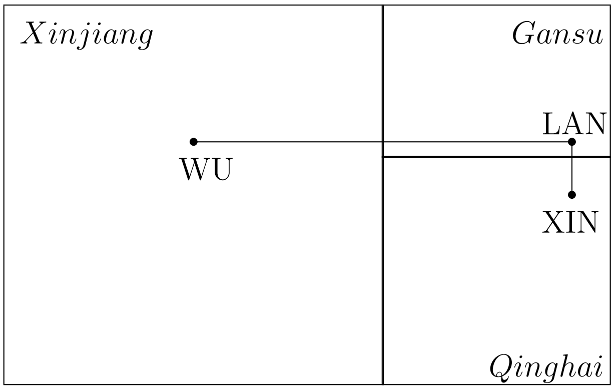

Figure 2: The Geography of the Setting

province off from those of the other provinces. Due to these natural impediments and the low quality of inter-province highways, rail transportation dominates the trades between Xinjiang and the rest of China. According to *Chinese Transportation Yearbook* [4] (pp. 522), around 95 percent of the trade volume (in terms of weight) between Xinjiang and the rest of China was carried out by the railroad.

Lanzhou (abbreviated as LAN below) is the economic center of Gansu province. Gansu is one of the three neighbor provinces of Xinjiang and is the only one of them that is connected to Xinjiang by railroad.[3] LAN is to the southeast of WU and the distance between them is around *1,200 miles.*

Xining (abbreviated as XIN below) is the economic center of Qinghai, a province neighbored to Xinjiang and Gansu. XIN is about 120 miles apart to the west of LAN. LAN and XIN are connected by a railroad and a highway, which have not had significant changes during our sample period (1992 - 1997).[4]

In this study I focus on the impact of LAN-WU railroad capacity on the LAN-WU price gaps. The LAN-XIN price gaps will be used to construct one of the reference groups for the LAN-WU price gaps.

---

[3] Qinghai and Tibet, the other two neighbor provinces to Xinjiang, are each connected to Xinjiang by a low quality road.

[4] There is no significant upgrading of the LAN-XIN rail and highway since 1980 according to the *Chinese Transportation Yearbook* [4].

3.2 The LAN-WU Railroad

The LAN-WU rail was built in the 1960s'. According to the *Map of Chinese Rail Transportation* [7] and *Chinese Transportation Yearbook* [4], there has been no major change to the railroad until the mid-1990's. A "capacity-doubling" project, which started on September 16, 1992 and ended in October 21, 1994, doubled the rails for about 80 percent of the railroad [5] and expanded its theoretical capacity from 12 to 25 million tons per year. The actual capacity increased significantly for 1995, and it took about five more years for the theoretical capacity to be fully utilized.

It is important to note that the "capacity-doubling" project is listed in the eighth national five-year-plan (1990 - 1995) of China. This will be useful later in my argument for the causal effect of railroad capacity on price gaps.

3.3 Neighbor Countries and Their Trade with Xinjiang

Xinjiang borders on eight countries — Mongolia, Russia, Kazakhstan, Kyrgyzstan, Tajikistan, Afghanistan, Pakistan, and India. All these countries are over 500 miles from WU, which is located around the center of Xinjiang. The *only* shipping method between Xinjiang and these countries was road-shipping until October 11, 1992, when a railroad between WU and Kazakhstan (460 miles) was completed.

Detailed information is not available on the trade volumes between Xinjiang and each of the eight countries. Nevertheless, evidences suggest that the trade volumes were negligible before 1992 due to the poor transportation conditions; since 1992, the WU-Kazakhstan trade has increased dramatically and dominated the foreign trade of Xinjiang. Table 1 illustrates this pattern. The first row shows Xinjiang's total import value for each year between 1990 and 1996. Obviously, the import tripled in 1992 and kept growing fast afterwards. This change is consistent with the completion of the WU-Kazakhstan railroad. In rows two through five, import volumes for steel, chemical fertilizer, industrial chemicals, and paper are reported. Among them, steel import seems to have increased the most. In 1992, steel import doubled; in 1993, steel import had a ten-fold increase. After 1993, steel import shrank but was still

---

[5] This portion is between WU and a city called Weiwu; the length of this portion is about 1,000 miles.

much larger than the before-1992 levels. This suggests that steel import from Kazakhstan constitutes a major portion of the import, especially that of steel, to Xinjiang. In the last row I also show the export value of Xinjiang, and similar patterns as above appear.

3.4 Advantages of the Setting

Two features of the setup simplifies the task of identification. First, rail-shipping is the dominating shipping methods in this case; therefore, the effects from other competing shipping methods are negligible. This significantly reduces the set of confounding factors (and also reduces the bias of the price-gap approach as discussed in the footnote). Second, the upgrading project changes the capacity of the railroad but did little to other attributes like speed or costs. This saves the need to disentangle the effects due to different attributes, and thus simplify predicting the effects of future capacity investment.

