Competition and the welfare gains from transportation infrastructure: Evidence from the Golden Quadrilateral of India*

Jose Asturias
School of Foreign Service in Qatar, Georgetown University

Manuel García-Santana
UPF, Barcelona GSE and CEPR

Roberto Ramos
Bank of Spain

May 13, 2016

Abstract

A significant amount of resources is spent every year on the improvement of transportation infrastructure in developing countries. In this paper, we investigate the effects of one such large project, the Golden Quadrilateral in India, on the income and allocative efficiency of the economy. We do so using a quantitative model of internal trade with variable markups. We find real income gains of 2.71% in the aggregate and that allocative efficiency accounts for 8% of these gains. The importance of allocative efficiency varies greatly across states, and can account for up to 19% of the overall gains. Thus, allocative efficiency can play an important role in determining both the size and distribution of gains from new infrastructure.

*We thank the CEPR for financial support of this project and the Private Enterprise Development in Low-Income Countries (PEDL) program. The authors would like to thank Reshad Ahsan, Chris Edmond, Ben Faber, Rob Feenstra, Thomas Holmes, Joe Kaboski, Timothy Kehoe, Virgiliu Midrigan, Michael Peters, Scott Petty, Andres Rodriguez-Clare, John Romalis, Jim Schmitz, and seminar participants at NBER ITI Winter meeting, NBER Growth Summer Institute, University of Oxford, Carlos III, Barcelona Summer Forum, CREI, University of Notre Dame, Stockholm School of Economics, CEMFI, Tsingua Macro Workshop, SED, ANU in Canberra, WAMS in Melbourne, University of Queensland, Indian Statistical Institute, Delhi School of Economics, REDg Workshop (Madrid), Georgetown Political Economy Lunch Group. The views expressed in this paper are of the authors and do not necessarily reflect the views of the Bank of Spain or the Eurosystem. This paper was previously circulated under the title "Misallocation, Internal Trade, and the Role of Transportation Infrastructure." Asturias: jga35@georgetown.edu; García-Santana: manuel.santana@upf.edu; Ramos: roberto.ramos@bde.es.

1 Introduction

There is by now extensive evidence that allocative efficiency is poor in developing countries. However, little is known about the forces driving these bad allocations. Many developing nations share the problem of substandard transportation infrastructure. Accordingly, the question arises of whether this could be part of the problem. With this in mind, we set out to investigate: Could it be that the poor quality of the transportation infrastructure in developing nations is causing resources to be allocated inefficiently? And if so, to what extent? If, indeed, a subpar transportation network is a significant driver of bad allocations, then designing policies to improve it in these nations could be of use in resolving the problem. This is particularly important since one of the most prominent policy tools in the developing world is building new infrastructure.

We use the example provided by a recent large-scale highway development project in India to delve into this issue from a quantitative perspective. The construction of the Golden Quadrilateral (GQ) provided India with 5,800 km of highway that connected India's four major metropolitan areas (Delhi, Mumbai, Chennai, and Calcutta). This highway is the equal of any similar road infrastructure built in the west. Construction was rapid for a project of this scale, starting in 2001 and 90% completed by the end of 2006.

Perhaps surprisingly, our paper is the first attempt to study how an improvement to infrastructure affects allocative efficiency. The goal of the quantitative exercise is twofold. First, we aim to quantify the income gains from the construction of the Golden Quadrilateral both in the aggregate as well as across states. Second, we seek to decompose these changes in income to find the portion accounted for by allocative efficiency. These results provide an understanding of the both the size and distribution of gains accounted for by improvements in allocative efficiency.

We use a quantitative trade model a la Atkeson and Burstein (2008) in which all of the states of India trade with each other. Firms compete oligopolistically, which implies that firms charge variable markups depending on the level of competition in a market. This framework is useful to study the effects of high transportation costs on the allocative efficiency of the economy. The reason is that the distribution of markups across firms can be mapped to the dispersion in the marginal revenue product of labor (MRPL). Firms with high markups are inefficiently small and have a high MRPL (and vice-versa). Changing transportation costs, by affecting the pattern of spatial competition, will thus impact the distribution of markups and allocative efficiency.

In order to discipline the parameters of our model, we use plant-level data of the country's manufacturing sector and transportation network. We derive a set of structural equations in order to estimate these parameters. In particular, we use a two-step approach that allows us to estimate transportation costs and the elasticity of the sectoral demand curve. This methodology provides a straight-forward way of identifying these parameters in a manner that is fully consistent with the model.

In the first step, we estimate transportation costs between Indian states. To do so, we show that transportation costs can be identified by comparing the prices charged across locations by firms that are monopolistic producers at the national level. This is the case because the prices charged by these firms only depend on transportation costs since the level of competition they face is constant across space. To implement this strategy, we first identify all the goods that are produced by only one plant in India. For these goods we regress the prices paid across destinations with the effective distance between origin and destination. This measure of effective distance is the lowest cost path given the infrastructure quality in place at the time.

In the second step, we estimate the elasticity of the sectoral demand curve. This parameter is important since it governs the size of markups for firms with a large degree of market power. We use the fact that, for goods produced by monopolistic producers, the model implies a standard gravity equation that relates internal flows to transportation costs. Using the transportation costs from the first step, we find the elasticity of the sectoral demand curve that is consistent with the gravity equation for monopolistic products in the data.

We use our calibrated model to quantify the effects of the construction of the GQ. To do so, we compare outcomes from the model when we feed in the estimated transportation costs with and without the GQ.

We find aggregate real income gains of 2.71%, equivalent to $4.1 billion per year. A back of the envelope calculation shows that these gains are large relative to the initial construction costs, which are $5.6 billion. Thus, our results imply that it would take less than two years for India to recover the initial construction cost. We also find a high degree of heterogeneity in income changes across states, including some states that lose.

We find that allocative efficiency accounts for 8% of the overall gains. We also find that there are large differences in the importance of allocative efficiency gains across states. In fact, allocative efficiency can account for up to 19% of the overall gains at the state level. These gains are concentrated in the largest states since these are the states with the lowest levels of allocative efficiency. This is due to the fact that the largest states tend to have lower wages, a standard result of Ricardian trade models. These lower wages provide a cost advantage to local firms, which allows them to charge high markups.

Finally, we conduct an empirical exercise to examine if there is evidence in the data for the main mechanisms of the model. We estimate a differences-in-differences specification in which we compare economic outcomes for districts close to the GQ with those that are far away, before and after the construction of the highway. First, we show that prices paid for intermediate inputs declined more for areas close to the GQ. Second, we find that the Olley and Pakes (1996) covariance term between size and productivity increased more for areas close to the GQ, suggesting that the GQ improved allocative efficiency. These empirical findings are consistent with the model output.

The remainder of the paper is organized as follows. In Section 2, we present the related literature. In Section 3, we describe the main characteristics of the road network in India. In

Section 4, we present the model. In Section 5, we describe the data used. In Section 6, we discuss the calibration of the model. In Section 7, we present and discuss our quantitative results. In Section 8, compare model output with reduced from exercises in the data. In Section 9, we present results from sensitivity exercises. Finally, in Section 10 we conclude.

2 Related Literature

Misallocation Our paper builds on the recent literature that emphasizes misallocation of resources across firms as one of the main sources of TFP differences across countries. In their highly influential paper, Hsieh and Klenow (2009) show that wedges between the marginal products of factors may account for up to 60% of the TFP gap between India and the United States. Other papers in this literature include Guner, Ventura, and Xu (2008), Restuccia and Rogerson (2008), and Peters (2013).

In contrast to the existing literature, we quantitatively study the effects of a policy that actually took place instead of a counterfactual that removes all misallocation in the economy. Thus, we quantify a specific distortion which might be behind the low levels of allocative efficiency in developing countries. Second, the construction of the GQ provides a quasi-natural experiment, which allows us to check that the main mechanisms of the model are present in the data.

Gains from new transportation infrastructure We also contribute to the literature that analyzes the income gains from new transportation infrastructure using general equilibrium models of trade. Papers in this area include Adamopoulos (2011), Donaldson (Forthcoming), Donaldson and Hornbeck (Forthcoming), Herrendorf, Schmitz, and Teixeira (2012), Redding and Turner (2015). Within the context of India, examples include Alder (2014) and Van Leemput (2015).

The existing work typically uses an Eaton and Kortum (2002) model to analyze gains from new transportation infrastructure. In that model, or in any of the other workhorse models considered by Arkolakis, Costinot, and Rodriguez-Clare (2012), there is no scope for gains from allocative efficiency. Our results show that improvements in allocative efficiency can be an important channel of income gains. Furthermore, our framework implies a different set of gains both in the aggregate as well as across states relative to models that are typically used.

Methodologically, we apply the two-step procedure used by Donaldson (Forthcoming). He uses data of products produced in only one location to identify transportation costs and the parameter governing the trade elasticity in his model. This two-step procedure is useful since it is a way of disciplining parameters of the model in a setting in which aggregate trade flows are not observed. We show that this procedure is consistent with a model of oligopolistic competition when applied to monopolistic producers. Thus, by exploiting our plant level data set, we are able to identify both transportation costs and the elasticity of substitution across sectors using a gravity approach.

Pro-competitive gains in international trade Lastly, this paper contributes to the active debate in international trade relating to the size of pro-competitive gains. Gains in allocative efficiency are equivalent to pro-competitive gains since they are due to changing markups. Prominent papers in this large literature include Arkolakis, Costinot, Donaldson, and Rodríguez-Clare (2015), de Blas and Russ (2015), Dhingra and Morrow (2014), Edmond, Midrigan, and Xu (2015), Epifani and Gancia (2011), Feenstra (2014), Feenstra and Weinstein (Forthcoming), Holmes, Hsu, and Lee (2014).[1] The most closely related paper to ours is that of Edmond, Midrigan, and Xu (2015). The authors use an Atkeson and Burstein (2008) model to study the size of pro-competitive gains in a context in which Taiwan trades with the rest of the world.

To the best of our knowledge, ours is the first paper to quantitatively study the size of pro-competitive gains in a setting with many non-symmetric economies. The fact that the economies are not symmetric plays a key role in determining both the size and the distribution of pro-competitive gains. First, pro-competitive gains are concentrated in states with low wages. Low wages in those states imply that firms located there will be able to charge high markups. Second, for some states we find that changing wages can account for large fractions of the pro-competitive gains.

Finally, we find that a setting with a-symmetric economies plays an important role in another determinant of income, which is related to the aggregate markup charged on exported goods relative to those that are imported. This can be interpreted as the effect of markups on a state's terms of trade, since it affects the price of exported vs. imported goods. We find that income changes through this channel can be quantitatively important in some states.

3 Roads in India and the Golden Quadrilateral

India has the second largest road network in the world, spanning approximately 3.3 million kilometers. It comprises expressways, national highways (79,243 km), state highways (131,899 km), major district highways, and rural roads. Roads play an important role in facilitating trade in India: approximately 65 percent of freight in terms of weight and 80 percent of passenger traffic are transported on roads.[2] National highways are critical since they facilitate interstate traffic and carry about 40 percent of the total road traffic.

At the end of the 1990s, India's highway network left much to be desired. The major economic centers were not linked by expressways, and only 4% of roads had four lanes. In addition to the limited lane capacity, more than 25% of national highways were considered to be in poor surface

[1] Workhorse international trade models with variable markups include: Bernard, Eaton, Jensen, and Kortum (2003) and Melitz and Ottaviano (2008).

[2] The importance of railroads has declined in India over time. Although in 1950 more than 80% of freight traveled by rail, this figure has steadily been decreasing. At present, rail carries mostly bulk freight such as iron, steel, and cement. Non-bulk freight represents only around 3 percent of total rail freight in terms of ton-km.

condition.

Congestion was also an important issue, with 25% of roads categorized as congested. This was due to poor road conditions, increased demand from growing traffic, and crowded urban crossings. Frequent stops at state or municipal checkpoints for government procedures such as tax collection or permit inspection also contributed to congestion (see World-Bank (2002)).

In order to improve this situation, the Indian government launched the National Highways Development Project (NHDP) in 2001. The goal of the initiative was to improve the performance of the national highway network. The first phase of the project involved the construction of the Golden Quadrilateral (GQ), a 5,800 km highway connecting the four major metropolitan areas via four- and six-lane roads. The four metropolitan centers that were connected are Delhi, Mumbai, Chennai, and Calcutta. Apart from the increase in the number of lanes, additional features of a high-quality highway system were constructed. These features include grade separators, over-bridges, bypasses, and underpasses.

The cost was initially projected to be 600 billion rupees (equivalent to $13.4 billion in 2006). As of October 2013, the total cost incurred by the Indian government was approximately half of the projected sum (250 billion rupees or $5.6 billion). In Section 7, we compare this cost with the benefits predicted by our model.

The second phase of the NHDP consists in the construction of the North-South and East-West corridor, a highway that aims to connect Srinagar in the north to Kanyakumari in the south and, Silchar in the east to Porbandar in the west. Although this second phase was approved in 2003, there have been many delays for its construction, and less than 10% of the work was completed by the end of 2006. Thus, we will not consider that project in our analysis.

Geospatial data We have geospatial data for all the National Highways of India, which was supplied by ML Infomap. We complement this data using information provided by the National Highways Authority of India (NHAI) on the completion dates of various portions of the GQ. The GQ consisted of 127 stretches and we have detailed information about the start and end points.[3] Figure I shows the evolution of the GQ (in red) in 2001 and 2006. Although the GQ was finished in 2013, more than 90 percent of the project was completed by 2006. We will link this geospatial data to manufacturing data for 2001 and 2006.

4 Model

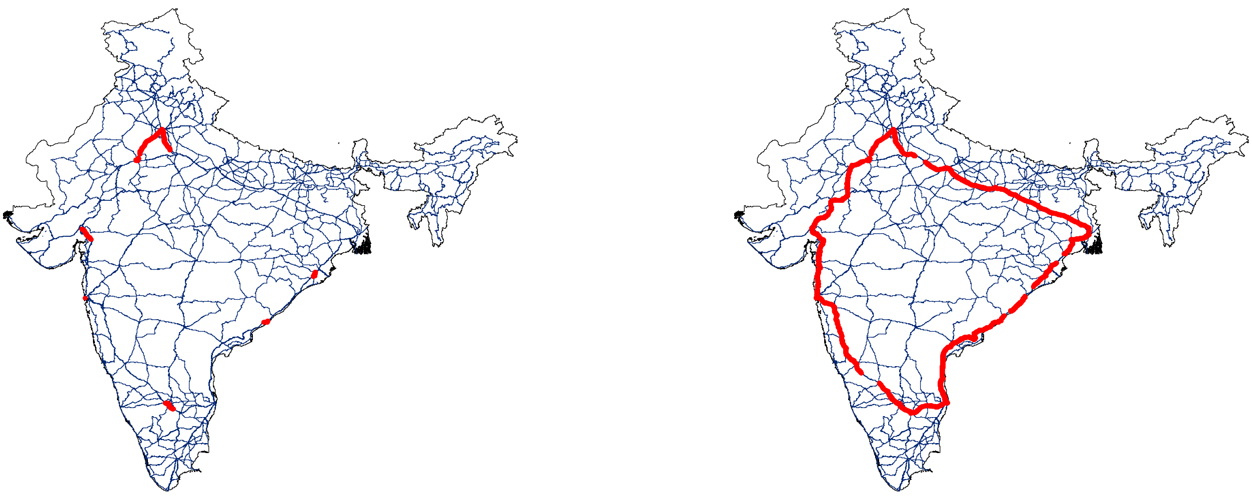
In this section, we present our static general equilibrium model of internal trade. We consider N asymmetric states trading with each other. In each state, there is a measure 1 of sectors. Within each sector, there is a finite number of firms that compete in an oligopolistic manner. Labor is

[3]See nhai.org/completed.asp and the Annual Reports of NHAI.

Figure I

Road Network in India and the GQ

A: GQ construction in 2001 B: GQ construction in 2006



Panel A of Figure I shows a map with the road network in India at the end of 2001, including the sections of the Golden Quadrilateral that were finished by then (around 10% of the total project). Panel B shows the same map but for 2006 (more than 90% of the total project).

immobile across states.[4]

4.1 Consumers

In each state n, there is a representative household with a utility function:

$$C_n = \left(\int_0^1 C_n(j)^{\frac{\theta-1}{\theta}} dj \right)^{\frac{\theta}{\theta-1}}, \tag{1}$$

where $C_n(j)$ is the composite good of sector j and $\theta > 1$ is the elasticity of substitution across composite goods of different sectors. The sector-level composite good is defined as:

$$C_n(j) = \left(\sum_{o=1}^{N} \sum_{k=1}^{K_{oj}} c_n^o(j,k)^{\frac{\gamma-1}{\gamma}} \right)^{\frac{\gamma}{\gamma-1}}, \tag{2}$$

where $c_n^o(j,k)$ is the good consumed by state n and provided by firm k in sector j shipped from state o, N is the number of states, K_{oj} is the number of firms that operate in sector j in state o, and $\gamma > 1$ is the elasticity of substitution between goods produced by different firms in the same sector. We assume that $\gamma > \theta$, which means that goods are more substitutable *within* sectors than *between* sectors.

[4] Interstate migration flows in India are among the lowest in the world. According to the 2001 Indian Population Census, around 96% of people report to be living in the state where they were born.

The budget constraint of the representative household in state n is given by:

$$\int_0^1 \left(\sum_{o=1}^{N} \sum_{k=1}^{K_{oj}} p_n^o(j,k) c_n^o(j,k) \right) dj = W_n L_n + \Pi_n, \tag{3}$$

where W_n is the equilibrium wage, L_n is the labor endowment, and Π_n is the income derived from the profits of firms located in n. Note also that $C_n = W_n L_n + \Pi_n$.

4.2 Firms

In each sector j in state o, there is a finite number of K_{oj} firms. Firms draw their productivity from a distribution with CDF $G(a)$. A firm with a productivity level a has a constant labor requirement of $1/a$ to produce one unit of good. Because firms do not pay a fixed cost to operate in a market, they sell to all N states.

To determine the firm's pricing rule, we first find the demand it faces. Equations (1), (2), and (3) generate the demand:

$$c_n^o(j,k) = \left(\frac{P_n}{P_n(j)} \right)^{\theta} \left(\frac{P_n(j)}{p_n^o(j,k)} \right)^{\gamma} C_n, \tag{4}$$

where

$$P_n(j) = \left(\sum_{o=1}^{N} \sum_{k=1}^{K_{oj}} p_n^o(j,k)^{1-\gamma} \right)^{\frac{1}{1-\gamma}} \tag{5}$$

is the price index for sector j in state n and

$$P_n = \left(\int_0^1 P_n(j)^{1-\theta} dj \right)^{\frac{1}{1-\theta}} \tag{6}$$

is the aggregate price index in state n. Intuitively, the relative demand for a differentiated good within a sector depends on the price of the good relative to the price of the composite good of the sector, and also on the price of the composite good of the sector relative to the aggregate price index.