4 The Data

The main data required for this study is the market prices in LAN and WU. I compile monthly transaction prices of 35 specific products from the journal *Chinese Prices* [1] spanning the period 1992-1996. From *Chinese Transportation Yearbook* [4] I also obtain information on the LAN-WU railroad shipment volumes. The official rail-shipping rates are complied from the *Chinese Price Yearbook* [2]. For certain commodities, e.g. steel, their yearly local output, local production capacity, local consumption, import volume, and export volume are obtained from the *Xinjiang Statistical Yearbook* [5] and *Chinese Steel Statistical Yearbook* [3]. Below I describe the data in more details.

4.1 Commodity Prices

Transaction prices of over fifty production and consumption goods have been surveyed on the *same day* of each month in 29 capital cities of China ever since 1992.[6] This price data set has

---

[6] Before 1995, the prices were surveyed on the 15th day of each month. After 1995, prices were surveyed on the 25th day of each month.

three main advantages compared with those used in other studies. First, the prices in this study are for highly disaggregated products. This disaggregation is important for this study since, even for the products in the same category, e.g. steel, they may have dramatically different local supply and demand and thus bear different effects from the rail capacity. Moreover, the disaggregated data are also more transparent and make it easy to crosscheck with other sources for data quality. The second advantage of the price data is that the prices are not the monthly average prices; instead, they were spot prices surveyed on the same day for all 29 cities. This is also important since averaging prices over time would tend to smooth the price series and eliminate the peaks of price gaps that are most likely to contain information on the rail capacity effects. The third advantage is that the prices are transaction prices, which may best indicate market conditions and reflect the effects of transportation conditions.

Among the products surveyed, only 35 have observations spanning the before- and after-rail-upgrading periods. Therefore, in the study below I shall focus on these 35 commodities. Among them are 12 industrial goods, including 70# gasoline, 0# diesel, 10-20mm round steel (normal carbon level), 19-24mm thread steel, 6.5mm hot-rolled steel rod, 1mm cold-rolled steel sheet, 1mm hot-rolled steel sheet, 20mm hot-rolled steel sheet, 2-6# angle steel, Sodium Hydroxide (98%), Sodium Carbonate (98.5%), and cement (normal). Among the rest are 23 agricultural products including flour (normal), rice (grade 2), corn flour (grade 2), soybean (grade 1), veggie oil (grade 2), Chinese cabbage (grade 1), cabbage (grade 1), Chinese chives (grade 1), cucumber (grade 1), tomato (grade 1), eggplant (grade 1), radish (grade 1), green pepper (grade 1), potato (grade 1), watermelon (grade 1), fresh pork (boneless), beef (boneless), mutton (boneless), chicken (medium), egg (fresh), belt fish (medium), silver carp (medium), and tofu.

4.2 LAN-WU Railroad Shipment Volumes

From *Chinese Transportation Yearbook* [4] I obtain annual shipment volumes of the LAN-WU railroad. The volumes are measured in weight and are given for the westbound (LAN-to-WU) and eastbound (WU-to-LAN) directions, respectively. In graph 3 I plot the shipment volumes together with the railroad's theoretical capacity during the period 1990-2001. Significantly different patterns emerge for *eastbound* and *westbound* shippings. In general, the eastbound volumes are closely related to the theoretical railroad capacity; moreover, it is visually obvious that the upgrading project completed in 1994 had a significant positive effect on the eastbound volumes. In sharp contrast, westbound volumes were much smaller than the theoretical capacity, and they seemed to be little affected by the rail-upgrading. These patterns suggest that the demand for LAN-WU eastbound shipments was so large that the rail capacity was binding. In contrast, the demand for westbound shipment was small enough such that it is not constrained by the railroad capacity; in other words, the price gaps of the goods shipped westbound should not be affected by the capacity change.

4.3 Rail-shipping Costs

Information on the marginal cost of rail-shipping is unavailable. In the exercise below I shall approximate the cost with the total official shipping rates, which include basic shipping rates and additional fees (e.g. uploading fee). This approximation may be reasonable since most railroads were run by the government and are not for profit. LAN-WU railroad basic shipping rates from 1990 to 2002 are summarized in table 2. Specifically, between 1990 and 1997 the basic shipping rates more than tripled; after 1997, the rates fluctuated and showed a slight upward trend. Detailed information on additional fees is not available, but evidences suggest that they are of about the same size as the basic shipping charges. In evaluating the economic benefit below, I shall check how sensitive the results are to different scenarios of the shipping costs.

5 The Impacts on Price Gaps

The empirical setting as described above allows me to construct the following natural experiment. My treatment group consists of LAN-WU *eastbound* price gaps — the price gaps of the commodities shipped from WU to LAN. By graph 3, I have shown that eastbound shipments were likely to be constrained by the railroad capacity, and thus should be affected by its changes. In terms of reference groups, I have two alternatives. One consists of LAN-WU *westbound* price gaps, which should not be affected by the railroad capacity changes as discussed above.[7] The other natural reference group is the price gaps between LAN and XIN. This is so since the transportation infrastructure between LAN and XIN has changed little during my sample period.

A difference-in-difference type of estimator is particularly suitable for the setting considered. The estimation details and results are presented below.

5.1 Shipping Directions

Before carrying out any analysis, I first need to know a commodity's shipping direction through the LAN-WU rail. This information is not directly available, but it may be inferred as below. First of all, I may examine the commodity's *net flow* as calculated below:

$$\begin{aligned} &Net\ Flow\ from\ Xinjiang\ to\ Other\ Provinces = Xinjiang's\ Local\ Production \\ &-Xinjiang's\ Local\ Consumption\ \ - Xinjiang's\ Export\ to\ Foreign\ Countries. \end{aligned}$$

If the net flow was positive for a product, then at least some of the commodity was shipped eastbound (WU to LAN); therefore, the product's price gaps should be affected by the LAN-WU railroad capacity change. On the other hand, however, if the net flow of a good was negative, I should not conclude that its price gaps was immune to the change of railroad capacity. This is because, even if a product's net flow was negative, it is still possible that some of this good was shipped eastbound.

Alternatively, instead of considering the net flow, I may examine the signs of observed price gaps. Intuitively, a commodity tends to flow from a low-price city to the high-price one.[8] It is important to note that, even if price gaps exist, trades may not happen. Therefore, if a

---

[7] More rigorously, I am implicitly assuming the westbound shipping cost is not affected by the rail capacity. If this assumption does not hold, then the rail-doubling project can affect the westbound price gaps *indirectly* unless the shipping cost is fully controlled for.

[8] A sufficient condition for this is competitive pricing. In this study, competitive pricing is not an unreasonable assumption for most of the products — especially agricultural products, so the signs of the price gaps can be used as a reference.

good was found to be cheaper in LAN than in WU, then it was either not shipped or shipped westbound; in either case, the good's price gaps should *not* be affected by the rail-capacity change. However, if the price is more expensive in LAN than in WU, it is hard to say whether the price gaps should or should not be affected by the rail-capacity change; it depends on whether the good was actually traded, as we do not know.

Based on the two methods above, I find strong evidences for the shipping directions of three commodities — gasoline, diesel, and thin steel sheet. First consider gasoline and diesel. As shown in table 3, their local productions significantly exceeded their local consumptions; moreover, their China-Kazakhstan trade volumes were negligible during my sample period.[9] Therefore, a large fraction of the excess local supply should have flowed *eastbound* through the LAN-WU railroad towards the other provinces. Turning to thin steel sheet, I find a different pattern. The local output of thin steel sheet was zero throughout my sample period; in contrast, its local consumption was quite substantial. Therefore, before 1993, this local consumption was mostly supplied by other provinces of China. After 1993, however, the import of thin steel sheet increased dramatically due to the completion of the WU-Kazakhstan railroad; by 1997, the import volume was *seven times* that of local consumption. It is thus obvious that, after 1993, thin steel sheet should also flowed *eastbound*.

As a further check, I also consider the signs of these products' price gaps — calculated by subtracting the price in WU from that in LAN. In the top-left panel of figure 4 I plot the products' average LAN-WU price gaps. They are all positive between 1992 and 1997, as is consistent with the flow directions as inferred above.

For other products, the net-flow approach is much less effective either because good data are not unavailable or because the implied net flows are too small to be robust to measurement errors. Nevertheless, the sign-of-price-gap approach can still be used to find *westbound* LAN-WU price gaps — prices that are higher in WU than in LAN. As shown earlier, the westbound price gaps should not be affected by the LAN-WU railroad capacity changes.

---

[9] Ideally, I need information on Xinjiang's foreign trade volumes. However, this is unattainable in this study. Nevertheless, since Kazakhstan borders upon no Chinese provinces other than Xinjiang and since Kazakhstan is the only foreign country to which Xinjiang is connected to by a railroad, the China-Kazakhstan trade volume can be used to approximate that between Xinjiang and foreign countries.