Firms within sectors compete à la Cournot. Firm k located in state o selling to state d takes the demand characterized by equation (4) and the quantity supplied by competitor firms in the sector as given and solves the following problem:

$$\pi_d^o(j,k) \;= \max_{c_d^o(j,k)} \; p_d^o(j,k) c_d^o(j,k) - \frac{W_o \tau_d^o}{a_o(j,k)} c_d^o(j,k), \tag{7}$$

where $a_o(j,k)$ is the productivity of firm k in sector j producing in state o, τ_d^o is the iceberg transportation cost to ship one unit of good from o to d. Note that, because of the constant returns to scale technology, the problem of a firm across all different destinations can be solved independently. The solution to this problem is:

$$p_d^o(j,k) = \frac{\epsilon_d^o(j,k)}{\epsilon_d^o(j,k) - 1} \frac{W_o}{a_o(j,k)} \tau_d^o, \tag{8}$$

where

$$\epsilon_d^o(k,j) = \left(\omega_d^o(j,k)\frac{1}{\theta} + (1-\omega_d^o(j,k))\frac{1}{\gamma}\right)^{-1}, \tag{9}$$

and $\omega_d^o(k,j)$ is the market share of firm k producing in state o in sector j selling to state d:

$$\omega_d^o(k,j) = \frac{p_d^o(j,k)c_d^o(j,k)}{\sum_{o=1}^{N}\sum_{k=1}^{K_{oj}} p_d^o(j,k)c_d^o(j,k)}. \tag{10}$$

The price that firms set in equation (8) is similar to the markup over marginal cost that is found in a setup with monopolistic competition. The difference is that the markups are endogenous here, and depend on the market structure of the sector. For example, suppose that there is only one firm in a given sector, then that firm will compete only with firms operating in other sectors and its demand elasticity will be equal to θ. This means that the firm faces the sector-level elasticity of demand. At the other extreme, suppose that a firm's market share is close to zero, then the firm will compete only with firms in its own sector and its elasticity of demand will be equal to γ. Notice that a given firm will generally have different market shares and hence charge different markups across different destinations.

The aggregate profits of firms in state n are characterized by:

$$\Pi_n = \int_0^1 \left(\sum_{d=1}^{N}\sum_{k=1}^{K_{nj}} \pi_d^n(j,k)\right) dj. \tag{11}$$

4.3 Balanced Trade and Labor-Clearing Condition

All states n must have balanced trade:

$$\int_0^1 \left(\sum_{o=1,\ o\neq n}^{N}\sum_{k=1}^{K_{oj}} p_n^o(j,k)c_n^o(j,k)\right) dj = \int_0^1 \left(\sum_{d=1,\ d\neq n}^{N}\sum_{k=1}^{K_{nj}} p_d^n(j,k)c_d^n(j,k)\right) dj. \tag{12}$$

The labor-clearing condition for state n is:

$$\int_0^1 \left(\sum_{d=1}^{N}\sum_{k=1}^{K_{nj}} \frac{c_d^n(j,k)}{a_n(j,k)}\tau_d^n\right) dj = L_n. \tag{13}$$

4.4 Definition of Equilibrium

Equilibrium. For all states n and n', sectors j, and firms k_{nj}, an equilibrium is a set of allocations of consumption goods $\{c_{n'}^n(j,k),\ C_n(j)\}$, firm prices $\{p_{n'}^n(j,k)\}$, sector prices $\{P_n(j)\}$, and aggregate variables $\{W_n,\ P_n,\ \Pi_n\}$ such that:

1. Given firm prices, sector prices, and aggregate variables, $\{c_{n'}^n(j,k)\}$ is given by (4), $C_n(j)$ by (2), and they solve the consumer's problem in (1), and (3).

2. Given aggregate variables, $p_{n'}^n(j,k)$ is given by (8), (9), and (10), and solves the problem of the firm in (7).

3. Aggregate profits satisfy (11), aggregate prices satisfy (6), and sector prices satisfy (5).

4. Trade flows satisfy (12).

5. Labor markets satisfy (13).

4.5 Misallocation in the Model

Misallocation in this setting arises due to *dispersion* in markups across producers: the marginal revenue product of labor (MRPL) of firms with high markups becomes inefficiently high, which implies that the goods produced by these firms are under-consumed relative to the goods produced by firms with low markups.[5] As emphasized by Peters (2013), this heterogeneity in markups can generate misallocation. The model is hence relevant to think about the cross-firm misallocation emphasized by Hsieh and Klenow (2009).

These papers have interpreted this misallocation as resulting from government policies that create idiosyncratic distortions at the firm level, which affect the optimal decision of firms. In our model, dispersion in MRPL is caused by dispersion in the market power, which translates into variations in markups: firms with higher productivity draws charge higher markups because they are able to capture bigger market shares. The constant-returns-to-scale technology implies that the MRPL of a firm operating in state o is $W_o{}^{\epsilon(j,k)}/(\epsilon(j,k)-1)$. Thus, firms with high productivity draws (and high markups) also have a high MRPL.

This misallocation is hence similar in nature to the one studied by Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), for the particular case in which the size of the idiosyncratic distortions of firms is positively correlated to their productivity. Firms with high productivity draws are smaller in size than in the case of perfect competition. Thus, India's aggregate welfare should increase by reallocating labor from firms with low productivity draws (low-markup firms) to firms with high productivity draws (high-markup firms).

4.6 A framework to decompose the effects of the GQ

We can apply the framework developed by Holmes, Hsu, and Lee (2014) (HHL) to decompose the changes in real income in our model in a way that highlights the various mechanisms at work. The framework allows us in particular to distinguish between Ricardian, allocative efficiency, and markups terms of trade effects from lowering transportation costs.

We now introduce some notation for the purpose of the decomposition. First, we define the aggregate markups on the goods sold. This reflects how much market power firms producing in a

[5] MRPL is the price of the good multiplied by the marginal product of labor. This is equivalent to the TFPR in Hsieh and Klenow (2009) since labor is the only factor of production, and the production function exhibits constant returns to scale.

state have when selling to other states. First, the revenue-weighted mean labor cost share for the products sold by state n is:

$$c_n^{sell} = \int_0^1 \left(\sum_{d=1}^{N} \sum_{k=1}^{K_{nj}} c_d^n(j,k) s_d^n(j,k) \right) dj,$$

where $s_d^n(j,k)$ is the share of state n's revenue that comes from goods produced by firm k in sector j and sold in state d. The aggregate markup on the goods sold can be expressed:

$$\mu_n^{sell} = \frac{R_n}{W_n L_n} = \frac{1}{c_n^{sell}},$$

where $R_n = W_n L_n + \Pi_n$, which is the state's total revenue. Note that there is an analogous expression at the firm level which is that the firm's markup is equal to the reciprocal of the labor share.

We next define the aggregate markups on the goods purchased by state n, which reflect how much market power firms located in other states have when selling to state n. The revenue-weighted mean labor cost for the products purchased by state n is:

$$c_n^{buy} = \int_0^1 \left(\sum_{o=1}^{N} \sum_{k=1}^{K_{oj}} c_n^o(j,k) b_n^o(j,k) \right) dj.$$

where $b_n^o(j,k)$ is the share of expenditures in state n on goods produced by firm k in sector j located in state o. The aggregate markups on the goods purchased are:

$$\mu_n^{buy} = \frac{1}{c_n^{buy}}.$$

Lastly, let P_n^{pc} be the aggregate price of state n if every firm engages in marginal cost pricing. P_n^{pc} is the aggregate price index that would emerge in a context of perfect competition. This price index depends on the factors that determine the marginal cost of firms: the distribution of firm productivity, the wages paid by firms, and the transportation costs that these firms face.

Using this notation, the real income in state n can be rewritten into the following components:

$$Y_n = \underbrace{W_n L_n}_{\text{Labor income}} * \underbrace{\frac{1}{P_n^{pc}}}_{\text{Prod. efficiency}} * \underbrace{\frac{\mu_n^{sell}}{\mu_n^{buy}}}_{\text{Markup ToT}} * \underbrace{\frac{P_n^{pc}}{P_n} \mu_n^{buy}}_{\text{Allocative efficiency}} \tag{14}$$

The first component is the aggregate *labor income*. The second component is the *productive efficiency* component of welfare. This component is simply the inverse of the price index if all firms charged the marginal cost. The third component is the *markups terms of trade*. This component compares the aggregate markups charged for the goods a state sells with those that it purchases. The last component is *allocative efficiency*. This term is related to the welfare loss that arises due to the dispersion in markups, which results in misallocation. In a situation in which there is no variations on markups, or when there is no misallocation, this index is equal to one. As misallocation increases, this index decreases.[6]

Combining the first two terms leads to an expression that is equal to real income if firms charge their marginal cost. This expression maps back to welfare in the large class of models considered by Arkolakis, Costinot, and Rodriguez-Clare (2012), in which the markups of firms remain unchanged. Thus, we consider changes in this component to be Ricardian effects. Given the expression in equation (14), we decompose the changes in real income into the following terms:

$$\Delta \ln Y_n = \underbrace{\Delta \ln W_n L_n + \Delta \ln \frac{1}{P_n^{pc}}}_{\text{Ricardian}} + \underbrace{\Delta \ln \frac{\mu_n^{sell}}{\mu_n^{buy}}}_{\text{Markup ToT}} + \underbrace{\Delta \ln \frac{P_n^{pc}}{P_n} \mu_n^{buy}}_{\text{Allocative efficiency}}$$

5 Manufacturing Plant-Level Data

In this section, we describe the construction of the data set used in the paper. We link plant-level data on the Indian manufacturing sector with geospatial data. We do so for two snapshots in time (2001 and 2006). Altogether, the data provides the necessary information to analyze how changes in infrastructure quality affect the manufacturing sector.

5.1 Annual Survey of Industries and National Sample Survey

We first construct a representative sample of the Indian manufacturing sector. To do so, we merge two separate sets of plant-level data: the Annual Survey of Industries (ASI) and the National Sample Survey (NSS). The ASI targets plants that are in the formal sector. It is the main source of manufacturing statistics in India and has been commonly used in the development literature.[7] It covers plants that have more than 10 workers if they have electricity and 20 if they do not. The information provided by the establishments is very rich, covering several operational characteristics such as sales, employment, wage bill, capital stock, and intermediate goods usage. The NSS covers all informal establishments in the Indian manufacturing sector. "Informal" refers to all manufacturing enterprises not included in the ASI. The process of merging the ASI and NSS data is straightforward since very similar questions are used to collect both sets of data. The final product contains 17 million manufacturing plants that employ 45 million workers.

It is important to note the huge differences in productivity between formal and informal plants. Informal plants account for approximately 80% of employment and only 20% of value-added. Thus, it is crucial to merge these data sets to have an accurate picture of the Indian manufacturing sector.

[6] It can be shown that this term is equal to the cost of one unit of utility under marginal cost pricing divided by the cost of acquiring one unit of utility with the equilibrium bundle under marginal cost pricing.

[7] See for instance Aghion, Burgess, Redding, and Zilibotti (2005), Chari (2011), Hsieh and Klenow (2009), and Bollard, Klenow, and Sharma (2013).

5.2 Prices and the Consumption of Intermediates in ASI-NSS

The ASI and NSS data contain detailed information about production and intermediate inputs usage. For each plant in our data, we observe the value and physical quantity of production and intermediate goods usage broken down by product. This means that we can compute the input prices paid by plants, which allows us to identify transportation costs.[8] To compute the price of inputs, we divide the expenditure on a particular good by physical units.

The product classification used in both the ASI and NSS is the Annual Survey of Industries Commodities Classification (ASICC). The ASICC contains approximately 5,400 different products, which are very narrowly defined. For instance, the ASICC distinguishes between different types of black tea such as leaf, raw, blended, unblended, and dust. In the processed mineral category, the ASICC distinguishes between 12 different types of coke.

6 Inferring Parameter Values

We calibrate our model to 2006, when the GQ was already in place. Our calibration strategy is as follows. Our model is characterized by (i) a set of bilateral iceberg costs between states (a 29 by 29 matrix of iceberg costs), (ii) the elasticity of substitution across sectors θ, (iii) the elasticity of substitution within sectors γ, (iv) the number of producers in state i and sector j K_{ij}, (v) the labor endowment of states, and (vi) the parameters governing the productivity distribution of firms.

Using structural equations from the model, we first estimate the transportation costs and the two elasticities (Sections 6.1, 6.2, 6.3, and 6.4). We next plug into the model the number of firms per state-sector that we observe in the data, and calibrate the labor endowment of the states and the productivity distribution to match the relevant statistics of the Indian manufacturing sector (Section 6.5).

6.1 Estimating Transportation Costs

The first step is to infer transportation costs. To do so, we use pricing data from intermediate inputs used across India as described in Section 5.2. Equation (8) shows that the prices charged by firms depend both on transportation costs and on firms' market shares in the destination market. In order to identify transportation costs, we exploit one implication of the model: variations in prices for nation-wide monopolists are due solely to variations in transportation costs across destinations. To see this formally, equation (8) and the fact that a monopolist has a market share of one in all destinations imply that the firm will charge:

[8] Although these data sets are starting to be widely used (see Garcia-Santana and Pijoan-Mas (2014) and Hsieh and Klenow (2009) for example), not much attention has been paid to the price information. A notable exception is Kothari (2013).

$$p_d^o(j,k) = \frac{\theta}{\theta-1}\frac{W_o}{a_o(j,k)}\tau_d^o. \tag{15}$$

Then, the relative price across destinations is:

$$\frac{p_d^o(j,k)}{p_{d'}^o(j,k)} = \frac{\tau_d^o}{\tau_{d'}^o},$$

which only depends on the ratio of transportation costs. Hence, through the lens of our model, the prices charged by monopolists across destinations reveal differences in transportation costs.

Empirically, we define a monopolist as a plant selling at least 95 percent of the value of a given 5-digit ASICC product nationally. Using the ASI and NSS for the years 2001 and 2006, we identify 165 products that are manufactured by monopolists. The largest category is "Manufacture of chemicals and chemical products," which contains around 40 percent of the identified products. This is consistent with the nature of the chemical industry, in which production is often concentrated in one plant due to economies of scale, with the product then shipped to many locations.[9]

Once the products manufactured by monopolists are identified, we use the price paid for intermediate inputs in order to estimate equation (15). The strategy is similar to the one used by Donaldson (Forthcoming), except we work with plant-level data and with a framework that accommodates oligopolistic competition. In our empirical specification, we parametrize τ_d^o with effective distance. This measure computes the least costly path to travel from origin to destination, taking the road network and the variation in road quality into account.

In order to compute effective distance, we first convert the national highway network into a graph. The graph consists of a series of nodes that are connected by arcs. In our case, a node is the most populous city in each district and an arc is the road that connects these cities. An arc is referred to as being GQ or non-GQ depending on whether it was completed in a specific year. Each road segment is assigned a cost:

$$\begin{aligned} \text{Effective Distance}_{n_2}^{n_1} &= \text{Road Distance}_{n_2}^{n_1} \text{ if GQ} = 0 \\ \text{Effective Distance}_{n_2}^{n_1} &= \alpha\text{Road Distance}_{n_2}^{n_1} \text{ if GQ } = 1, \end{aligned} \tag{16}$$

where n_1 and n_2 are nodes, and α indicates the effective distance of the GQ relative to stretches of road that are not GQ. We use a value of $\alpha = 0.52$, which is based on average speeds calculated by the World Bank.[10] We then use Dijkstra's shortest-path algorithm to construct a matrix of

[9]A description of the production structure of the chemical industry in India can be found at http://smallb.in/sites/default/files/knowledge_base/reports/IndianChemicalIndustry.pdf

[10]The value of α is based on the fact that the average speed on a national highway is between 30 and 40 km/h according to World-Bank (2002). By contrast, the average speed on the GQ is estimated to be around 75 km/h. This can be computed by calculating the predicted average speed traveling from a random sample of origins-destinations over GQ roads using Google Maps (see Alder (2014)).

lowest-cost routes between all the districts for the years 2001 and 2006. The sets of bilateral effective distances in these two years are different since the algorithm internalizes the fact that traveling on a better quality road, that is completed stretches of the Golden Quadrilateral, is less costly.

We take equation (15) to the data by regressing prices on our measures of effective distance. We use a flexible specification of effective distance in order to capture non-linearities in transportation costs. Such a flexible specification is commonly used to estimate the parameters of trade models using gravity equations, such as in Eaton and Kortum (2002). We estimate equation (15) as follows:

$$\log p^o_{d,t}(j,k) = \sum_{\ell=1}^{10} \beta_\ell \mathbb{I}\{\text{Effective Distance}^o_{d,t} \in \text{decile } \ell\} + \sum_{o,t} \delta_{o,t} + \sum_{j,t} \alpha_{j,t} + \epsilon^o_{d,t}(j,k) \tag{17}$$

where $p^o_d(j,k)$ is the weighted average of the prices paid by plants in district d in year t for the intermediate input j produced by a monopolist located in district o, $\mathbb{I}$ is an indicator function that takes value 1 if the condition within brackets is satisfied, $\delta_{o,t}$ are district of origin fixed effects that vary by year, $\alpha_{j,t}$ are product fixed effects that vary by year, and $\epsilon^o_{d,t}(j,k)$ is the error term. The origin fixed effects control for local wages and the product fixed effects control for firm productivity. Note that origin and product fixed effects are time dependent, which implies that the identification comes from the cross-sectional variation. We estimate equation (17) at the district level instead of the state level, in order to exploit all possible variation in the data.[11]

Table I presents the results from estimating equation (17). Column (1) uses data from 2001 and 2006, whereas column (2) only uses data from 2006. We use column (1) as our baseline specification. In both cases, we find that prices increase significantly over long distances. For example, prices are 51% higher in the 10th decile than in the 1st decile. The 10th decile includes districts located more than 1,800 kilometers away in effective distance, which is approximately the road distance from New York City to Des Moines, Iowa.

Although the overall pattern is that prices increase over long distances, the estimates are non-monotonic over shorter distances. For example, in column (1), the coefficient associated with the third decile is 7 percentage points lower than the coefficient in the second decile. In order to avoid having non-monotonic transportation costs in the model, we assume that the relationship between iceberg costs and effective distance is given by a discrete monotonically increasing cubic function $g(\text{Decile}^o_d)$, where Decile^o_d indicates the corresponding decile between o and d. We first normalize iceberg costs in the first decile to 1. The resulting iceberg costs from the regression are

$$\hat{\tau}^o_d = e^{\hat{\beta}_{\text{Decile}^o_d}} \tag{18}$$

[11] In order to avoid noisy estimates, we clean the data in several dimensions. First, we exclude input items whose description refers to "other" or "non elsewhere classified" products. Second, we exclude goods that are consumed in at least five districts. Finally, we identify unit misreporting in several goods, which generates large jumps in prices. See the Appendix for more details.

We then find the parameters of the cubic polynomial $g(\text{Decile}_d^o)$ that best fit the iceberg costs by equation (18).

In Figure II, we plot both sets of iceberg costs. The smoothed iceberg costs indicate that there are indeed significant non-convexities with respect to effective distance. For example, we find that there is an initial sharp increase between deciles 1 and 2. Then, there is a subsequent flattening out starting in the third decile. Lastly, we see another large increase in deciles 9 and 10.

Figure II

Smoothed Iceberg Costs using ASI-NSS

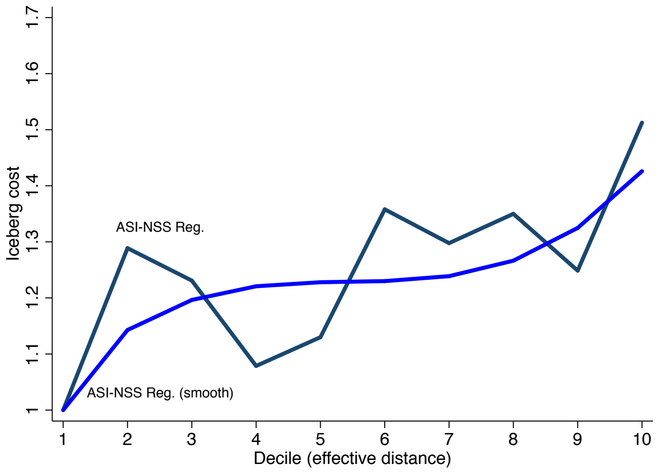


Figure II shows the transportation costs implied by the estimated coefficients of column (1) of Table I - "ASI-NSS Reg." and a monotonic cubic function that best fits the estimated coefficients - "ASI-NSS Reg. (smooth)". This is $g(x_d^o) = 0.9 + 0.176x_d^o - 0.0317(x_d^o)^2 + 0.002(x_d^o)^3$, where x_d^o is a discrete variable that indicates the decile of effective distance.