5.2 Graphical Analysis

Above I have shown that gasoline, diesel, and thin steel sheet were shipped *eastbound* through the LAN-WU rail. For convenience, I shall call them *class A* products; their price gaps between LAN and WU constitute a treatment group. The remaining 31 products will be referred to as *class B* and their price gaps form two alternative comparison groups: one consists of their *westbound* LAN-WU price gaps — prices that are higher in WU than in LAN, and the other consists of their LAN-XIN price gaps (in both directions).

A distinct pattern of the treatment group, as shown by the top-left panel of figure 4, is the sharp drop of price gaps around the end of 1994.[10] The timing of this change fits quite well with the completion of the LAN-WU upgrading project, suggesting a relation between the drop of price gap and the change of capacity. If this is the case, the reference group should not show the similar pattern of dropping price gaps around the end of 1994. Below I examine this hypothesis.

The remaining three panels of figure 4 plot the average *westbound* LAN-WU price gaps for three categories of class B products — industrial goods, vegetables-fruits, and meat-eggs. The patterns are strikingly different from that of the treatment group: for all three categories, their price gaps showed stable fluctuations during the sample period.

Next I turn to the average LAN-XIN price gaps, as are plotted in figure 5 for class A and class B products. In general, they show significantly different patterns from that of the treatment group. Specifically, the class B LAN-XIN price gaps were also quite stable, showing slight increasing trends over time. As to the average class A LAN-XIN price gaps, I do find a significant drop during the sample period but it happened much earlier than the completion of the upgrading project.

To summarize, the graphical evidences suggest that the LAN-WU rail upgrading project had a significant *negative* impacts on the price gaps of the treatment group.

---

[10] Price gaps seemed to increased quickly before 1994; its causes will be examined in the "robustness check" subsection.

5.3 Difference-in-difference Estimates

The graphical analysis is quantified in this section with a difference-in-difference approach. First of all, I consider the following regression for each of the treatment and control groups:

$$G_{it} = \theta_0 + \theta_p P_{it} + \theta_c C_t + \theta_{prod} D^{products} + \theta_{oct94} D^{oct94} + \epsilon_{it} \tag{4}$$

Here G_{it} is the price gap of a good i at time t. P_{it} indicates the good's price level at time t; in my estimation P_{it} is measured by the price at LAN. C_t is the official rail-shipping rate at time t. $D^{products}$ is a set of dummy variables used to control product-specific fixed effects. D^{oct94} is a dummy variable that is zero for the period January 1993 - October 1994 and is one for the period November 1994 - December 1996.[11] θ_{oct94} thus estimates the change of price gaps before and after the rail-upgrading project. For the moment, assume that the error term ϵ_{it} behaves well.

Table 4 summarizes the OLS estimates. In particular, I consider the price gaps of each of the following four groups as dependent variables — class A LAN-WU (the treatment group), class B LAN-WU, class A LAN-XIN, and class B LAN-XIN. Furthermore, for each group I consider two alternative regressions — one in the original specification (4) and the other in its simplified form, which omits the regressors P_{it} and C_{it}. The estimate of θ_{oct94} is negative and highly significant for the treatment group and is insignificant for the comparison groups. This is certainly consistent with my earlier graphical analysis.

With the above estimates, I then perform a difference-in-difference (D-in-D) type of calculation as below. This calculation is supposed to purge the effects of confounding factors and provide a *net consequence* of the upgrading project on the treatment group price gaps.

Recall that for each of the four groups I have performed two regressions. Now for each group I pick the regression with a higher adjusted R-squared and put its estimate of θ_{oct94} in table 5. For example, for the treatment group the *simple* regression(column 1 of table 4) has a lower adjusted R-squared than the *original* regression (column 2) has, so I pick the estimate for θ_{oct94} from the latter and put it in the top-left cell of table 5. Repeating this for the three

---

[11] Note that information before 1993 is omitted to avoid the confounding effect on price gaps of opening the WU-Kazakhstan railroad and of changing price regulations in 1992, as will be discussed in more details later.

reference groups, I have the four top-left cells of table 5.

Next I subtract column 2 from column 1 to form the last column, and I subtract row 2 from row 1 to form the last row. Among the resulted figures, three are of particular interest and I label them as $DD1$, $DD2$, and DDD. $DD1$ is a D-in-D estimate assuming that confounding factors had similar impacts on the price gaps of all products. $DD2$ assumes that confounding factors affect the LAN-WU and LAN-XIN price gaps similarly. Finally, DDD combines $DD1$ and $DD2$ and requires weaker assumptions. In fact, all these three estimates are close to the original estimate -291.2. Specifically, DDD is -362.1, meaning that the LAN-WU rail-upgrading project squashed the treatment group price gaps by -362.1 yuan per ton on average. This is about 10 percent of the average price level of the treatment group.

5.4 A Robustness Check

Figure 6 plots the price gap time series for each of the four class A products — gasoline, diesel, hot-rolled thin steel sheet, and cold-rolled thin steel sheet. The price gaps of *all* the products experienced sharp drops around mid-1994, suggesting that my findings above was not due to the dominating behavior of any particular product in class A. Quantitatively, I estimate regression (4) for each of the four products and report the results in table 6. Consistent with figure 6, all price gaps deceased significantly after October 1994. The estimates range from -111.1 yuan per ton to -450 yuan per ton.

This evidence also helps us rule out one alternative story that could explain the sudden decrease of price gaps. Specifically, one may argue that it was the opening of Wu-Kasakshtan rail in October 1992 that has generated the patterns observed for the class A price gaps. If this were true, then the price gaps of the thin steel sheet should change while those of gasoline and diesel should not. This is so since, as I have shown, little petroleum products were traded between Xinjiang and Kasakshtan while large quantity of thin steel sheets were. This obviously contradicts with the findings above.

6 The Direct Benefit of Capacity-upgrading

With the results above, the formula (2) may now be applied to infer the short-term direct benefit of the upgrading project in 1995 and 1996. To proceed, I assume that the upgrading project has the same effect on the treatment group and on the prices of other products shipped *eastbound*. This is necessary since only the effects on the treatment group prices have been estimated.

Table 7 summarizes the calculation. The average price gap of the treatment group was 846 yuan per ton in 1993 and 1994. Using the DDD estimate of 362 yuan per ton, I can calculate the average price gap as $846 - 362 = 484$ yuan per ton in 1995 and 1996. As to annual rail capacity, it was about 7 million tons before the project was completed. By 1995 and 1996, the average annual capacity was about 10 million tons per year (this is approximated by the *actual* shipping volume). Therefore, the actual capacity gain in 1995 and 1996 was about $2 \times (10 - 7) = 6$ million tons. Finally, according to table 2, the basic shipping rate for 1995 and 1996 was bout .07 yuan per ton kilometer. This implies that the basic shipping cost through the LAN-WU rail was about 126 yuan per ton, calculated by multiplying .07 with 1,800 kilometers (the length of LAN-WU rail). As mentioned earlier, additional shipping fees should also be included in the total shipping cost and they were about the same size as the basic shipping cost. Therefore, I consider the following three scenarios of marginal LAN-WU operating costs — 100, 150, and 200 yuan per ton. The estimated *short-term direct benefit* under these three scenarios are 3.39, 3.09, and 2.79 billion yuan during 1995-1996. Therefore, the benefit estimates are not very sensitive to the choice of additional shipping fees.

To see how significant these benefits were, I also calculate their implied rate of return. This requires information on the cost of the capacity-upgrading project, which is not directly known but may be inferred as below. According to the *Chinese Transportation Yearbook* (1994, pp.61, pp. 420, and pp. 423), by the end of 1993 half of the upgrading project was completed and about 2.37 billion yuan has been spent. From this I infer that the total project costed between 4 and 5 billion yuan. This implies that the total *nominal* rate of return of the project for the two years 1995 and 1996 was between $2.79/5 = 56\%$ and $3.39/4 = 84\%$.

The inflation rate of the region was between 15 and 20 percent in 1995 and was around 10 percent in 1996 (*Chinese Price Yearbook* 1996 and 1997). This implies a compound inflation rate of about 30 percent. Therefore, the real rate of the return for the project was between 26 and 54 percent during the two years. Taking a geometric mean, this implies an annual *real* return of between 12 and 24 percent!

Even though long-term and indirect benefits have not been taken into account, the estimated rate of return is significant! Considering the particularly low real interest rate in China, the project is obviously profitable.

7 Concluding Remarks

In this study I found that the capacity-upgrading project of a railroad between two Chinese cities generated significant benefit. Specifically, the *real* annual rate of return (calculated from the short-term direct benefit) to the project is estimated to range from 12% to 24% for the two years following the capacity-expansion. This exercise takes advantage of a natural experiment in which the railroad underwent a sudden and predetermined change and in which two natural control groups can be constructed. Without this favorable setting the exercise would be much more expensive, as may have been a reason limiting the development of the literature. This study also shows a strong promise of the *price-gap approach*. Without this approach, the quantification would not have been possible since the existing shipping-cost-based methods are not able to evaluate *capacity-related* infrastructure investment.

A natural next step is to apply the price-gap approach to other settings. China is a particular interesting country to study due to its fast growth and ambitious plans of public capital investment. It is important to note that the approaches employed in this study are not limited to China. The identification techniques in this paper can be readily applied to settings in other countries. For example, the U.S. transportation administration has been using a detailed and regularly updated database of over ten thousand highways across the U.S.. Given the size of this database, there should be occasions in which natural experiments can be constructed to evaluate the economic return. This suggests a potential direction for my future research.