Panel A of Figure III shows a map of the transportation costs from the district of New Delhi (located in the National Capital Territory of Delhi). The legend on the map shows transportation costs divided into quartiles. The figure also shows that only a small portion of the GQ had been upgraded by this point (depicted in red). The first thing to notice is the concentric circles around New Delhi. This means that the further the destination, the higher the transportation costs. These circles also show that straight-line distances are highly correlated with the shortest path on the highway system. The reason is that the highway system is dense, as can be seen in Figure I. Next, we look at transportation costs in the year 2006 (Panel B of Figure III), after significant portions of the upgrade of the GQ had been completed. The color categories for the map have not changed compared to Panel A, so that the colors are comparable across maps. The lighter colors reflect a general decrease in transportation costs.

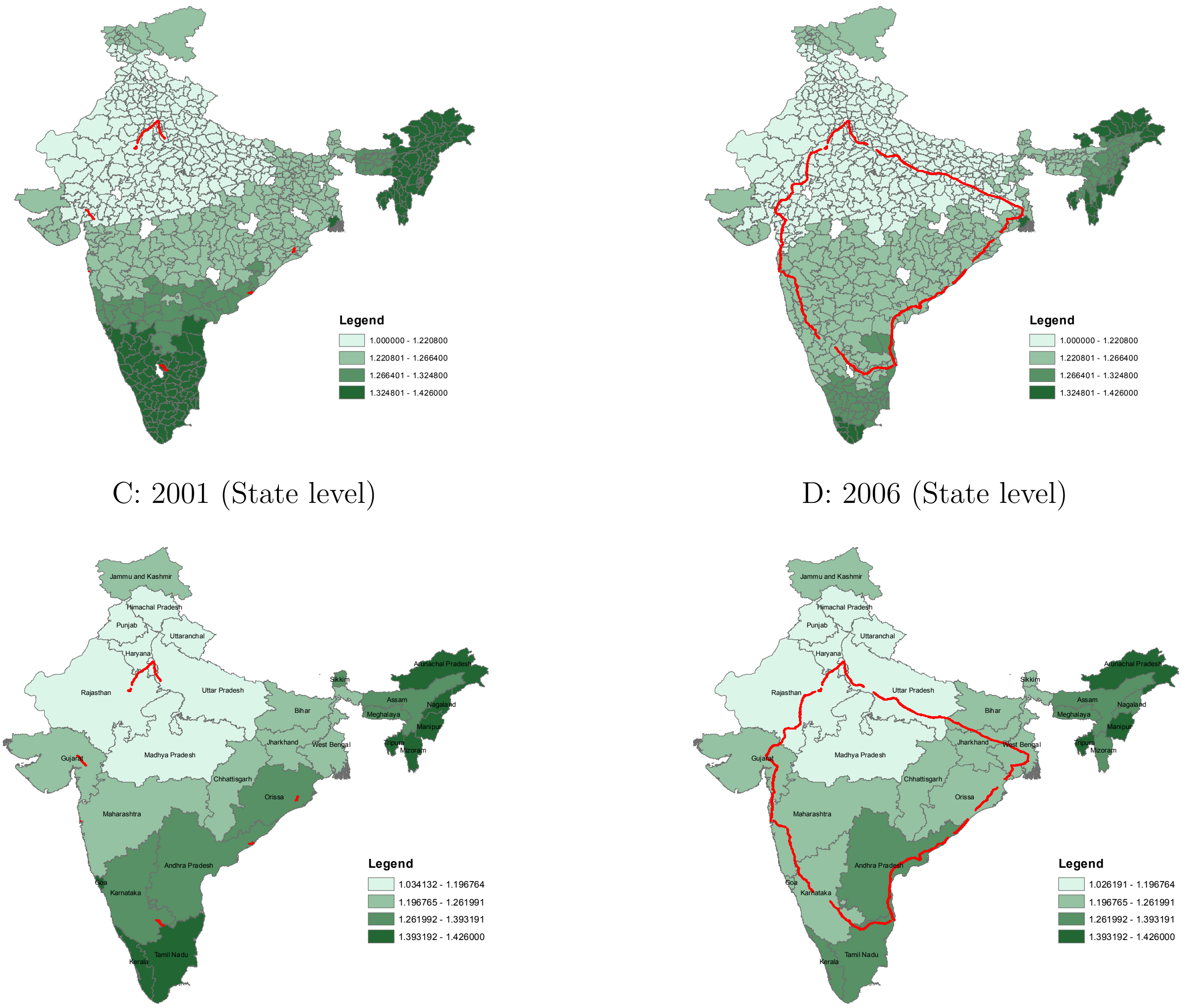
Geographic location of monopolists When estimating equation (8), it is important to consider whether there are unobserved factors that could be correlated with both distance and reported price. For example, destination markets that are far away could have characteristics that generate

Figure III

Estimated Transportation Costs from Delhi

A: 2001 (District level) B: 2006 (District level)

C: 2001 (State level) D: 2006 (State level)



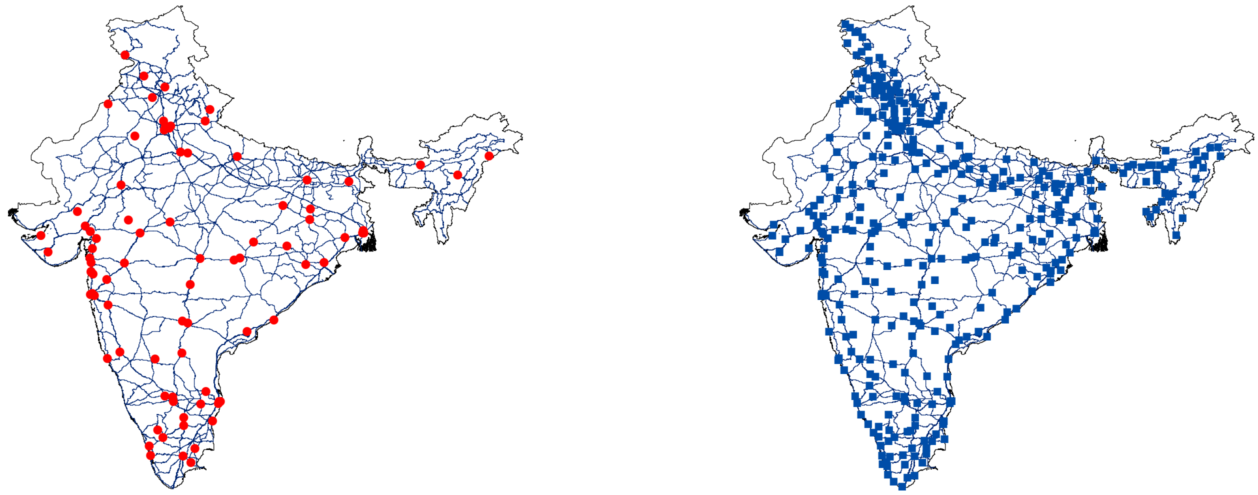
Panel A of Figure III shows the estimated transportation costs from Delhi at the district level for 2001; Panel B of Figure III shows the estimated transportation costs from New Delhi at the district level for 2006; Panel C of Figure III shows the estimated transportation costs from Delhi at the state level for 2001; Panel D of Figure III shows the estimated transportation costs from Delhi at the state level for 2006. The transportation costs have been estimated as explained in section 6.1.

higher/lower prices. For our specification, the ideal would be to have monopolist producers located in various parts of the country. Thus, the likelihood that there are unobservable characteristics systematically correlated with distance would be reduced. In Figure IV, we show a map of the location of all monopolist producers (Panel A) and the plants that utilize the products produced by the monopolists (Panel B). Reassuringly, we find that they are highly spread out geographically. As an additional check, we estimate a specification in which we include per capita income and average compensation per employee as additional controls, none of which enter significantly into the estimation.

FIGURE IV

ROAD NETWORK IN INDIA AND THE GQ

A: Monopolist plants B: Plants that utilize monopolist products



Panel A of Figure IV shows the location of monopolist plants. Panel B shows the location of plants that report using a monopolist product.

Imported intermediate inputs A key consideration when estimating equation (8) is that we have identified plants that are monopolists. One potential source of competition is that of foreign plants. Thus, we check the robustness of our pricing regressions by excluding goods in districts where monopolist producers may face foreign competition. To do so, we use the fact that plants report domestic and foreign intermediate inputs separately. We exclude any observations in districts where imports account for more than 5% of total input usage. This corresponds to 11 percent of the good-district observations. Excluding these observations yields a very similar profile in the association between prices and effective distance. See the Appendix for more details.

Comparison with direct measures of transportation prices We have also assembled an additional data set on transportation costs in India. The data contains prices charged by GIR Logistics, a large transportation logistics firm in India.[12] In particular, we have collected and digitized pricing quotes for transporting a shipping container of size 20 ft x 8 ft x 8.5 ft via truck for approximately 900,000 origin-destination city pairs in August 2014.

In order to compare the transportation costs implied by our estimates of equation (8) with those charged by GIR Logistics we proceed as follows. First, we construct prices at the district

[12] For details, see http://www.girlogistics.in/road-transportation.htm

level by calculating the simple average across cities. We then select the same pairs of districts as those used in column (1) in order to make the two sets of transportation costs comparable. After that initial preparation of the data, we estimate the following regression:

$$\log\, p^o_{d,GIR} = \sum_{\ell=1}^{10} \beta_{\ell,GIR}\mathbb{I}\{\text{Effective Distance}^o_{d,2006} \in \text{decile } \ell\} + \sum_o \delta_{o,GIR} + \epsilon^o_{d,GIR}, \tag{19}$$

where $p^o_{d,GIR}$ is the price charged by GIR Logistics to transport a container, $\delta_{o,GIR}$ is a set of origin fixed effects, and $\epsilon^o_{d,GIR}$ is an error term. The results of this regression can be found in column (3) of Table I.

Note that the estimates of $\beta_{\ell,GIR}$ cannot be compared to β_ℓ in columns (1) and (2). The coefficient $\beta_{\ell,GIR}$ measures changes in transportation prices, whereas β_ℓ measures changes in product prices at the destination. Thus, we convert the transportation costs estimated with the GIR Logistics data into iceberg costs. To do so, we use the following formula:

$$\tau^o_{d,GIR} = \frac{V + \hat{p}^o_{d,GIR}}{V} \tag{20}$$

where $\tau^o_{d,GIR}$ is the implied iceberg cost using the GIR Logistics data, V is the value of shipments, and $\hat{p}^o_{d,GIR}$ is the transportation cost estimate from equation (19). It is important to note that V pins down the level of iceberg costs. For example, as the value of shipments increases, the level of implied iceberg costs declines.

We now want to find V in a way that makes it comparable to the ASI-NSS data. First, we pick a V and find $\tau^o_{d,GIR}$ for all deciles using equation (20). We then divide the iceberg costs of all deciles by the first one in order to normalize transportation costs. As mentioned previously, in our analysis we have normalized the first decile of iceberg costs to one. We then find V such that the average normalized iceberg costs across deciles is equal to that of the ASI-NSS data.

The results of this exercise can be found in Figure V. In both cases, transportation costs increase more than linearly starting in deciles 7 and 8.

Predicted vs actual transportation cost estimates in 2001 We now investigate the ability of our empirical specification estimated using data from 2006 to predict transportation costs in 2001. We estimate equation (17) using log effective distance as the independent variable with data from the year 2006. The results of this estimation can be found in column (4) of Table I. We then regress the observed prices in 2001 on the predicted prices from the previous regression. The results can be found in column (5). We find that the coefficient on the predicted price is 0.80, indicating that the empirical model has strong predictive powers.

6.2 Estimating the Across-Sector Elasticity of Substitution (θ)

The next step consists in estimating the elasticity of substitution across sectors. The identification strategy is to compare the differences in the transportation costs of the goods produced by

Figure V

ASI-NSS vs. GIR Logistics

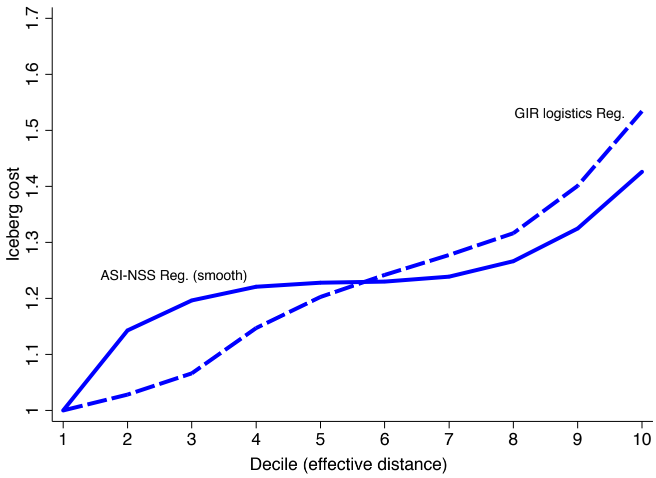


Figure V "ASI-NSS Reg. (smooth)" coefficients with the ones estimated in column (3) of Table I - "GIR logistics Reg."

monopolists across destinations with the trade flows across these destinations. This strategy is similar to that used by Donaldson (Forthcoming).

Formally, we derive a gravity equation implied by the model for the trade flows of monopolists. Combining equations (4) and (15), we derive the following condition for the trade flow values:

$$\begin{aligned}\log c_d^o(j,k)p_d^o(j,k) =& (1-\theta)\log W_o + (\theta-1)\log a_o(j,k) + \log P_d^\theta Y_d \\ &+ (1-\theta)\log \tau_d^o + (1-\theta)\log\frac{\theta-1}{\theta}.\end{aligned} \tag{21}$$

The model predicts that higher transportation costs reduce trade flows, and the strength of this relationship depends on the value of θ. The intuition behind this identification strategy is that if small differences in transportation costs across destinations are associated with big differences in trade flows, then the value of θ must be high (and vice versa). It is also important to note that this straightforward relationship only holds when firms are monopolists.

We estimate equation (21) as follows:

$$\log \mathrm{Sales}_{d,t}^o(j,k) = \zeta \log \hat{\tau}_{d,t}^o + \sum_{o,t}\delta_{o,t} + \sum_{j,t}\alpha_{j,t} + \sum_{d,t}\lambda_{d,t} + \epsilon_{d,t}^o(j,k) \tag{22}$$

where $\mathrm{Sales}_{d,t}^o(j,k)$ is the value of sales of product j in year t consumed in district d and produced by a monopolist located in district o, $\hat{\tau}_{d,t}^o$ is the predicted iceberg transportation cost found in Section 6.1, $\delta_{o,t}$ are district of origin dummies that vary by year, $\alpha_{j,t}$ are product dummies that vary by year, $\lambda_{d,t}$ are district of destination dummies that vary by year, and $\epsilon_{d,t}^o(j,k)$ is the error term. The origin dummies control for local wages. The product dummies control for firm productivity.

The destination dummies control for market size and aggregate prices at the destination. As in Section 6.1, this specification uses cross-sectional variation to identify parameters. Furthermore, we estimate it at the district level in order to fully exploit the variation in the data.

Table II presents the results of estimating equation (22). Columns (1) to (2) show the coefficient associated with the predicted transport costs constructed from the coefficients of columns (1) to (2) of Table I, respectively. We find that higher transportation costs are associated with lower trade flows at statistically significant levels. Our estimates range from 0.83 to 0.99. Thus, a 10 percent increase in transport costs is associated to a 8-10 percent decrease in trade flows. In column (3) we assume that all variations in transport infrastructure are translated into prices, which means introducing effective distance directly into equation (22). We find that a 10 percent increase in transport costs is associated to a 2 percent decrease in trade flows. Given these estimates, we take a value of 1.99 for θ as a benchmark, which is the most conservative one in terms of its implications for the size of allocative efficiency gains.

6.3 Estimating the Within-Sector Elasticity of Substitution (γ)

We now estimate the within-sector elasticity of substitution. To do so, we derive the following condition from the model between a firm's labor share and its sectoral share for a given destination:

$$\frac{W_o l_d^o(j,k)}{\tilde{p}_d^o(j,k) c_d^o(j,k)} = 1 - \frac{1}{\gamma} - \left(\frac{1}{\theta} - \frac{1}{\gamma}\right) \omega_d^o(j,k) \tag{23}$$

where $\tilde{p}_d^o(j,k)$ is the factory gate price of the good. This condition implies that firms with a higher sectoral share at a destination have a lower labor share. The reason is that firms with higher sectoral shares charge higher markups, which result in lower labor shares. See the Appendix for more details.

In the data, we do not observe the market shares of firms by destination. However, a similar condition can be derived for goods that are only produced in one state. The reason is that in these sectors, the market shares of firms are constant across destinations. Empirically, we find that approximately 12% of goods are produced only in one state. Using data from these plants producing these products, we estimate equation (23) as follows:

$$LS_{o,t}(j,k) = \beta \omega_t^o(j,k) + \sum_{o,t} \delta_{o,t} + \sum_{j,t} \alpha_{j,t} + \epsilon_t^o(j,k) \tag{24}$$

where $LS_{o,t}(j,k)$ and $\omega_t^o(j,k)$ are the labor and sectoral shares respectively in state o, $\delta_{o,t}$ are district of origin dummies that vary by year, $\alpha_{j,t}$ are product dummies that vary by year, and $\epsilon_t^o(j,k)$ is the error term. Note that γ will be given by $\hat{\theta}/\left(1+\hat{\beta}\hat{\theta}\right)$. In contrast to the estimation of transportation costs and the elasticity of substitution across sectors, we estimate equation (24) at the state level in order to maximize the number of products produced in one location.

We present the results in Table III. Columns (1) and (3) show the results when we include only labor remuneration on the left-hand side of the equation, whereas in columns (2) and (4) we also include capital remuneration on the left-hand side of the equation. This second type of specification controls for across-firm variations in capital intensity. In columns (1) and (2), we include the pool of observations for the years 2001 and 2006, and control for time-variant state and product fixed effects. In columns (3) and (4), we only include observations for the year 2006.

We find a strong correlation between the labor shares and sectoral shares of firms in the data. The point estimates are similar in magnitude across the four specifications, ranging from -0.34 to -0.49. The implied values for γ in columns 1-4 are 6.17, 25.30, 10.67, and 107.38, respectively. Note that small changes in β can lead to large changes in γ, given the functional form that relates these two variables. We choose 10.67 as a benchmark, since it is the closest to the value used by Atkeson and Burstein (2008) and Edmond, Midrigan, and Xu (2015).

6.4 State-Level Transportation Costs

It is necessary to aggregate the district-to-district transportation costs to the state level since the model that we simulate is based on interstate trade. We do so in two steps. In the first step, we choose a district and find the average transportation cost to a destination state. This process yields district-to-state transportation costs. In the second step, we obtain the state-to-state transportation costs by finding the average over districts in the origin state. All averages are weighted by population.