Another natural extension of this study is to employ the favorable empirical setting in this project to study *something* more interesting — the long-term and indirect effects of infrastructure, e.g. productivity gain due to specialization. Relevant data are available in two strains. In the aggregate level, measures of productivity and investment behaviors are available for two-digit industries in the three cities — LAN, WU, and XIN — ever since 1989. Therefore, economic activities aggregated at industry-city levels can be employed to estimate the long-term effects of the rail-upgrading project. In the disaggregate level, firm level panel data in the three cities are also going to be available soon. Once the data are ready, the aggregate analysis can be refined to obtain more convincing and detailed evidences.

References

[1] Chinese price (monthly), 1990–2000.

[2] Chinese price yearbook, 1990–2000.

[3] Chinese steel yearbook, 1990–2000.

[4] Chinese transportation yearbook, 1990–2000.

[5] Xinjiang statistical yearbook, 1990–2000.

[6] *How Federal Spending for Infrastructure and Other Public Investments Affects the Economy.* The Congress of United States, Congressional Budget Office, Washington, D.C., July 1991.

[7] The railway transportation atlas of china (chinese). Chinese Map Publisher, 1994.

[8] David A. Aschauer. Is public expenditure productive? *Journal of Monetary Economics*, 23(2):177–200, 1989.

[9] Edward M. Gramlich. *Benefit-cost Analysis of Government Programs.* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1981.

[10] Peter D. McClelland. Social rates of return on american railroads in the nineteenth century. *The Economic History Review, New Series*, 25(3):471–488, August 1972.

[11] Patrick K. O'Brien. *The New Economic History of the Railways.* New York: St. Martin's Press. London: Croom Helm., 1977.

Table 1: Import and Export of Xinjiang

Unit: yuan/ton×kilometer

Imports of Xinjiang	**90**	**91**	**92**	**93**	**94**	**95**	**96**
Total Value ($10^6 US\$$)	75.0	96.2	296.5	427.0	464.4	659.2	853.9
Steel (ton)	48,019	30,954	65,476	711,759	335,434	356,869	581,945
Chemical Fertilizer (ton)	209,967	298,184	700,187	275,759	191,045	398,680	787,300
Industrial Chemicals (ton)	701	4,262	18,268	3,529	9,951	35,668	17,176
Paper (ton)	4	0	2,649	9,985	4,884	79	52,146
Exports of Xinjiang							
Total Value ($10^6 US\$$)	335.3	363.2	453.9	495.1	576.1	768.8	549.8

Source: *Xinjiang Statistical Yearbook*.

Table 2: Average Basic Shipping Rates of LAN-WU and LAN-XIN Rails

Unit: yuan/ton*kilometer

Time	**LAN-WU**	**LAN-XIN**
Before 3/1/91	.0265	.0265
3/1/91 - 6/30/92	.029	.029
7/1/92 - 9/30/92	.0385	.0385
10/1/92 - 6/30/93	.043	.0385
7/1/93 - 12/25/95	.058	.0535
12/26/95 - 1/31/96	.083	.0785
2/1/96 - 3/31/96	.0802	.0785
4/1/96 - 6/1/97	.0872	.0855

Table 3: Local Production, Consumption, and Trade of the Three Products

Unit: ton

Year	**Product**	**Loc Pro**	**Loc Cons**	**Export**	**Import**
91	Gasoline	1,410,000	785,500	-	-
	Diesel	1,560,000	956,400	-	-
	Thin Sheet	0	98,941	-	-
92	Gasoline	1,590,000	954,600	0	0
	Diesel	1,810,800	1,032,000	0	0
	Thin Sheet	0	100,152	0	1,941
93	Gasoline	1,866,300	1,144,300	0	0
	Diesel	2,005,100	1,056,000	0	3,312
	Thin Sheet	0	97,316	0	135,945
94	Gasoline	1,745,900	1,052,200	-	-
	Diesel	2,140,700	1,162,800	-	-
	Thin Sheet	0	97,316	0	135,945
95	Gasoline	1,790,000	1,019,600	308	0
	Diesel	2,318,800	1,212,000	490	0
	Thin Sheet	0	93,727	0	213,124
96	Gasoline	1,983,200	1,043,000	592	0
	Diesel	2,643,400	1,271,700	0	0
	Thin Sheet	0	77,373	0	626,474
97	Gasoline	2,084,300	973,700	0	0
	Diesel	3,044,200	1,250,600	0	0
	Thin Sheet	0	86,999	0	700,155

Table 4: Estimates of Model Specification (4)

Unit: yuan/ton

	LAN-WU Class A (Eastbound)		**LAN-WU Class B (Westbound)**		**LAN-XIN Class A**		**LAN-XIN Class B**	
	Simple	Original	Simple	Original	Simple	Original	Simple	Original
θ_{oct94}	-291.2^{**} (70.2)	-295.6^{**} (101.9)	70.2 (99.1)	44.1 (120.9)	-39.1 (27.9)	-27.3 (49.1)	38.2 (43.3)	-40.0 (51.0)
θ_p		$.513^{**}$ (.146)		$-.022$ (.035)		.034 (.060)		$.128^{**}$ (.014)
θ_c		-1540 (3311)		3090 (3702)		-1159 (1374)		-2990^{*} (1470)
Obs.	88	57	491	491	125	96	944	944
Adj R^2	.41	.60	.471	.470	.04	-.001	.51	.56

Note: To conform with theory, in the column three and four regressions I have used the westbound price gaps of all products as the dependent variable, so it also includes one observation from the class A products.

Table 5: *Changes* of Price Gaps after October 1994 and Difference-in-difference Analysis

Unit: yuan/ton

	Class A (Col. 1)	**Class** B (Col. 2)	(Col. 1 − Col. 2)
LAN-WU (Row 1)	−291.2	70.2	−361.2 (DD1)
LAN-XIN (Row 2)	−39.1	−40.0	.9
(Row 1 − Row 2)	−252.1 (DD2)	110.2	−362.1 (DDD)

Table 6: Estimates of Model Specification (4) (simple) for Class A

Unit: yuan/ton

	Gasoline	**Diesel**	**Hot-rolled Thin Sheet**	**Cold-rolled Thin Sheet**
θ_{oct94}	-368.0^{**} (102.2)	-111.1 (124.4)	-396.6^{*} (175.1)	-450^{*} (149.1)
Obs.	30	31	19	8
Adj R^2	.29	-.01	.19	.54

Note: The regression specification is $G_{it} = \theta_0 + \theta_{prod} D^{products} + \theta_{oct94} D^{oct94} + \epsilon_{it}$.

Table 7: The Benefit of the Capacity-doubling Project in 1995 and 1996

	Mode 1	**Mode 2**	**Mode 3**
Pre-price-gap (yuan/ton)	846	846	846
Post-price-gap (yuan/ton)	484	484	484
Pre-volume (million tons per year)	7	7	7
Post-volume (million tons per year)	10	10	10
Rail Cost (yuan/ton)	100	150	200
Direct Benefit (billion yuan)	3.39	3.09	2.79

Note 1: Formula (2) for calculating the benefit is reproduced below

$$Incremental\ Benefit = [(P_1^h - P^o) + (P_2^h - P^o))](K_2 - K_1)/2 + K_1 r_1 - K_2 r_2$$

Note 2: The benefit calculated is the total benefit for 1995 and 1996. Therefore, the total capacity gain, $K' - K$, is $(10-7) \times 2 = 6$ million tons. Moreover, since the rail-shipping cost is little affected by the project, $r = r'$.