In Figure III, we map transportation costs from the National Capital Territory of Delhi. Panel C shows transportation costs in 2001 where the legend colors reflect the quartile of each destination state. Panel D shows transportation costs in 2006. We hold the legend categories fixed so that the maps are comparable. The lighter colors in 2006 reflect declines in transportation costs from the GQ. Figure VI shows the percentage decline in these transportation costs. States in the quartile with the largest declines tend to be far from Delhi and in a location that can utilize the GQ for transportation between these locations. The states in the top quartile benefit from a decline of 4.9 to 13.1 percent in transportation costs. For the bottom quartile, this figure ranges from 0 to 0.3 percent.

6.5 Calibrating the Remaining Parameters

Labor endowment There are large differences in the economic size of states. For example, Maharashtra (the largest state) accounts for 23 percent of manufacturing value-added, while Arunachal Pradesh (the smallest state) accounts for only 0.01 percent. In order to capture these differences in the model, we first normalize the labor endowment of Aruchanal Pradesh to 1. We then set the labor endowments of each state, L_n, so that the model matches the ratio of manufacturing value-added observed in the data across states.

Figure VI

Percentage change in transportation costs from Delhi

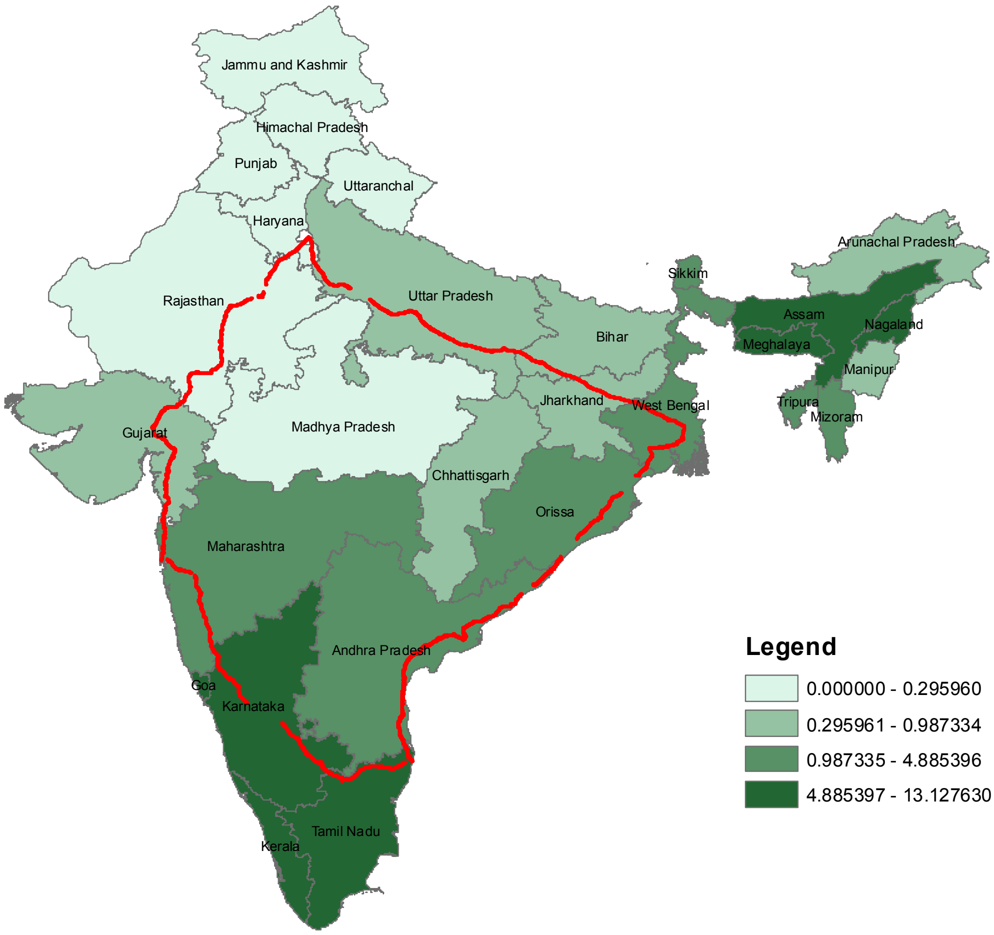


Figure VI shows the percentage change in transportation costs from Delhi due to the construction of the GQ at the state level.

Number of producers by state and sector The number of firms that operate in each state is important to determine the nature of gains from lower transportation costs. To illustrate this idea, consider a two-state example in which these two states go from autarky to trading with each other. If firms in these two states produce in entirely different sectors, the effects from trade will be purely Ricardian since markups will remain unchanged. However, if the two states produce goods in the same sectors, then there may be effects on allocative efficiency.

We set the number of firms in sector j of state n, K_{nj}, to match the number of plants observed in the data. Since there are no fixed costs in the model, firms always choose to operate. Thus, there is no entry and exit of firms after changes in transportation costs. Abstracting from firm entry and exit in these kinds of models does not affect the final results quantitatively, as discussed by Atkeson and Burstein (2008) and Edmond, Midrigan, and Xu (2015). The reason is that a reduction in transportation costs leads to the entry and exit of low-productivity firms, which do not significantly affect the markups that large firms charge. Furthermore, the data does not show significant changes in the number of firms operating in each state. For example, the auto-correlation of the number of producers per sector-state between 2001 and 2006 is 0.98. We did not see large changes in the number of active sectors either (average change of 3 percent), or in the total number of firms (average change of 2 percent) by state over this period.

Finally, in order to reduce the computational burden, we limit the number of firms operating in each sector of a state to 50. This means that we set the maximum number of producers per sector to 1,450 (29 x 50).

Productivity distribution We use a Pareto distribution for the productivity draws. The tail parameter, α, is calibrated in equilibrium to match the fact that the top 5% of firms in manufacturing value-added account for 89% of value-added.

Within-industry productivity across states The correlation of firm productivity draws across states is important to determine the size of allocative efficiency gains. The reason is that if firms across states have a similar productivity, then there is a high degree of head-to-head competition. Thus, lowering transportation costs will have a larger impact on the distribution of markups.

We assume that firms across states have perfectly correlated productivity draws. To implement this, we first find the maximum number of plants present in any state for each industry. We make this number of draws from a Pareto distribution. We then sort the productivities in descending order. If a state has one firm, we select the first productivity on the sorted list. If a state has ten firms, we select the first ten productivities on the sorted list. This setup ensures that the firms with the highest productivity face head-to-head competition. Note that this does not imply that the sectors are symmetric across states. The reason is that states have a different number of operating firms. Furthermore, states have different wages and transportation costs, which affect their marginal cost.

It is important to determine whether the model generates reasonable levels of head-to-head competition given the assumption of perfectly correlated productivity draws. We create a "similarity" index that measures the similarity in size among the largest firms across states. We focus on the largest firms since they are the ones that drive most of the dispersion in markups as we will show in Section 7. To calculate the index, for each sector and state we identify the firm with the largest value-added. Then, we regress the log of the value-added of these firms on sector dummies. The R squared of that regression, which we use as our index, indicates the extent to which large firms in each sector are of similar size. For example, an R squared of one indicates that the largest firms across states are exactly the same size. We find an index of 0.45 in the data and 0.49 in the model. A similar picture emerges when we use employment as a measure of firm size. In that case, we find a similarity index of 0.46 in the data and 0.43 in the model. This exercise indicates that the model generates levels of head-to-head competition that are in line with the data.

Furthermore, we can check whether the trade elasticity implied by the model is consistent with other estimates in the literature. More head-to-head competition implies a larger trade elasticity. The reason is that when producers in different locations have similar levels of productivity, small changes in trade costs have larger effects on trade flows between states. We calculate the trade elasticity implied by our model by estimating the following regression

$$S_d^o = \sigma \log \tau_d^o + \sum_o \delta_o + \sum_d \lambda_d + \epsilon_d^o, \tag{25}$$

where S_d^o is the bilateral trade from state o to d, σ is the trade elasticity, δ_o is a set of state-origin fixed effects, and λ_d is a set of state-destination fixed effects. The trade elasticity implied by our model is 4.74, which is similar to recent estimates found in the literature. For example, Head and Mayer (2014) carry out a meta-analysis of empirical estimates of trade elasticity and find a median estimate of 5.03.

We will show in the sensitivity analysis that the case of non-correlated draws cannot match the data in the two dimensions listed above.

6.6 Discussion of Markups

Distribution of markups faced by consumers Table IV summarizes the distribution of markups charged to consumers in each state. We find an average markup of 1.11 in all destinations, which is very close to the lowest possible markup of $\gamma/(\gamma-1)$. Furthermore, our results indicate that the bulk of firms do not have enough market shares to charge significantly higher markups. For example, even in the 95th percentile of the markup distribution, we do not see a large increase in the minimum markup. Markups become significantly larger in the 99th percentile, which ranges from 1.25 to 1.29 across states. There are two ingredients in the model that can explain this markup distribution. First, we use a Pareto distribution for productivity draws. Second, the model implies a convex relationship between sectoral shares and markups. Thus, the few firms with large market shares also have the high markups.

Markups compared to empirical studies Empirical studies in industrial organization find that the bulk of firms have modest markups and that a minority of firms have markups that are significantly higher. For example, De Loecker, Goldberg, Khandelwal, and Pavcnik (2016) estimate the median markup by sector. They use the Prowess data set which covers medium to large Indian firms. The authors find a median markup of 1.18 across sectors. They also find a mean markup of 2.24 across sectors indicating a skewness in the right tail of the markup distribution. De Loecker and Warzynski (2012) find similar results using data from Slovenian manufacturing. They find a median markup of 1.17-1.28 across sectors depending on the specification. Furthermore, they find a standard deviation of 0.50 across all specifications, also implying a skewness in the right tail of the distribution.

Our calibrated model matches the fact that most firms have small markups and that a few firms have very large markups. On the other hand, our model does not quantitatively capture the high end of the markup distribution found in these empirical studies. For example, the highest possible markup that firms can charge in the model is $\theta/(\theta-1)$ or 2.01. In Section 9, we perform a sensitivity analysis by lowering θ to allow for higher markups.

Where are the high markup firms located? An important dimension of the markup distribution is related to the location of firms. In the model, large states have lower wages. For example, the 5 largest states have an average wage of 0.44, where we have normalized the wage of the smallest state to 1. States ranked 10-20 and 25-29 in size have an average wage of 0.50 and 0.86, respectively. As a result, firms located in large states tend to have a cost advantage and thus charge higher markups. Another implication is that large states have lower levels of allocative efficiency. The reason is that, again, the local firms can charge relatively high markups. The allocative efficiency index from equation (14) indicates a loss in welfare of 2% to 4% across states due to a dispersion in markups, with the low states having the highest losses due to poor allocative efficiency.

Figure VII shows the location of the firms whose markups on goods purchased in Arunachal Pradesh (the smallest state) and Maharashtra (the largest state) are in the top 1%. Note that the state of origin is ranked from largest to smallest. We see that, in both cases, firms with the highest markups are primarily located in Maharashtra and other big states where wages tend to be lower. For instance, around 50% of the firms charging the markups in the top 1% in the market of Maharashtra are local firms.

7 Quantifying the Impact of the GQ

In this section, we quantify the aggregate and state-level effects of the construction of the GQ. To this end, we compare the outcomes from our calibrated model in 2006 with the outcomes when we remove the GQ. To remove the GQ, we use the estimates from Section 6.1 to determine the changes in transportation costs. For illustrative purposes, we present all the results as changes from before to after the construction of the GQ (2001 to 2006).

In order to quantify the effects of the GQ, we begin with our baseline calibration described in Table V. We then feed the iceberg transportation costs in 2001 into the calibrated model. To estimate these transportation costs, we first find the new effective distance between districts in 2001 by re-computing the shortest path between them taking into account the road network at the time. We use these effective distances, along with estimates from equation (17), to find the new iceberg transportation costs. Finally, we re-aggregate the district-to-district transportation costs

Figure VII

Spatial distribution of the top 1% firms in terms of markups (Model)

(A) Firms selling to Arunachal Pradesh (B) Firms selling to Maharashtra

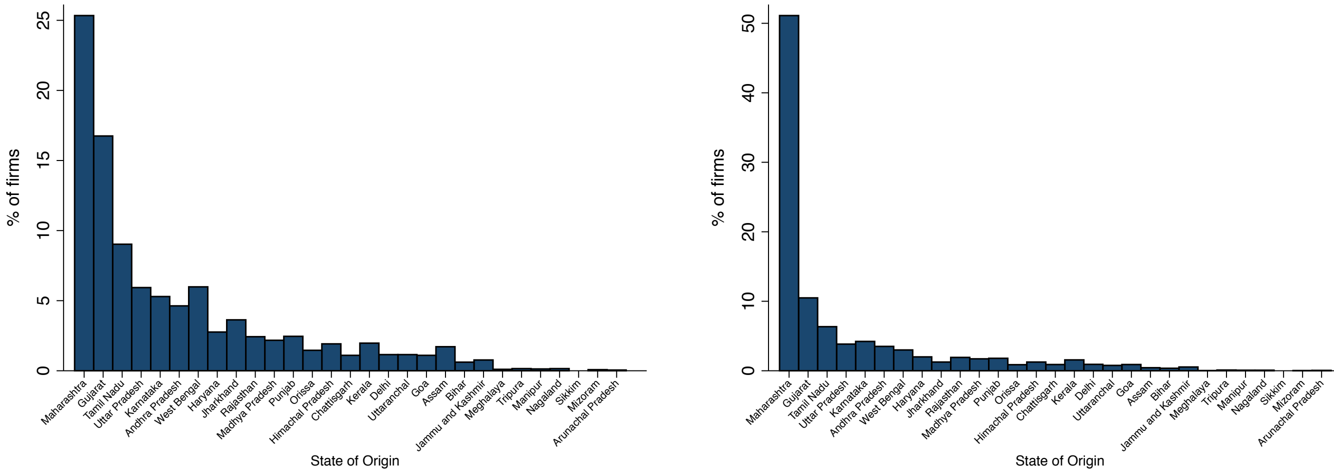


Figure (VII) shows the distribution of states in which the top 1% of firms in terms of markups operate. Panel A refers to the markups charged on goods sold in Arunachal Pradesh. Panel B refers to the markups charged on goods sold in Maharashtra.

to state-to-state transportation costs as described in Section 6.4.

7.1 Welfare Changes

First, we consider the aggregate change in real income resulting from the GQ. We find that real income increased by 2.71% as shown in Table VI.[13] In our quantitative exercise, we only consider the manufacturing sector which had a value-added of $152.8 billion in 2006 (around 16% of Indian GDP). Thus, the static benefit from the GQ project is $4.1 billion. Note that these static benefits accrue to the economy *each year*. At the same time, estimates indicate that the government spent approximately $5.6 billion on this project. Thus, we find that the benefits over a two-year period exceed the initial construction costs.

The welfare effects of the GQ are very heterogeneous across states. Table VI lists the welfare effects across states in descending order of size. Overall, large states have gained more from the reduction in transportation costs. Small states have seen modest gains and, in some cases, have even lost. This is driven by the fact that, due to its placement, the GQ has lowered transportation costs primarily for large states. Panel A of Figure VIII shows a map of the welfare effects across states. The map shows that most of the states that have lost are located in the Northeast, which are the states located farthest from the GQ. The states in the Northeast that have experienced losses include: Arunachal Pradesh, Manipur, Mizoram, Nagaland, and Tripura. Finally, the state

[13]The model generates changes in real income at the state level. To aggregate these changes to the national level, we compute the percentage change in real income for all states weighted by their size.

of Chattisgarh has experienced a loss that is very close to zero.

7.2 Decomposition of Welfare Changes

We use the Holmes, Hsu, and Lee (2014) decomposition to break down the effect of the GQ on changes in real income into Ricardian, markups terms of trade, and allocative efficiency effects. Table VI shows these components at the aggregate and state level. Table VII shows the change in these three components and Figure VIII shows the geographical distribution across India.

7.2.1 Allocative efficiency

We find that, for India as a whole, the allocative efficiency component accounts for 8% of the aggregate gains (0.21% of the 2.71% total gains). Lower transportation costs have generally led to welfare-enhancing changes in markups since the allocative efficiency effects are positive in all but four states. This quantitative result is informative since the theory is ambiguous as to whether declines in transportation costs lead to gains in allocative efficiency. We find that the largest states see the strongest allocative efficiency gains. For example, allocative efficiency gains in Maharashtra account for 19% of the overall gains (0.33% of 1.77%). The reason is that large states have low levels of allocative efficiency since local firms tend to have the highest markups. We also find a strong concentration of gains in allocative efficiency among the largest states. For example, the average gain in allocative efficiency is 0.29% for the five largest states. This number declines significantly after the largest five states.

We carry out two sets of simulations to disentangle whether improvements in allocative efficiency come from: 1) changing transportation costs, which are a direct effect of the GQ, or 2) changing wages, which are the result of general equilibrium effects. To do so, we report two sets of simulations. In the first set of simulations, we change transportation costs for only one bilateral pair of states, while holding all other transportation costs and wages fixed. Changes in allocative efficiency in these simulations are reported in Table VIII. The column indicates the state for which we report the percentage change in allocative efficiency and the row indicates the state for which we change the transportation costs. Consider the case of Chattisgarh, which is the median state in terms of size. The column for that state indicates positive effects from reducing transportation costs to all but four states. Thus, we find that transportation costs tend to have a positive effect. The reason is that local firms, which have relatively high markups, must lower their markups when transportation costs decline. One notable exception is the direct effect of Chattisgarh reducing transportation costs with Maharashtra. In this case, this direct effect of lowering transportation costs is negative. The reason is that firms located in Maharashtra have relatively high markups since wages in that state are low enough to compensate for transportation costs.

In the second set of simulations, we change wages in one state and hold all other wages and transportation costs fixed. The changes in allocative efficiency are reported in Table IX. The row indicates the state for which we change wages and the column indicates the state for which we report the percentage change in allocative efficiency. It is important to note that wages rose for states close to the GQ. The reason is that the exports of these states became more competitive relative to other states. Thus, wages must rise in order to satisfy the balanced trade condition. As before, consider the case of Chattisgarh. Allocative efficiency rises if we let wages only in Chattisgarh rise since local firms, which have relatively high markups, lower their markups. If we inspect the rest of that column, we find that rising wages in other states negatively affects allocative efficiency in Chattisgarh. One exception is the increase in wages in Maharashtra, which improves allocative efficiency in Chattisgarh.

The last row of Table VIII shows the change in allocative efficiency in each state if we change all transportation costs and hold wages fixed. The last row of Table IX indicates the change in allocative efficiency if we change all wages and hold transportation costs fixed. Let's come back to the case of Chattisgarh. The change in allocative efficiency which comes from the direct effect of changes in transportation costs is positive (0.050%). In contrast, the indirect effect implies a negative change (-0.029%). This means that for the particular case of Chattisgarh, ignoring the general equilibrium effects through wages would lead to overestimating the size of allocative efficiency gains.

We find that the relative importance of the direct effect vs indirect effect in explaining allocative efficiency gains varies across states. In particular, we find that the direct effect accounts for the bulk of changes in allocative efficiency in big states. Let's consider the case of Maharashtra. In this case, the gains from changes in relative wages are -0.067%, whose absolute value represents around 17% of the gains that come from the direct effect. For the case of small states, however, the importance of the indirect effect is bigger. Take for instance the case of Aruchanal Pradesh. In this state, we find that the gain from the indirect effect (0.015%) is larger than the gain from the direct effect.