Figure 3: Total Railroad Capacity and Actual Transportation Volume

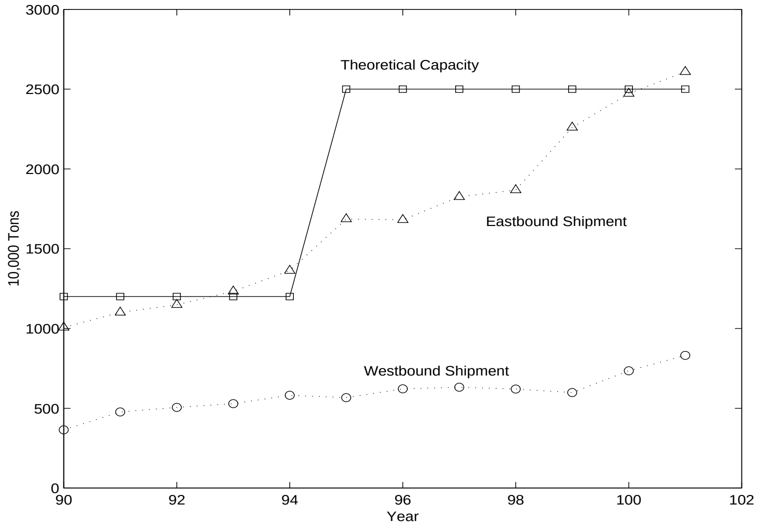

Figure 4: Average LAN-WU Price Gaps: Class A and Westbound Class B

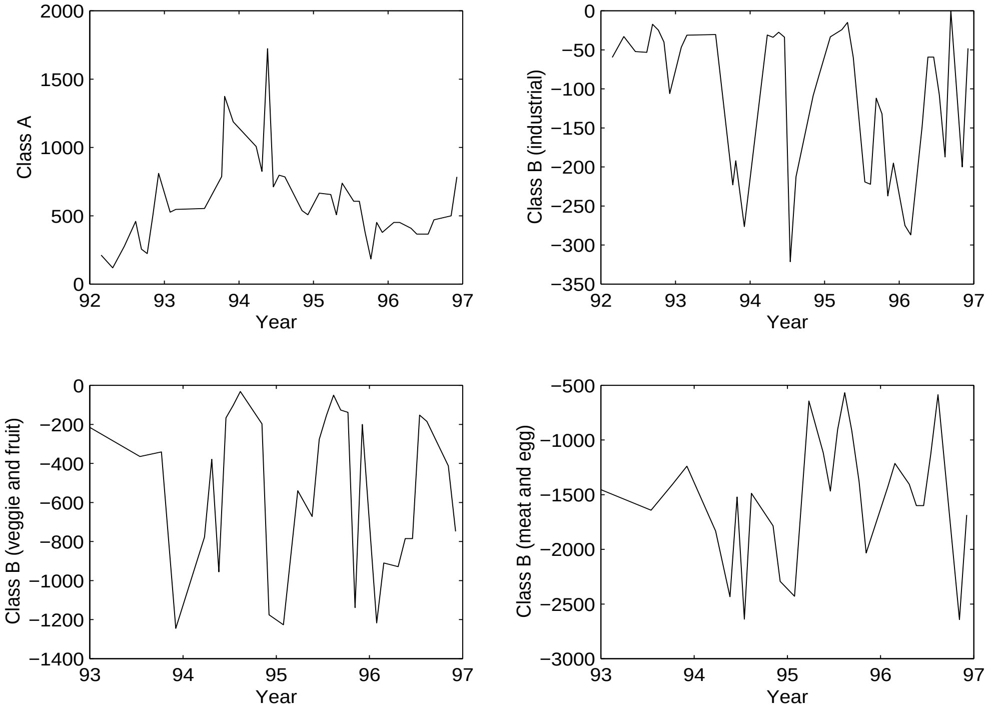

Figure 5: Average LAN-XIN Price Gaps

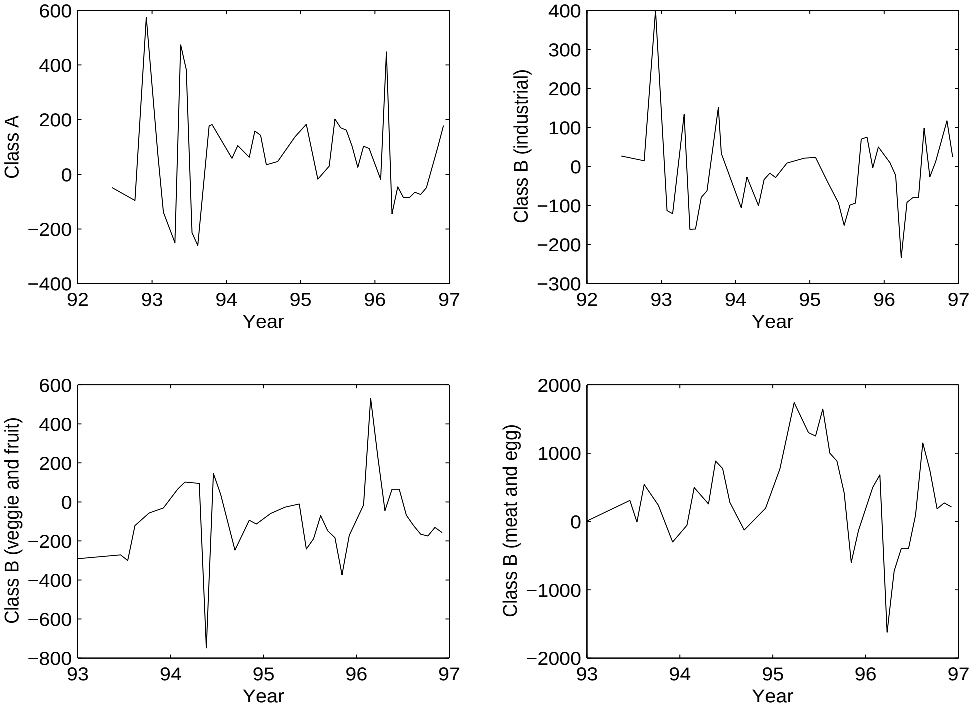

Figure 6: LAN-WU Price Gaps of Class *A* Products