7.2.2 Markups Terms of Trade

We find that welfare gains from changes in markups terms of trade are quantitatively important. There is a significant re-shuffling of income across states due to the markups terms of trade component. We use the term re-shuffling since this term is zero in the aggregate. For example in the case of Mizoram, welfare gains from the improvement in its markups terms of trade are 0.22%. This implies that, had this effect not been present, welfare losses would have been around 20% bigger in absolute value. This effect is due to the fact that the increase in wages for states close to the GQ forced firms located in those states to lower their markups in Mizoram. Thus, the aggregate markup on the goods imported by Mizoram declined.

7.2.3 Change in the distribution of markups

Figure IX shows the percentage change in markups that firms charge in Arunachal Pradesh and Maharashtra, the smallest and the biggest states. To construct the figure, we first find the markups of firms in the various of the markup distribution before the construction of the GQ. For each of these percentiles, we then find the average percentage change in markups after the construction of the GQ. As mentioned before, most firms charge markups that are very close to the lowest possible one. That is why we only see quantitatively relevant changes starting in the 90th percentile. We also find that the largest decline is in the 99th percentile of the distribution. This result is consistent with other works which find that, after trade reforms, declines in markups are more pronounced among firms with initially high markups. For instance, De Loecker, Goldberg, Khandelwal, and Pavcnik (2016) find that, for high-markup products (above the 90th percentile), the same reduction in tariffs results in an additional 4.40 percent decline in markups. Lastly, we find that the markups in the 90th and 95th percentiles of the distribution increase. On the other hand, in both cases, we find that this increase is smaller than the decline among the firms in the 99th percentile. The model also indicates that the largest declines in markups are in Maharashtra. This finding is consistent with the fact that Maharashtra experienced the largest gains in allocative efficiency.

7.2.4 Ricardian

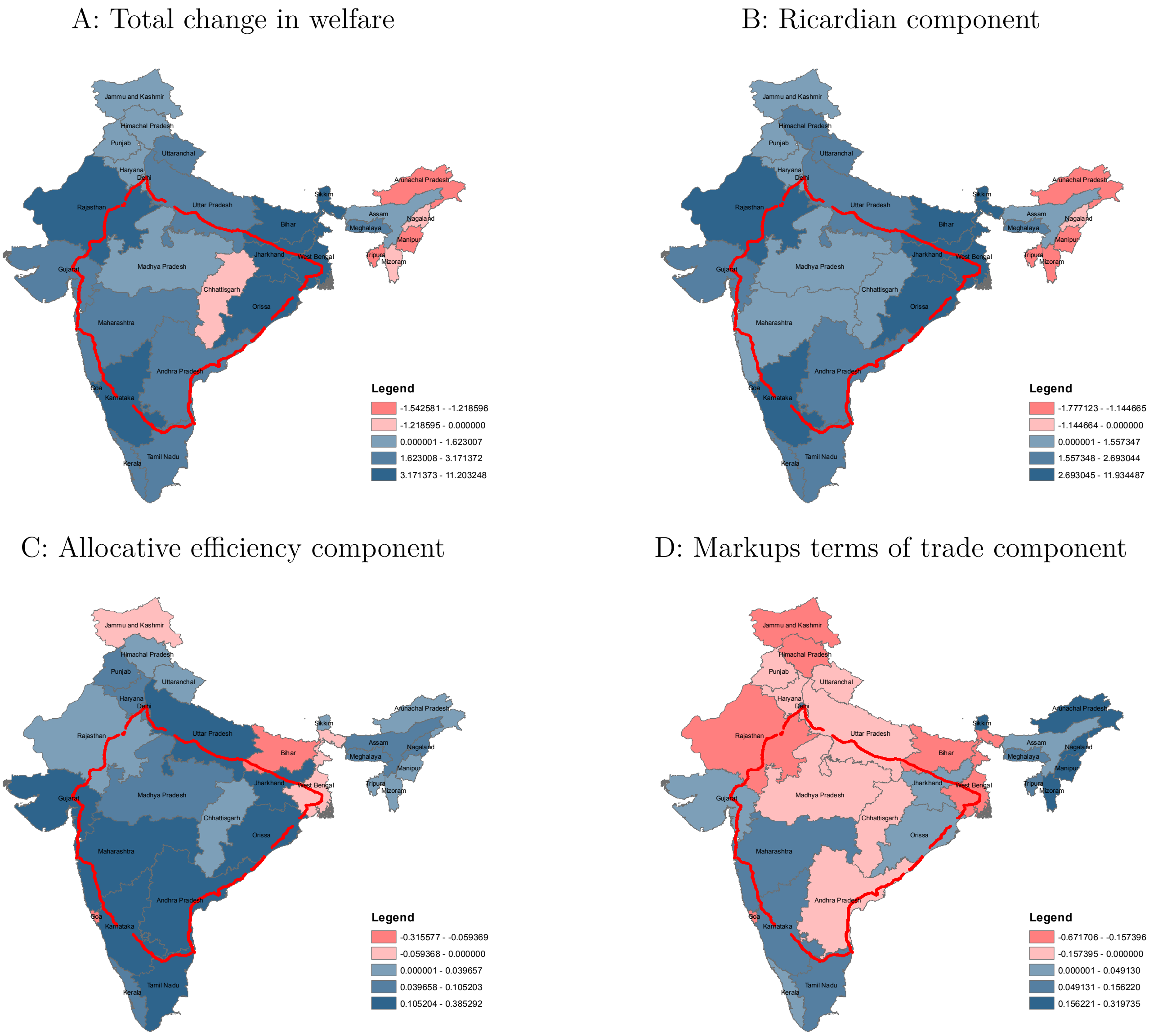
We now turn to the Ricardian component of welfare across states. This term is generally positive and large across states. It also explains the negative effects for the states in the Northeast. The fact that the Ricardian term is negative for Northeastern states comes from the fact that the price index under marginal cost pricing increases. The reason is that it becomes more expensive to purchase goods at marginal cost from states close to the GQ. The only two factors that affect a firm's marginal cost to serve a destination are the transportation costs that it faces and wages. First, we know that the GQ lowers transportation costs for some destinations and leaves those for others unchanged. Second, states close to the GQ trade more intensively with each other. The result is that they become more open and their wages increase. Thus, the increase in wages in the states close to the GQ outweighs the benefits of the GQ in terms of lower transportation costs.

We find that the negative Ricardian effect induced by changing trade patterns is mitigated by the markups terms of trade term. As mentioned above, the average state in the Northeast that lost had an average gain of 0.25% in markups terms of trade.

Predictions of the model about trade diversion and creation We now study the changes in state-to-state trade patterns induced by the construction of the GQ. The fact that reductions in transportation costs are not uniformly distributed across states leaves room for trade diversion and creation. To study this possibility, we compute bilateral trade flows all Indian state pairs implied by the model before and after the GQ. We define the total trade between state i and state

Figure VIII

Percentage change in real income after GQ with HHL components



Panel A of Figure VIII shows the percentage change in real income after the decrease in transportation costs due to the construction of the GQ; Panel B shows the Ricardian component of the change in welfare; Panel C shows the allocative efficiency component of the change in welfare; Panel D shows the markups terms of trade component of the change in welfare. The numbers represented in this map correspond to the ones presented in columns 2-6 of Table VI.

j as

$$\text{Total Trade}_{i,j} = \text{exports}_j^i + \text{exports}_i^j,$$

where exports_j^i and exports_i^j are the total exports from state i to state j and the total exports from state j to state i, respectively.

Table X shows summary statistics of the change in trade patterns for state pairs according to

Figure IX

distribution of the change in markups (Model)

(A) Firms selling to Arunachal Pradesh   (B) Firms selling to Maharashtra

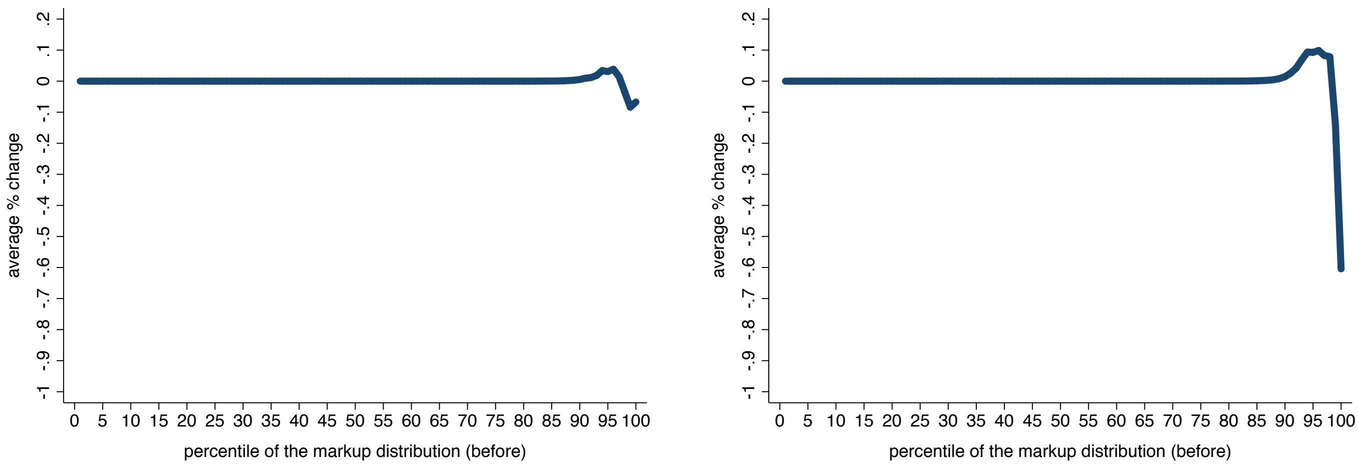


Figure IX shows the average percentage change in markups across firms across different percentiles of the distribution of markups before the construction of the GQ in the model. Panel A refers to the markups charged on goods sold in Arunachal Pradesh. Panel B refers to the markups charged on goods sold in Maharashtra.

their access to the GQ.[14] In the first row, we include state pairs in which both states are crossed by the GQ. In the second row, we include state pairs in which only one of the states is crossed by the GQ. In the third row, we include state pairs in which neither state is crossed by the GQ.

We find that on average trade increases considerably more between state pairs in which both states are either crossed by the GQ or not crossed by the GQ. We also find evidence of trade diversion. For instance, for the median trade relationship, trade flows between state pairs crossed by the GQ increase by 4.36%. For state pairs in which neither state is crossed by the GQ, the median increase in trade is 0.41%. On the other hand, the median change in trade between state pairs in which only one of the states is crossed by the GQ is -1.12%.

7.3 Comparison with Monopolistic Competition

In this section, we compare the results from our framework with those of a standard model of monopolistic competition. As a first step, we take the changes in trade flows from our baseline case as given. We then use the Arkolakis, Costinot, and Rodriguez-Clare (2012) (ACR) framework to study the welfare implications in a model of monopolistic competition given these trade flows. In particular, we use the fact that the percentage change in real income under monopolistic competition is $(1/\epsilon)(\lambda/\lambda')$, where ϵ is the trade elasticity, and λ and λ' are the shares of spending on the local good before and after the GQ. The results are reported in Table XI. We find that

[14] This analysis is similar in spirit to papers that study trade diversion such as Krueger (1999) and Bayoumi and Eichengreen (1998).

the gains in the aggregate using the ACR formula are 2.28% (vs 2.71% in the baseline). These results suggest that, in situations in which firms compete oligopolistically, using the observed trade flows along with the ACR formula will underestimate the gains from lowering trade costs. In this particular case, the gains would be 19% higher.

As a second step, we re-do the entire exercise with a model of monopolistic competition. To do so, we set $\theta = \gamma$, which becomes the elasticity of substitution in the model of monopolistic competition. Note that, in this case, all firms have the same markup. Furthermore, we set the elasticity of substitution to 5.74, so that the model generates the trade elasticity of 4.74 as in the benchmark (see section 6.5). In this exercise, we re-calibrate the labor endowments and tail parameter of the Pareto distribution to match the same statistics as in the benchmark case.

We find that the model of monopolistic competition generates aggregate gains of 2.83% (vs 2.73% for our benchmark case). Thus, the aggregate gains are 4% higher in the model of monopolistic competition. However, the distribution of gains across states is very different. Figure X shows a scatterplot of the percentage difference in the gains from the benchmark model and that of monopolistic competition. We plot this against the log of the ratio of the real GDP of the state and that of the smallest state. We find that the smallest and largest states gain more in the benchmark case. The states with the higher gains are precisely those which see larger effects from allocative efficiency and markups terms of trade in the baseline case.

Figure X

Baseline model vs Monopolistic competition model

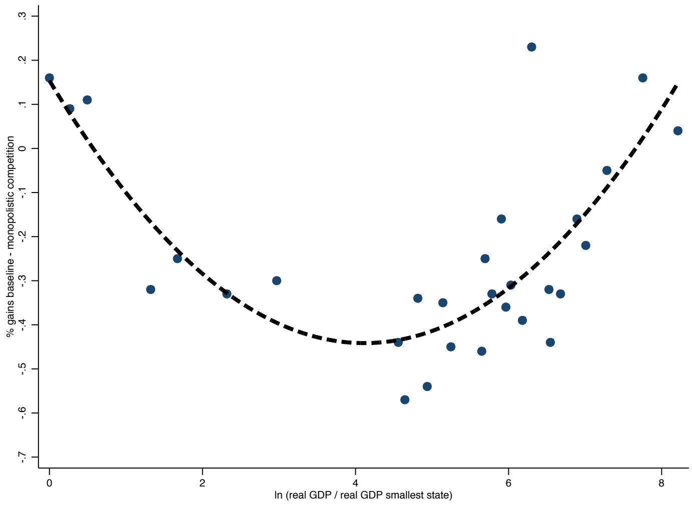


Figure X shows the gains in real income for each state predicted by our baseline model minus those of a re-calibrated model of monopolistic competition. We plot this difference against the relative size of the state.

8 Reduced-Form Evidence on the Effects of the GQ

In this section, we compare the predictions of the model and data regarding two outcomes, prices and allocative efficiency. To do so, we use reduced-form approaches on the data and compare the results with output from the model.

8.1 Impact of the GQ on Prices

One of the challenges to identify the causal impact of transportation infrastructure is its non-random placement. For example, transportation infrastructure may be placed in areas with characteristics that are correlated to economic outcomes of interest. For example, infrastructure may be placed in areas that are expected to have high future growth. An identification strategy used in the latest empirical literature has exploited the fact that infrastructure projects often aim to connect historical cities or large economic centers.[15] In our particular case, the stated goal of the GQ was to connect the major urban centers (Delhi, Kolkata, Chennai, and Mumbai). Thus, we estimate a difference-in-differences specification in which we compare economic outcomes for districts close to the GQ with those that are far away. We exclude the major urban centers or nodal districts since these areas were explicitly targeted by policymakers.

We run the following difference-in-differences regression:

$$\Delta \log P_{jd} = \beta_1 \Delta \mathrm{GQ}_d + \sum_j \alpha_j + \epsilon_{jd}, \tag{26}$$

where $\Delta \log P_{jd}$ is the change in log price of input j in district d between 2001 and 2006, GQ_d is a dummy variable taking value 1 if district d has been "treated" by the GQ, α_j are product fixed effects, and ϵ_{jd} is an error term.[16] Thus, $\Delta \mathrm{GQ}_d$ takes value 1 if a district was within the specified distance of the GQ in 2006 but not in 2001. Distance is calculated as the shortest straight-line distance between the district and a treated portion of the GQ. We use categories of distance ranging from 25 to 300 kilometers. Standard errors are clustered at the district level in order to account for the possible serial correlation of price shocks within districts.

The estimates of equation (26) can be found in columns (1) and (2) of Table XII. The results in column (1) include nodal districts. Column (2), which is our baseline specification, excludes nodal districts. Each column shows the estimate of β_1 under different specifications of distance.

We find that prices declined significantly for areas close to the GQ relative to those farther away. For example, we find that for districts located within 25 km of the GQ, input prices declined by almost 60 percentage points more than in districts located farther away. Furthermore, we find

[15] See Banerjee, Duflo, and Qian (2012) for an early example of this empirical strategy. This strategy has also been applied to the GQ by Alder (2014), Datta (2012), and Ghani, Goswami, and Kerr (2016).

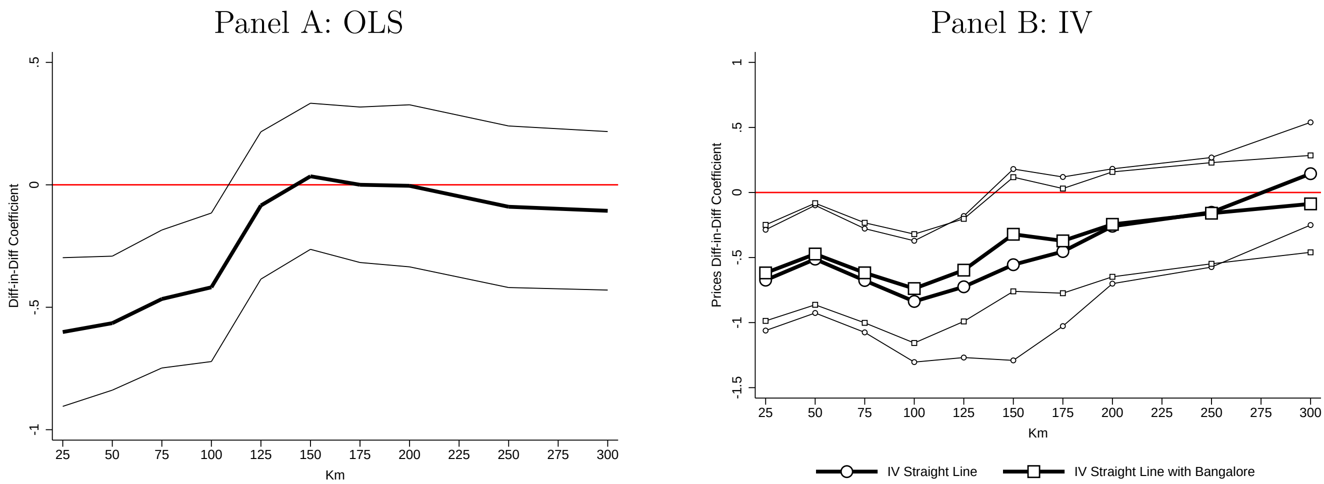
[16] We compute a weighted average of the prices paid by plants consuming the input in the district, excluding products with evidence of unit misreporting. We have data for 920 inputs consumed in 325 districts. See the Appendix for more details.

that this effect dissipates as we increase the treatment distance. This trend towards zero can be seen in Panel A of Figure XI, in which we plot the coefficients in column (2) in steps of 25 km. Finally, we find that the exclusion of the nodal cities does not significantly change the estimates.

Figure XI

Prices and the Golden Quadrilateral

Differences-in-Differences



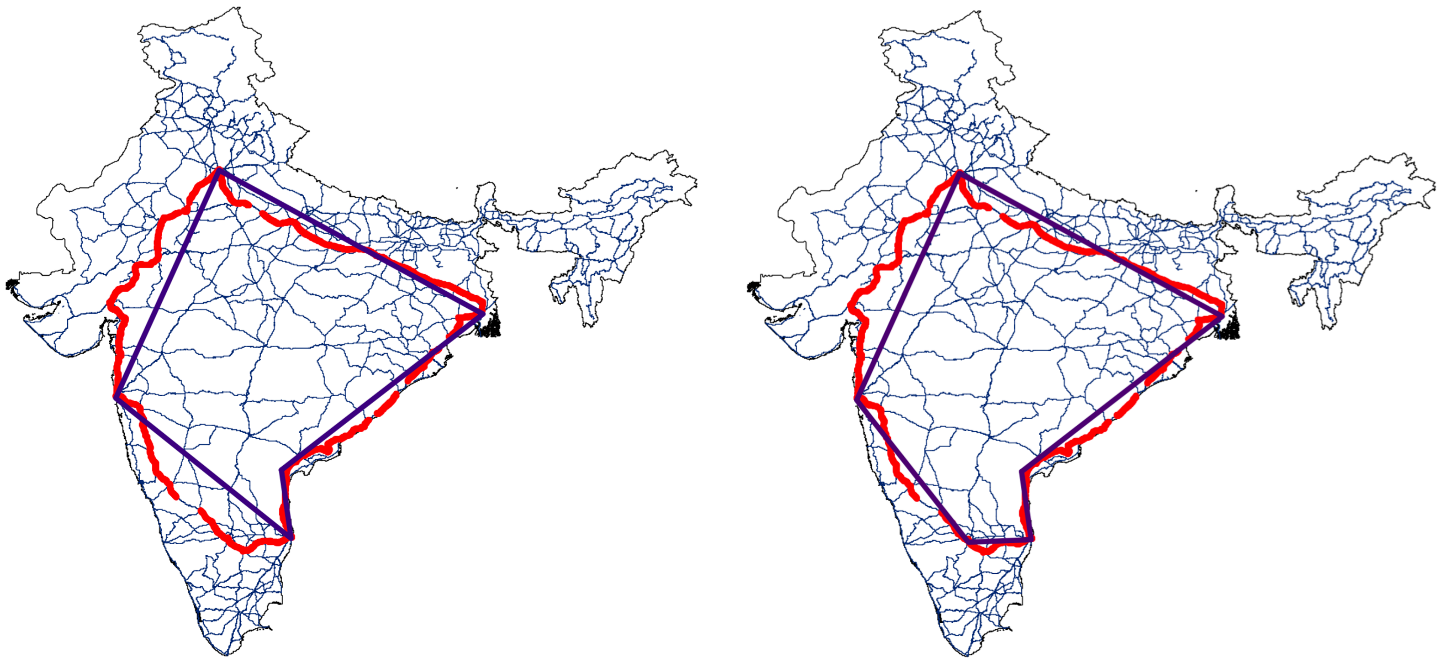
Notes: Figure XI shows the estimates of equation (26) at each category of distance. The dependent variable is the log change in the price of input j between 2001 and 2006 in district d. The coefficients depicted are those associated to the connectivity of the district, defined as whether the district is within a certain distance from the GQ in 2006 and 2001. Nodal districts are excluded. Panel A displays OLS coefficients and Panel B IV estimates. The instruments are the distance to the straight line connecting the four and five vertices of the GQ (Delhi, Chennai, Mumbai, Calcutta, and Bangalore). 95% confidence intervals stemming from robust standard errors clustered at the district level are drawn in thinner lines.

Recent empirical work, such as the work by Faber (2014), has emphasized the non-random placement of infrastructure even outside non-nodal areas. Thus, we check the robustness of our results by instrumenting the distance to the GQ that we use in estimating equation (26). In particular, we instrument it with the straight lines that connect the four nodal cities, shown in Panel A of Figure XII. These straight lines resemble the lowest-cost path connecting the nodal cities. The identifying assumption is that the distance to the straight line affects districts only through how likely they are to be close to the GQ network. We add a second IV specification with a straight line connecting the city of Bangalore, shown in Panel B of Figure XII. These set straight-line IVs were used by Ghani, Goswami, and Kerr (2016). Columns (3) and (4) of Table XII show the results of these IV specifications. We find that the effects of the GQ on prices follow a similar pattern and are somewhat higher in absolute value. In Panel B of Figure XI, we show that the overall pattern remains the same as in the baseline case.

Figure XII

Road Network in India, the GQ and the straight line GQ

A: Straight line IV B: Straight line IV through Bangalore



Panel A of Figure XII shows a map with the road network in India in 2006, including the sections of the Golden Quadrilateral that were finished by then (around 95% of the total project) and the IV straight line. Panel B shows the same map but making the straight line going through Bangalore.

8.2 The GQ and the Evolution of Allocative Efficiency

In order to explore the implications of the GQ on allocative efficiency, we analyze the changes in the Olley and Pakes (1996) within-industry covariance term between size and productivity before and after the GQ upgrades. Bartelsman, Haltiwanger, and Scarpetta (2013) show that this covariance term is a robust measure of misallocation, both theoretically and empirically.

Let ω_{ij} be a measure of productivity of firm i in industry j and θ_{ij} be the share of firm i in the industry. We define aggregate industry productivity Ω_j as:

$$\Omega_j = \bar{\omega}_j + \sum_{i=1}^{K_j} \left(\theta_{ij} - \bar{\theta}_j\right)\left(\omega_{ij} - \bar{\omega}_j\right)$$

where $\bar{\omega}_j$ is the unweighted average firm productivity, $\bar{\theta}_j$ is the unweighted average firm industry share, and K_j is the number of firms. Therefore, aggregate productivity can be decomposed into two terms: the unweighted average productivity and the within-industry covariance between size and productivity, which measures allocative efficiency. The intuition behind the covariance term is that as allocative efficiency improves, more productive firms should have higher industry shares. It is important to note that the covariance term remains constant in the models typically used to evaluate transportation infrastructure.

We compute the covariance term for each industry-district in both 2001 and 2006.[17] Firm productivity is calculated as log value added per worker and the shares are industry employment shares.[18] We estimate the following regression:

$$\Delta \text{OP}_{jd} = \beta \Delta \text{GQ}_d + \sum_j \alpha_j + \epsilon_{jd}, \tag{27}$$

where ΔOP_{jd} is the change between 2001 and 2006 in the Olley and Pakes covariance term in industry j in district d, GQ_d is a dummy variable taking value 1 if district d has been "treated" by the GQ, α_j are product fixed effects, and ϵ_{jd} is an error term. As before, distance is calculated as the shortest straight-line distance between the district and a treated portion of the GQ.

Table XIII shows the results of estimating equation (27). As before, column (1) includes nodal cities in the estimation and column (2) excludes them. We find improvements in allocative efficiency in the districts that became treated by the GQ compared with districts further away. The improvement in the covariance term is a decreasing function of distance, and converges rapidly towards zero. This pattern can be seen in Panel A of Figure XIII, which shows the coefficient of β found in column (2).

For robustness, we estimate the specification using the straight-line IVs described in Section 8.1. We find that the results do not change as shown in columns (3) and (4) of Table XIII. The confidence intervals increase in size due to the use of the IV. However, the results preserve the relationship that allocative efficiency increased more in districts close to the GQ. The coefficients of column (4) are plotted in Panel B of Figure XIII.

8.3 Model Output

In this section, we compare the empirical evidence regarding prices and allocative efficiency with the outcome of the model. In the results from the model, both the role of variable markups and general equilibrium effects induced by changes in transportation costs are considered.

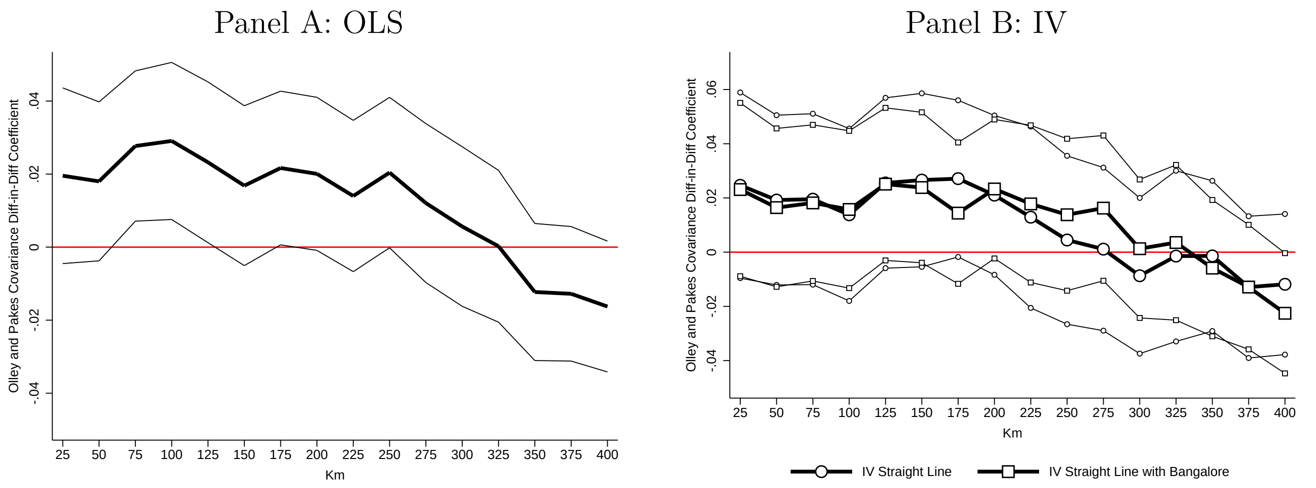
The estimates of the effect of the GQ on prices, presented in column (2) of Table XII, imply that prices in districts within 25-100 km of the GQ decreased by around 1.73-1.92 times as much as in the average district. A difference-in-differences regression with the model-simulated data shows that the decrease in prices charged in states crossed by the GQ is 1.94 times bigger than in the

[17] Each industry is a four-digit National Industrial Classification (NIC), which follows the procedures of the United Nations' International Standard Industrial Classification (ISIC). We map the 2006 codes, expressed in NIC 2004, to the 2001 codes, expressed in NIC 1998. Note that in this exercise, we consider industries (NIC) instead of products (ASICC). This allows us to use a level of aggregation that is sufficiently coarse to compute the covariance terms with a large number of observations and, at the same time, provide enough variation. This aggregation procedure is not possible with 5-digit ASICC data. Moreover, using ASICC entails the problem of dealing with multi-product plants.

[18] We restrict the sample to cells containing at least 10 plants. We also trim the 1% tails of the distribution of covariance changes in order to reduce the influence of outliers. We have 117 industries and 466 districts.

Figure XIII

Allocative efficiency and the Golden Quadrilateral Differences-in-Differences



Notes: Figure XIII shows the estimates of equation (27) at each category of distance. The dependent variable is the change between 2001 and 2006 of the Olley and Pakes (1996) within-industry cross-sectional covariance between labor share and labor productivity in industry (NIC) j and district d. The coefficients depicted are those associated to the connectivity of the district, which takes value one if the district was within a certain distance of the Golden Quadrilateral in 2006 but not in 2001, and zero otherwise. Nodal districts are excluded. Panel A displays OLS coefficients and Panel B IV estimates. The instruments are the distance to the straight line connecting the four and five vertices of the GQ (Delhi, Chennai, Mumbai, Calcutta, and Bangalore). 95% confidence intervals stemming from robust standard errors clustered at the district level are drawn in thinner lines.

average state. Therefore, the effect of the GQ on the model-simulated data is of the same order of magnitude as that found in the data. Regarding allocative efficiency, the reduced-form evidence, presented in column (2) of Table XIII, indicates that for districts within 25-100 km of the GQ, the OP covariance term improved by around 2.05-2.55 times as much as in the average district. The model predicts a smaller effect: the improvement in allocative efficiency is 1.29 times as high in states crossed by the GQ as in the average state.

Overall, we find that this evidence supports the differential evolution in prices and allocative efficiency predicted by model. It is important to note a few considerations. First, in the data, we carry out the analysis at the district level, in order to exploit as much variation as possible, whereas the model is calibrated at the state level. Second, the model compares two steady states, whereas in the data it may take some time for the full adjustment to take place.

9 Sensitivity Analysis

We now examine the sensitivity of our results by considering versions of our model in which we change the value of some of the crucial parameters. We first examine the implications of setting a lower value for the elasticity of substitution across sectors, θ. Second, we study a version of the model in which productivity shocks are uncorrelated across states.

For all these cases, we keep the rest of the parameters which we estimate outside the model constant, and re-calibrate the labor endowment for each state i, L_i, and the shape parameter of the Pareto distribution, α. To match the fact that the top 5% of plants in manufacturing account for 89% of value-added, the model requires a shape parameter of 1.61 in the case of $\theta = 1.24$, and 4.42 in the case of uncorrelated draws (vs 2.72 in our benchmark calibration).

We find that the aggregate gains are remarkably stable across specifications. The share of allocative efficiency gains is similar to the benchmark calibration in the case of the lower θ. However, allocative efficiency gains disappear in the case of uncorrelated productivity draws across states.

A lower elasticity of substitution across sectors We set $\theta = 1.24$, which is the value estimated by Edmond, Midrigan, and Xu (2015) using Taiwanese data. In this economy, the maximum markup a firm can charge is 5.17 (vs 2.01). There is more misallocation than in the benchmark economy: the allocative efficiency index ranges from 0.89 to 0.92 across states, whereas in the benchmark calibration it ranges from 0.96 to 0.98. The reason is that the lower θ implies that firms with large market shares charge higher markups, increasing the dispersion of markups.

It is interesting to note that the results do not change significantly relative to our baseline case. In this specification, allocative efficiency gains increase to 0.25% (vs 0.21%). The share of allocative efficiency gains increases to 8.3% of the gains (vs 7.7%). At the state level, allocative efficiency gains represent up to 21% of the overall gains (vs 19%).

A value of 1.24 for θ would imply a trade elasticity for monopolists that is too low compared to the one we estimate using Indian data.

Uncorrelated productivity draws We next examine how our results change if we have uncorrelated productivity draws across locations. We find that aggregate gains increase to 3.10% (vs 2.71%). Furthermore, allocative efficiency gains do not account for any of the aggregate gains. However, the case of uncorrelated productivity draws does not match the data along two important dimensions. First, when we calculate our similarity index for this economy, we find a value of 0.27. This is lower than 0.49, which we obtained in our baseline calibration, and lower than 0.45, which we measure in the data. Furthermore, the trade elasticity is 2.86 in the case of uncorrelated productivity draws, which is too low relative to those estimated in the literature. On the other hand, the trade elasticity of 4.74 implied by the baseline calibration is consistent with estimates from the literature.

The degree of head-to-head competition that firms are confronted with is too low in the case of uncorrelated draws.

10 Conclusions

The goal of this paper is to quantitatively evaluate the welfare effects of improving transportation infrastructure. We use the case of the GQ in India, which is an important road infrastructure project in the early 2000s. First, we quantify the aggregate gains from the GQ and the gains across states. Then, we decompose these gains to determine that which can be accounted for by changes in allocative efficiency. These finds highlight the fact that gains from allocative efficiency can change both the aggregate gains and the distribution of the gains that arise from new transportation infrastructure. Furthermore, we use the natural experiment of the GQ to find evidence for the main model mechansisms in the data.

From our point of view, the future direction of the misallocation literature is to identify concrete mechanisms that drive misallocation. A recent example of this kind of work is David, Hopenhayn, and Venkateswaran (2015), which studies misallocation within a framework that has information frictions. Furthermore, a natural way of approaching this agenda is to use natural experiments to shed light on promising mechanisms. This paper aims to contribute to this literature along both of these dimensions.

Table I

Impact of Road Distance and Infrastructure Quality on Prices

	2001 & 2006	2006	Transport Prices	2006	2001
	(1)	(2)	(3)	(4)	(5)
Dep. Variable: All columns except 4: Log price at district of destination. Column 4: Log price of transportation cost.					
Effective Distance 2nd decile	0.2855**	0.2429*	0.2675***		
	(0.1196)	(0.1440)	(0.0719)		
Effective Distance 3th decile	0.2120*	0.1797	0.5394***		
	(0.1168)	(0.1442)	(0.0714)		
Effective Distance 4th decile	0.0981	0.0618	0.9513***		
	(0.1206)	(0.1554)	(0.0662)		
Effective Distance 5th decile	0.1305	0.0114	1.1635***		
	(0.1351)	(0.1582)	(0.0658)		
Effective Distance 6th decile	0.3538***	0.3784**	1.2869***		
	(0.1320)	(0.1731)	(0.0663)		
Effective Distance 7th decile	0.3009**	0.1835	1.3895***		
	(0.1390)	(0.1747)	(0.0663)		
Effective Distance 8th decile	0.3491**	0.2615	1.4892***		
	(0.1510)	(0.1814)	(0.0676)		
Effective Distance 9th decile	0.2476*	0.3279*	1.6771***		
	(0.1485)	(0.1914)	(0.0680)		
Effective Distance 10th decile	0.5107***	0.5770***	1.9164***		
	(0.1439)	(0.1990)	(0.0698)		
Log Effective Distance				0.0765***	
				(0.0293)	
Predicted Price in 2001					0.7952*
					(0.4594)
Origin-Year Fixed Effects	YES	YES	-	YES	-
Product-Year Fixed Effects	YES	YES	-	YES	-
Origin Fixed Effects	-	-	YES	-	YES
Product Fixed Effects	-	-	-	-	YES
Observations	1,999	1,460	1,372	1,460	539
R-squared	0.87	0.86	0.82	0.86	0.88
Number of products	165	119	-	119	53
Number of origins	86	63	75	63	38
Number of destinations	367	338	319	338	171

Notes: This table shows the estimation of equation (17). The dependent variable is the log price of a product manufactured by a monopolist at destination. The variable of interest is the effective distance between the district where the product is manufactured and the district of destination. Effective distance is defined as the lowest cost path between both districts, taking into account road distance and infrastructre quality. Specifically, going across the Golden Quadrilateral reduces road distance 48 per cent, relative to roads not in the Golden Quadrilateral. The lowest path is computed by means of road networks and applying the Dijkstra's search path alogrithm. Columns (1) and (2) use our ASI-NSS data for 2001 & 2006 and 2006, respectively. The dependent variable of column (3) is the log price of shipping a container from origin to destination, according to GIR Logistics. Column (4) introduces effective distance linearly and estimate the regression with data of 2006, and column (5) compares the estimated prices in 2001 from the coefficient of column (4) to the actual prices in 2001. All pooled specifications include origin-year and product-year fixed effects. Robust standard errors are in parenthesis. Significance levels: *: 10%; **: 5%; ***: 1%.

Table II
Gravity Equations for Monopolists

	ASI-NSS Years 2001 & 2006	ASI-NSS Year 2006	Transport Costs as Effective Distance
	(1)	(2)	(3)
Dep. Variable: Log value of sales at district of destination			
$\log \tau_d^o$	-0.9917*	-0.8378	-0.2085***
	(0.5821)	(0.5588)	(0.0647)
Origin-Year Fixed Effects	YES	YES	YES
Destination-Year Fixed Effects	YES	YES	YES
Product-Year Fixed Effects	YES	YES	YES
Observations	1,999	1,460	1,999
R-squared	0.58	0.54	0.59
Number of Origins	86	63	86
Number of Destinations	367	338	367
Number of Products	165	119	165

Notes: This table shows the estimates of equation (22). The dependent variable is the log value of sales at destination of products manufactured by monopolists. The variable of interest is the predicted values of equation (17), namely the predicted transport costs across districts. In columns (1) and (2) the predicted transport costs are those derived from column (1) and (2) of Table I, i.e. $\log \hat{\tau}^o_{d,t} = \hat{\beta}_k$ if Effective Distance$^o_{d,t} \in$ decile k. In column (3) we assume the functional form $\log \hat{\tau}^o_{d,t} = \hat{\beta}$Effective Distance$^o_{d,t}$, and impose $\beta = 1$. Hence, we include effective distance directly as a covariate in equation (22). This amounts to assuming that the differences in effective distance (transport infrastructure) fully translate into prices. All specifications include origin-year, destination-year, and product-year fixed effects. Robust standard errors are in parenthesis. Significance levels: *: 10%; **: 5%; ***: 1%.

Table III
Labor Shares vs Sectoral shares

	2001-2006		2006	
	Labor	Capital+Labor	Labor	Capital+Labor
	(1)	(2)	(3)	(4)
Dep. Variable: Share in firm's value added				
Firm's sectoral share	-0.3407***	-0.4633***	-0.4088***	-0.4932***
	(0.076)	(0.0937)	(0.0941)	(0.1058)
State-Year Fixed Effects	YES	YES	YES	YES
Product-Year Fixed Effects	YES	YES	YES	YES
Observations	2,257	1,510	1,166	1,008
R-squared	0.86	0.88	0.87	0.89

Notes: Table III shows the estimates of an OLS regression of firms' labor shares against sectoral shares (equation 24) for products (measured at the ASICC 5-digit level) that are operated only in one state. Column (1) shows the results including the pool of observations for the years 2001 and 2006. Column (2) shows the results for the same regression but includes capital remuneration on the left hand side. Columns (3) and (4) are the equivalent but include only observations for the year 2006. Robust standard errors are in parenthesis: *: 10%; **: 5%; ***: 1%. The implied γ's in columns 1-4, which are given by $\gamma = \hat{\theta} / \left(1 + \hat{\beta}\hat{\theta}\right)$ (with $\hat{\theta} = 1.99$), are 10.67, 107.38, 6.17, and 25.30 respectively.

Table IV

Markups in the Model (by destination market)

state	std			mean			p95			p99			log p99/p50		
	before	after	% change	before	after	% change	before	after	% change	before	after	% change	before	after	% change
Maharashtra	0.0470	0.0464	-1.1765	1.1071	1.1071	-0.0029	1.122	1.123	0.086	1.294	1.287	-0.527	0.163	0.157	-3.296
Gujarat	0.0464	0.0459	-1.1177	1.1071	1.1070	-0.0027	1.123	1.124	0.073	1.288	1.282	-0.430	0.158	0.153	-2.763
Tamil Nadu	0.0463	0.0457	-1.2603	1.1071	1.1070	-0.0031	1.122	1.124	0.129	1.286	1.278	-0.637	0.156	0.150	-4.157
Uttar Pradesh	0.0448	0.0445	-0.5815	1.1070	1.1070	-0.0014	1.126	1.127	0.076	1.265	1.261	-0.325	0.140	0.137	-2.353
Karnataka	0.0457	0.0451	-1.3596	1.1070	1.1070	-0.0033	1.123	1.125	0.183	1.274	1.266	-0.607	0.147	0.141	-4.221
Andhra Pradesh	0.0449	0.0447	-0.6006	1.1070	1.1070	-0.0014	1.126	1.127	0.068	1.264	1.261	-0.257	0.139	0.136	-1.867
West Bengal	0.0447	0.0448	0.1893	1.1070	1.1070	0.0004	1.127	1.127	0.020	1.266	1.266	0.027	0.140	0.141	0.190
Haryana	0.0447	0.0443	-0.7286	1.1070	1.1069	-0.0017	1.128	1.129	0.093	1.261	1.257	-0.301	0.137	0.134	-2.224
Jharkhand	0.0446	0.0446	0.1594	1.1070	1.1070	0.0004	1.129	1.128	-0.083	1.262	1.264	0.177	0.137	0.139	1.283
Rajasthan	0.0446	0.0444	-0.3586	1.1070	1.1069	-0.0009	1.127	1.128	0.042	1.260	1.259	-0.074	0.136	0.135	-0.550
Madhya Pradesh	0.0444	0.0443	-0.2230	1.1069	1.1069	-0.0005	1.128	1.129	0.033	1.258	1.257	-0.097	0.135	0.134	-0.723
Punjab	0.0445	0.0444	-0.3487	1.1070	1.1069	-0.0008	1.128	1.129	0.050	1.258	1.256	-0.161	0.135	0.133	-1.203
Orissa	0.0444	0.0442	-0.4290	1.1069	1.1069	-0.0010	1.128	1.129	0.066	1.260	1.257	-0.240	0.136	0.133	-1.786
Himachal Pradesh	0.0446	0.0445	-0.1865	1.1070	1.1070	-0.0005	1.128	1.128	0.010	1.261	1.259	-0.119	0.136	0.135	-0.873
Chattisgarh	0.0445	0.0444	-0.1653	1.1070	1.1069	-0.0004	1.128	1.128	0.011	1.260	1.259	-0.077	0.136	0.135	-0.566
Kerala	0.0451	0.0449	-0.4800	1.1070	1.1070	-0.0012	1.125	1.126	0.073	1.265	1.263	-0.169	0.140	0.138	-1.217
Delhi	0.0445	0.0442	-0.7717	1.1070	1.1069	-0.0018	1.128	1.129	0.082	1.259	1.256	-0.250	0.135	0.132	-1.869
Uttaranchal	0.0445	0.0443	-0.3492	1.1069	1.1069	-0.0008	1.128	1.129	0.058	1.258	1.256	-0.147	0.134	0.133	-1.100
Goa	0.0452	0.0454	0.3751	1.1070	1.1070	0.0009	1.124	1.124	0.011	1.267	1.270	0.235	0.142	0.144	1.644
Assam	0.0443	0.0441	-0.3723	1.1069	1.1069	-0.0009	1.128	1.129	0.075	1.259	1.257	-0.210	0.135	0.133	-1.565
Bihar	0.0441	0.0445	0.9571	1.1069	1.1070	0.0022	1.130	1.128	-0.126	1.255	1.263	0.610	0.132	0.138	4.524
Jammu and Kashmir	0.0441	0.0442	0.2175	1.1069	1.1069	0.0005	1.130	1.129	-0.028	1.255	1.256	0.019	0.132	0.132	0.147
Meghalaya	0.0442	0.0440	-0.4393	1.1069	1.1069	-0.0011	1.129	1.130	0.095	1.256	1.254	-0.188	0.133	0.131	-1.427
Tripura	0.0440	0.0439	-0.1610	1.1069	1.1069	-0.0004	1.130	1.130	0.034	1.256	1.255	-0.096	0.133	0.132	-0.725
Manipur	0.0440	0.0440	-0.1422	1.1069	1.1069	-0.0003	1.129	1.130	0.026	1.256	1.255	-0.101	0.133	0.132	-0.762
Nagaland	0.0441	0.0439	-0.2565	1.1069	1.1069	-0.0006	1.129	1.130	0.045	1.256	1.254	-0.147	0.133	0.131	-1.115
Sikkim	0.0440	0.0440	0.0758	1.1069	1.1069	0.0001	1.130	1.130	0.018	1.254	1.254	-0.038	0.131	0.131	-0.289
Mizoram	0.0440	0.0439	-0.1386	1.1069	1.1069	-0.0003	1.130	1.130	0.029	1.256	1.254	-0.090	0.132	0.131	-0.683
Arunachal Pradesh	0.0440	0.0440	-0.1212	1.1069	1.1069	-0.0003	1.129	1.130	0.034	1.256	1.254	-0.097	0.132	0.131	-0.733

Notes: Table IV shows some moments of the unconditional markup distribution generated by the model; std, mean, p95, p99, and log p99/p50 refer to the standard deviation, simple mean, 95th percentile, 99th percentile and the percentage difference between the 99th percentile and the median of the markups charged to the goods purchased by each state respectively.

Table V

Parameter values (benchmark calibration)

Param.	Definition	Value
(A) Parameters estimated with structural equations		
τ_d^o	Iceberg transportation costs between states	varies by state pair
θ	Elasticity of substitution across sectors	1.99
γ	Elasticity of substitution within sector	10.67
(B) Parameters taken directly from data		
K_{ij}	Number of firms operating in sector j of country i	varies by state/sector
(C) Parameters calibrated in equilibrium		
L_i	Labor endowment of the states	varies by state
α	Shape parameter Pareto	2.72

Notes: Table V refers to a our benchmark calibration. We explain how we estimate the parameters τ_d^o, θ, and γ, in sections 6.1, 6.2, and 6.3 respectively. We set the value K_{ij} to match the number of plants observed in the data. We calibrate L_n to the relative manufacturing value added across states. We calibrate α to to match the fact that the top 5% of plants in manufacturing accounted for 89% of value-added in 2006 (see section 6.5 for details).

Table VI

Changes in Real Income Resulting from the GQ

state	size	income change	descomposition			
			η_w	$\eta_{P_{pc}}$	η_{ToT}	η_{ae}
India		2.71	2.10	0.42	0.00	0.21
Maharashtra	100.00	1.77	1.67	-0.38	0.15	0.33
Gujarat	64.58	3.04	2.34	0.27	0.05	0.39
Tamil Nadu	40.80	2.42	2.07	-0.05	0.07	0.33
Uttar Pradesh	28.88	2.09	1.81	0.23	-0.09	0.13
Karnataka	26.05	4.05	2.76	0.92	0.12	0.25
Andhra Pradesh	20.26	1.91	1.63	0.18	-0.01	0.11
Haryana	18.10	1.26	1.42	-0.20	-0.07	0.10
West Bengal	17.86	6.86	4.15	3.27	-0.50	-0.06
Jharkhand	16.24	8.29	4.37	3.76	0.02	0.15
Rajasthan	12.00	3.60	2.48	1.25	-0.16	0.03
Madhya Pradesh	10.71	0.53	0.94	-0.42	-0.02	0.04
Orissa	10.05	3.35	2.30	0.92	0.01	0.12
Punjab	9.65	1.47	1.39	0.09	-0.06	0.04
Himachal Pradesh	9.00	1.42	1.49	0.11	-0.22	0.04
Chattisgarh	8.76	-0.00	0.75	-0.71	-0.07	0.02
Kerala	7.02	1.79	1.47	0.20	0.02	0.10
Delhi	4.34	1.03	1.07	-0.27	0.13	0.10
Uttaranchal	4.33	1.65	1.47	0.19	-0.04	0.04
Assam	3.45	1.56	1.40	0.03	0.04	0.09
Goa	3.27	11.20	6.07	5.86	-0.46	-0.28
Bihar	2.49	7.24	4.23	3.99	-0.67	-0.32
Jammu and Kashmir	2.39	0.46	0.86	-0.18	-0.18	-0.04
Meghalaya	0.56	2.12	1.64	0.27	0.11	0.10
Tripura	0.24	-1.54	-0.20	-1.54	0.16	0.04
Manipur	0.11	-1.49	-0.22	-1.56	0.25	0.04
Nagaland	0.07	-0.62	0.20	-1.10	0.22	0.05
Sikkim	0.03	6.00	3.27	2.44	0.27	0.02
Mizoram	0.02	-1.09	0.18	-1.52	0.22	0.03
Arunachal Pradesh	0.01	-1.34	0.00	-1.68	0.32	0.03

Table VI shows the percentage change in real income and the decomposition of the Holmes, Hsu, and Lee (2014) index for the 29 Indian states; η_w represents the % change in labor income component of the index; $\eta_{P_{pc}}$ represents the % change in the *marginal cost price* component; η_{ToT} represents the % change in the *markups terms of trade* component; and η_{ae} represents the % change in the *allocative efficiency* component.

Table VII

Changes in the Components of the HHL Decomposition

state	w_n		$\frac{1}{P_n^{pc}}$		$\frac{\mu_n^{sell}}{\mu_n^{buy}}$		$\frac{P_n^{pc}}{P_n}\mu_n^{buy}$	
	before	after	before	after	before	after	before	after
Maharashtra	0.4216	0.4287	0.0121	0.0121	1.0272	1.0287	0.9620	0.9652
Gujarat	0.4233	0.4333	0.0117	0.0118	1.0159	1.0164	0.9655	0.9692
Tamil Nadu	0.4406	0.4498	0.0114	0.0114	0.9984	0.9991	0.9684	0.9716
Uttar Pradesh	0.4728	0.4815	0.0113	0.0114	0.9986	0.9978	0.9760	0.9773
Karnataka	0.4653	0.4783	0.0113	0.0114	0.9845	0.9856	0.9740	0.9764
Andhra Pradesh	0.4928	0.5009	0.0112	0.0113	0.9894	0.9893	0.9774	0.9785
Haryana	0.4785	0.4853	0.0113	0.0113	0.9805	0.9798	0.9779	0.9790
West Bengal	0.4841	0.5046	0.0109	0.0112	0.9867	0.9818	0.9761	0.9756
Jharkhand	0.4032	0.4212	0.0110	0.0114	0.9956	0.9958	0.9754	0.9769
Rajasthan	0.5005	0.5131	0.0112	0.0113	0.9866	0.9850	0.9783	0.9786
Madhya Pradesh	0.5003	0.5050	0.0113	0.0112	0.9737	0.9736	0.9789	0.9793
Orissa	0.4686	0.4795	0.0110	0.0111	0.9753	0.9754	0.9783	0.9795
Punjab	0.5189	0.5261	0.0110	0.0110	0.9903	0.9898	0.9789	0.9793
Himachal Pradesh	0.4724	0.4796	0.0109	0.0109	0.9803	0.9782	0.9778	0.9782
Chattisgarh	0.4477	0.4511	0.0111	0.0110	0.9698	0.9691	0.9780	0.9781
Kerala	0.5316	0.5395	0.0108	0.0108	0.9860	0.9862	0.9769	0.9779
Delhi	0.5615	0.5675	0.0113	0.0113	0.9668	0.9680	0.9789	0.9799
Uttaranchal	0.5117	0.5192	0.0110	0.0110	0.9834	0.9830	0.9790	0.9794
Assam	0.5022	0.5093	0.0102	0.0102	0.9789	0.9792	0.9783	0.9793
Goa	0.5177	0.5501	0.0109	0.0116	0.9595	0.9552	0.9765	0.9738
Bihar	0.5287	0.5515	0.0108	0.0112	0.9711	0.9646	0.9799	0.9768
Jammu and Kashmir	0.5435	0.5482	0.0104	0.0104	0.9935	0.9917	0.9802	0.9798
Meghalaya	0.5190	0.5276	0.0101	0.0102	0.9267	0.9277	0.9795	0.9805
Tripura	0.6265	0.6252	0.0099	0.0098	0.9940	0.9955	0.9801	0.9805
Manipur	0.6902	0.6887	0.0099	0.0098	1.0241	1.0267	0.9800	0.9804
Nagaland	0.7254	0.7269	0.0100	0.0098	1.0179	1.0202	0.9799	0.9804
Sikkim	0.7508	0.7757	0.0104	0.0106	1.0086	1.0113	0.9802	0.9804
Mizoram	0.9749	0.9767	0.0099	0.0097	1.0936	1.0960	0.9801	0.9804
Arunachal Pradesh	1.0000	1.0000	0.0099	0.0097	1.1888	1.1926	0.9801	0.9804

Table VII shows the level of the four different components of the Holmes, Hsu, and Lee (2014) index for the 29 Indian states before and after the construction of the GQ; w_n is the wage (note that we have excluded labor endowment, which is constant); P_n^{pc} is the aggregate price index in state n if all firms charged marginal cost; μ_n^{buy} represents the expenditure-weighted average markup charged on goods purchased in state n; μ_n^{sell} represents the revenue-weighted average markup charged on goods produced in state n; P_n is the aggregate price index in state n.

Table VIII

Effects of a partial change in transportation costs

	Arunachal P.	Mizoram	Sikkim	Nagaland	Manipur	Tripura	Meghalaya	Jammu and K.	Bihar	Assam	Goa	Uttaranchal	Delhi	Kerala	Chattisgarh	Himachal P.	Orissa	Punjab	Madhya P.	Rajasthan	Jharkhand	Haryana	WestBengal	Andhra P.	Karnataka	Uttar Pradesh	Tamil Nadu	Gujarat	Maharashtra
Arunachal P.	0.0000	-0.0000	-0.0000	-0.0000	-0.0000	0.0000	0.0000	0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
Mizoram	0.0000	0.0000	-0.0000	-0.0000	-0.0000	0.0000	0.0000	0.0000	0.0000	-0.0000	-0.0000	0.0000	0.0000	-0.0000	0.0000	0.0000	0.0000	-0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0000	0.0001	-0.0000	-0.0000	0.0000
Sikkim	0.0000	-0.0000	0.0000	-0.0000	-0.0000	0.0000	0.0000	0.0003	0.0000	-0.0000	0.0001	0.0006	0.0003	-0.0000	0.0002	0.0007	0.0001	0.0006	0.0006	0.0006	0.0000	0.0004	0.0000	0.0006	0.0004	0.0002	0.0003	0.0008	0.0007
Nagaland	0.0000	-0.0000	-0.0000	0.0000	-0.0000	0.0000	0.0000	0.0000	0.0000	-0.0000	-0.0000	0.0000	0.0003	-0.0000	0.0001	0.0000	0.0001	-0.0000	0.0002	0.0001	0.0001	0.0001	0.0000	0.0000	-0.0000	0.0002	-0.0000	-0.0000	0.0000
Manipur	0.0000	-0.0000	-0.0000	-0.0000	0.0000	0.0000	0.0000	0.0000	-0.0000	-0.0000	-0.0000	0.0000	-0.0000	-0.0000	-0.0000	0.0000	-0.0000	-0.0000	-0.0001	-0.0000	-0.0000	-0.0000	-0.0000	0.0000	-0.0000	-0.0001	-0.0000	-0.0000	0.0000
Tripura	0.0000	-0.0000	-0.0000	-0.0000	-0.0000	0.0000	0.0000	0.0000	0.0000	-0.0000	-0.0000	0.0000	0.0001	-0.0000	0.0000	0.0000	0.0001	-0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	-0.0000	0.0001	-0.0000	-0.0000	0.0000
Meghalaya	0.0000	-0.0000	-0.0000	-0.0000	-0.0000	0.0000	0.0000	0.0000	0.0000	-0.0000	-0.0000	0.0004	0.0004	-0.0000	0.0005	0.0002	0.0003	0.0002	0.0005	0.0003	0.0002	0.0004	0.0001	0.0001	-0.0000	0.0005	-0.0000	0.0002	0.0001
Jammu and K.	0.0000	-0.0000	0.0015	-0.0000	-0.0000	0.0000	0.0000	0.0000	0.0024	0.0002	0.0008	0.0000	-0.0000	-0.0000	0.0014	0.0000	0.0011	-0.0000	0.0004	0.0000	0.0026	-0.0000	0.0031	0.0001	0.0003	0.0003	-0.0000	0.0015	0.0018
Bihar	0.0001	0.0002	0.0001	0.0002	0.0003	0.0001	0.0001	0.0017	0.0000	0.0001	0.0023	0.0005	0.0003	0.0008	0.0005	0.0013	0.0004	0.0008	0.0012	0.0006	0.0219	0.0004	0.0021	0.0013	0.0022	0.0041	0.0024	0.0022	0.0030
Assam	0.0000	-0.0000	-0.0000	-0.0000	-0.0000	0.0000	0.0000	0.0002	0.0002	0.0000	-0.0000	0.0027	0.0035	-0.0000	0.0036	0.0016	0.0023	0.0018	0.0043	0.0023	0.0013	0.0030	0.0004	0.0009	0.0000	0.0044	0.0002	0.0010	0.0008
Goa	0.0000	-0.0000	0.0010	-0.0000	-0.0000	0.0000	0.0000	0.0018	0.0057	-0.0000	0.0000	0.0082	0.0093	0.0003	0.0038	0.0065	0.0066	0.0075	0.0048	0.0081	0.0049	0.0093	0.0050	0.0014	0.0210	0.0078	0.0010	0.0048	0.0183
Uttaranchal	0.0003	-0.0000	0.0052	0.0006	-0.0000	0.0002	0.0059	0.0000	0.0012	0.0041	0.0077	0.0000	-0.0000	0.0008	0.0026	0.0000	0.0038	-0.0000	0.0002	0.0000	0.0015	-0.0000	0.0039	0.0034	0.0052	0.0002	0.0018	0.0011	0.0049
Delhi	0.0010	0.0026	0.0042	0.0076	0.0010	0.0020	0.0091	0.0000	0.0012	0.0095	0.0089	0.0000	0.0000	0.0030	0.0010	0.0000	0.0025	-0.0000	0.0005	0.0000	0.0009	-0.0000	0.0016	0.0038	0.0064	0.0009	0.0044	0.0014	0.0026
Kerala	0.0000	-0.0000	-0.0000	-0.0000	-0.0000	0.0000	0.0000	0.0000	0.0059	-0.0000	0.0009	0.0025	0.0095	0.0000	0.0069	0.0008	0.0171	0.0017	0.0075	0.0140	0.0122	0.0077	0.0129	0.0026	0.0007	0.0051	0.0005	0.0166	0.0046
Chattisgarh	0.0005	0.0007	-0.0011	0.0012	0.0008	0.0007	-0.0003	0.0024	0.0014	0.0019	0.0079	0.0016	0.0005	0.0039	0.0000	0.0031	-0.0000	0.0019	0.0001	0.0017	0.0008	0.0009	0.0017	0.0005	0.0028	0.0016	0.0074	0.0043	0.0034
Himachal P.	0.0000	-0.0000	0.0120	-0.0000	-0.0000	0.0000	0.0057	0.0000	0.0055	0.0056	0.0145	0.0000	-0.0000	0.0006	0.0062	0.0000	0.0082	-0.0000	0.0005	0.0000	0.0067	-0.0000	0.0112	0.0057	0.0107	0.0002	0.0020	0.0026	0.0138
Orissa	0.0027	0.0046	0.0010	0.0043	0.0055	0.0048	0.0026	0.0053	0.0021	0.0043	0.0183	0.0078	0.0036	0.0172	0.0004	0.0084	0.0000	0.0071	0.0015	0.0073	0.0045	0.0053	0.0180	0.0059	0.0092	0.0038	0.0156	0.0085	0.0090
Punjab	0.0000	-0.0000	0.0153	-0.0000	-0.0000	0.0000	0.0135	0.0000	0.0054	0.0093	0.0167	0.0000	-0.0000	0.0013	0.0060	0.0000	0.0086	0.0000	0.0005	0.0000	0.0061	-0.0000	0.0121	0.0073	0.0119	0.0003	0.0035	0.0023	0.0116
Madhya P.	0.0045	0.0067	0.0160	0.0106	0.0077	0.0069	0.0192	0.0014	0.0088	0.0174	0.0124	0.0007	0.0008	0.0069	0.0004	0.0008	0.0020	0.0005	0.0000	0.0030	0.0070	0.0006	0.0099	0.0033	0.0056	0.0161	0.0071	0.0044	0.0034
Rajasthan	0.0012	0.0006	0.0154	0.0041	0.0006	0.0008	0.0187	0.0000	0.0046	0.0146	0.0231	0.0000	-0.0000	0.0137	0.0048	0.0000	0.0096	0.0000	0.0032	0.0000	0.0052	0.0001	0.0110	0.0127	0.0198	0.0030	0.0173	0.0308	0.0089
Jharkhand	-0.0098	-0.0223	-0.0035	-0.0223	-0.0165	-0.0153	-0.0093	-0.0098	-0.3168	-0.0101	0.0174	-0.0056	-0.0015	0.0076	-0.0014	-0.0037	-0.0105	-0.0053	-0.0056	-0.0014	0.0000	-0.0015	-0.3761	-0.0006	0.0110	-0.0148	0.0209	0.0226	0.0289
Haryana	0.0029	0.0024	0.0140	0.0120	0.0012	0.0033	0.0236	0.0000	0.0040	0.0234	0.0544	0.0000	-0.0000	0.0146	0.0047	0.0000	0.0121	-0.0000	0.0009	0.0001	0.0045	0.0000	0.0110	0.0194	0.0369	0.0017	0.0232	0.0053	0.0206
West Bengal	0.0007	0.0023	0.0003	0.0026	0.0012	0.0008	0.0007	0.0271	0.0173	0.0015	0.0259	0.0195	0.0058	0.0246	0.0046	0.0253	0.0380	0.0230	0.0171	0.0181	-0.0337	0.0108	0.0000	0.0154	0.0268	0.0120	0.0323	0.0309	0.0256
Andhra P.	0.0005	0.0001	0.0296	0.0019	-0.0000	0.0002	0.0114	0.0019	0.0242	0.0118	0.0114	0.0233	0.0202	0.0064	0.0025	0.0143	0.0116	0.0187	0.0074	0.0283	0.0131	0.0210	0.0175	0.0000	0.0068	0.0225	0.0304	0.0303	0.0086
Karnataka	0.0000	-0.0000	0.0223	-0.0000	-0.0000	0.0000	0.0000	0.0047	0.0423	0.0003	0.0490	0.0376	0.0440	0.0009	0.0089	0.0315	0.0170	0.0356	0.0128	0.0465	0.0350	0.0443	0.0356	0.0046	0.0000	0.0402	0.0120	0.0367	0.0353
Uttar Pradesh	0.0122	0.0139	0.0029	0.0128	0.0141	0.0143	0.0001	0.0005	0.0094	0.0102	0.0728	0.0003	0.0002	0.0169	0.0058	0.0004	0.0093	0.0005	0.0167	0.0047	0.0264	0.0012	0.0137	0.0299	0.0498	0.0000	0.0386	0.0120	0.0341
Tamil Nadu	0.0000	-0.0000	0.0197	-0.0000	-0.0000	0.0000	0.0000	0.0000	0.0504	0.0018	0.0078	0.0183	0.0411	-0.0015	0.0157	0.0078	0.0140	0.0141	0.0177	0.0557	0.0471	0.0385	0.0471	-0.0003	0.0011	0.0425	0.0000	0.0667	0.0150
Gujarat	0.0000	-0.0000	-0.0072	-0.0000	-0.0000	0.0000	-0.0070	-0.0333	0.0047	-0.0032	0.0564	-0.0092	-0.0081	0.0380	0.0005	-0.0092	0.0171	-0.0080	-0.0088	-0.1238	0.0272	-0.0049	0.0420	0.0319	0.0406	0.0008	0.0845	0.0000	0.0980
Maharashtra	-0.0030	-0.0004	-0.0622	-0.0051	-0.0000	-0.0010	-0.0123	-0.0181	-0.0291	-0.0115	-0.4513	-0.0679	-0.0249	-0.0328	-0.0246	-0.0392	-0.0397	-0.0477	-0.0203	-0.0196	0.0254	-0.0222	0.0258	-0.0171	-0.0575	-0.0145	0.0030	0.0339	0.0000
All states	0.0135	0.0104	-0.0182	0.0282	0.0153	0.0172	0.0669	-0.0305	-0.3530	0.0815	-0.2688	0.0329	0.1051	0.1244	0.0508	0.0443	0.1071	0.0487	0.0598	0.0147	0.0577	0.1168	-0.1890	0.1346	0.2311	0.1419	0.3413	0.3562	0.3927

Notes: Table VIII shows the effect of partial changes in transportation costs on allocative efficiency across the different Indian states. In particular, the element in row i and column j shows the effect on allocative efficiency of state j of decreasing transportation costs between state i and j (according to the counterfactual transportation network under the presence of the GQ) and keeping the rest of transportation costs in the network unchanged.

Table IX

Effects of a partial change in wages costs

	Aruchanal P.	Mizoram	Sikkim	Nagaland	Manipur	Tripura	Meghalaya	Jammu and K.	Bihar	Assam	Goa	Uttaranchal	Delhi	Kerala	Chattisgarh	Himachal P.	Orissa	Punjab	Madhya P.	Rajasthan	Jharkhand	Haryana	West Bengal	Andhra P.	Karnataka	Uttar Pradesh	Tamil Nadu	Gujarat	Maharashtra	India
Arunachal P.	0.0000	-0.0000	-0.0000	-0.0000	-0.0000	0.0000	0.0000	0.0000	-0.0000	-0.0000	-0.0000	0.0000	-0.0000	-0.0000	0.0000	0.0000	0.0000	-0.0000	-0.0000	0.0000	-0.0000	-0.0000	0.0000	0.0000	-0.0000	-0.0000	-0.0000	-0.0000	0.0000	-0.0000
Mizoram	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
Sikkim	-0.0004	-0.0004	-0.0004	-0.0004	-0.0004	-0.0004	-0.0004	-0.0003	-0.0004	-0.0003	-0.0003	-0.0004	-0.0003	-0.0003	-0.0004	-0.0003	-0.0003	-0.0003	-0.0003	-0.0003	-0.0003	-0.0003	-0.0003	-0.0003	-0.0003	-0.0003	-0.0003	-0.0003	-0.0003	-0.0003
Nagaland	-0.0000	-0.0000	-0.0000	-0.0001	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
Manipur	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	-0.0000	-0.0000	0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
Tripura	0.0000	0.0001	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Meghalaya	-0.0004	-0.0004	-0.0004	-0.0005	-0.0004	-0.0003	0.0042	-0.0001	-0.0005	-0.0008	-0.0001	-0.0002	-0.0002	-0.0001	-0.0003	-0.0001	-0.0002	-0.0001	-0.0002	-0.0001	-0.0004	-0.0001	-0.0004	-0.0001	-0.0001	-0.0002	-0.0001	-0.0001	-0.0001	-0.0002
Jammu and K.	-0.0006	-0.0006	-0.0004	-0.0006	-0.0006	-0.0006	-0.0005	0.0047	-0.0006	-0.0006	-0.0003	-0.0010	-0.0006	-0.0003	-0.0005	-0.0013	-0.0003	-0.0012	-0.0006	-0.0007	-0.0006	-0.0007	-0.0005	-0.0003	-0.0003	-0.0007	-0.0003	-0.0007	-0.0004	-0.0005
Bihar	-0.0027	-0.0031	-0.0029	-0.0032	-0.0031	-0.0029	-0.0031	-0.0024	-0.0070	-0.0034	-0.0015	-0.0022	-0.0020	-0.0013	-0.0027	-0.0026	-0.0023	-0.0022	-0.0021	-0.0022	-0.0085	-0.0021	-0.0041	-0.0016	-0.0015	-0.0040	-0.0019	-0.0024	-0.0023	-0.0026
Assam	0.0036	0.0025	-0.0015	0.0049	0.0035	0.0021	0.0047	-0.0009	-0.0020	0.0285	-0.0005	-0.0011	-0.0011	-0.0007	-0.0015	-0.0009	-0.0013	-0.0009	-0.0013	-0.0009	-0.0019	-0.0010	-0.0021	-0.0006	-0.0005	-0.0016	-0.0007	-0.0008	-0.0007	-0.0007
Goa	-0.0062	-0.0061	-0.0046	-0.0056	-0.0060	-0.0060	-0.0051	-0.0057	-0.0059	-0.0052	-0.0176	-0.0071	-0.0067	-0.0084	-0.0065	-0.0068	-0.0063	-0.0068	-0.0070	-0.0072	-0.0051	-0.0070	-0.0055	-0.0067	-0.0128	-0.0066	-0.0089	-0.0088	-0.0194	-0.0106
Uttaranchal	-0.0013	-0.0012	-0.0013	-0.0011	-0.0013	-0.0012	-0.0012	-0.0021	-0.0016	-0.0013	-0.0014	0.0076	-0.0015	-0.0009	-0.0016	-0.0024	-0.0013	-0.0023	-0.0014	-0.0017	-0.0014	-0.0019	-0.0016	-0.0010	-0.0011	-0.0018	-0.0009	-0.0021	-0.0018	-0.0015
Delhi	-0.0012	-0.0013	-0.0018	-0.0016	-0.0012	-0.0012	-0.0018	-0.0021	-0.0015	-0.0019	-0.0013	-0.0025	-0.0027	-0.0008	-0.0016	-0.0023	-0.0013	-0.0022	-0.0017	-0.0026	-0.0014	-0.0035	-0.0014	-0.0010	-0.0011	-0.0024	-0.0009	-0.0017	-0.0013	-0.0016
Kerala	-0.0043	-0.0042	-0.0031	-0.0040	-0.0041	-0.0041	-0.0037	-0.0033	-0.0031	-0.0038	-0.0050	-0.0029	-0.0030	0.0081	-0.0037	-0.0029	-0.0042	-0.0028	-0.0035	-0.0033	-0.0033	-0.0029	-0.0035	-0.0047	-0.0053	-0.0029	-0.0070	-0.0040	-0.0045	-0.0040
Chattisgarh	-0.0001	-0.0002	0.0008	-0.0000	-0.0002	-0.0002	0.0006	-0.0005	-0.0010	0.0001	-0.0015	-0.0003	-0.0004	-0.0010	0.0162	-0.0007	0.0001	-0.0005	-0.0004	-0.0008	-0.0013	-0.0006	-0.0015	-0.0008	-0.0013	-0.0009	-0.0018	-0.0019	-0.0026	-0.0012
Himachal P.	-0.0024	-0.0025	-0.0021	-0.0025	-0.0025	-0.0024	-0.0024	0.0043	-0.0033	-0.0027	-0.0033	-0.0011	-0.0016	-0.0025	-0.0031	0.0328	-0.0028	0.0046	-0.0028	-0.0027	-0.0033	-0.0015	-0.0035	-0.0025	-0.0029	-0.0033	-0.0025	-0.0047	-0.0047	-0.0025
Orissa	-0.0033	-0.0030	-0.0006	-0.0022	-0.0028	-0.0030	0.0002	-0.0041	-0.0048	-0.0021	-0.0064	-0.0034	-0.0035	-0.0061	-0.0037	-0.0044	0.0453	-0.0039	-0.0041	-0.0045	-0.0046	-0.0039	-0.0023	-0.0063	-0.0064	-0.0045	-0.0097	-0.0061	-0.0081	-0.0049
Punjab	-0.0036	-0.0035	-0.0040	-0.0033	-0.0035	-0.0035	-0.0038	-0.0040	-0.0043	-0.0037	-0.0033	-0.0052	-0.0045	-0.0021	-0.0039	-0.0041	-0.0032	0.0125	-0.0038	-0.0040	-0.0038	-0.0050	-0.0039	-0.0024	-0.0027	-0.0047	-0.0022	-0.0044	-0.0039	-0.0033
Madhya P.	-0.0023	-0.0023	-0.0023	-0.0023	-0.0023	-0.0023	-0.0021	-0.0025	-0.0030	-0.0024	-0.0026	-0.0026	-0.0025	-0.0021	-0.0029	-0.0029	-0.0024	-0.0025	0.0100	-0.0028	-0.0028	-0.0026	-0.0029	-0.0023	-0.0023	-0.0035	-0.0025	-0.0037	-0.0036	-0.0027
Rajasthan	-0.0068	-0.0067	-0.0069	-0.0067	-0.0067	-0.0066	-0.0069	-0.0067	-0.0085	-0.0074	-0.0086	-0.0078	-0.0060	-0.0065	-0.0081	-0.0083	-0.0070	-0.0073	-0.0078	0.0193	-0.0077	-0.0078	-0.0078	-0.0064	-0.0078	-0.0094	-0.0076	-0.0157	-0.0097	-0.0085
Jharkhand	0.0504	0.0622	0.0526	0.0680	0.0627	0.0610	0.0666	0.0238	0.1505	0.0554	-0.0086	0.0262	0.0170	0.0014	0.0217	0.0165	0.0378	0.0208	0.0212	0.0117	0.1342	0.0144	0.1564	0.0098	-0.0019	0.0319	-0.0100	-0.0145	-0.0219	0.0119
Haryana	-0.0050	-0.0049	-0.0028	-0.0045	-0.0050	-0.0047	-0.0028	-0.0010	-0.0063	-0.0047	-0.0091	-0.0011	0.0231	-0.0057	-0.0063	-0.0022	-0.0055	0.0005	-0.0058	-0.0036	-0.0066	0.0394	-0.0067	-0.0062	-0.0073	-0.0067	-0.0065	-0.0105	-0.0112	-0.0054
West Bengal	-0.0140	-0.0098	-0.0122	-0.0065	-0.0096	-0.0092	-0.0076	-0.0201	-0.0238	-0.0150	-0.0197	-0.0217	-0.0195	-0.0210	-0.0237	-0.0232	-0.0050	-0.0216	-0.0221	-0.0211	0.1018	-0.0208	0.0906	-0.0229	-0.0215	-0.0249	-0.0258	-0.0235	-0.0237	-0.0133
Andhra P.	-0.0079	-0.0077	-0.0072	-0.0074	-0.0077	-0.0075	-0.0069	-0.0065	-0.0086	-0.0076	-0.0126	-0.0074	-0.0073	-0.0120	-0.0095	-0.0068	-0.0087	-0.0069	-0.0085	-0.0081	-0.0087	-0.0075	-0.0094	0.0375	-0.0125	-0.0084	-0.0164	-0.0110	-0.0140	-0.0088
Karnataka	-0.0106	-0.0107	-0.0123	-0.0112	-0.0108	-0.0106	-0.0119	-0.0122	-0.0147	-0.0128	0.0217	-0.0131	-0.0125	-0.0084	-0.0134	-0.0137	-0.0102	-0.0132	-0.0128	-0.0140	-0.0144	-0.0133	-0.0146	-0.0116	0.0947	-0.0160	-0.0248	-0.0228	-0.0325	-0.0132
Uttar Pradesh	-0.0034	-0.0016	0.0013	0.0021	-0.0009	-0.0012	0.0082	-0.0003	0.0063	0.0027	-0.0153	-0.0005	0.0015	-0.0115	-0.0076	-0.0072	-0.0075	-0.0075	-0.0017	-0.0089	-0.0060	-0.0047	-0.0103	-0.0113	-0.0136	0.0775	-0.0134	-0.0169	-0.0183	-0.0061
Tamil Nadu	0.0024	0.0016	-0.0087	-0.0008	0.0011	0.0013	-0.0065	-0.0104	-0.0112	-0.0070	-0.0182	-0.0124	-0.0109	0.0366	-0.0050	-0.0126	0.0011	-0.0116	-0.0099	-0.0122	-0.0104	-0.0118	-0.0103	0.0108	0.0046	-0.0146	0.1639	-0.0237	-0.0313	0.0016
Gujarat	0.0253	0.0233	0.0279	0.0178	0.0226	0.0228	0.0173	0.0523	0.0151	0.0129	-0.0191	0.0374	0.0291	0.0072	0.0085	0.0343	-0.0008	0.0339	0.0278	0.0825	0.0033	0.0238	0.0002	-0.0005	-0.0069	0.0065	-0.0245	0.2024	-0.0490	0.0237
Maharashtra	0.0390	0.0352	0.0383	0.0301	0.0339	0.0342	0.0191	0.0211	0.0276	0.0175	0.1187	0.0429	0.0282	0.0441	0.0472	0.0298	0.0343	0.0338	0.0441	0.0231	0.0078	0.0241	0.0062	0.0346	0.0434	0.0185	-0.0005	-0.0037	0.1878	0.0590
All wages	0.0150	0.0240	0.0170	0.0273	0.0241	0.0235	0.0251	-0.0005	0.0655	0.0091	-0.0316	-0.0019	-0.0091	-0.0163	-0.0297	-0.0114	0.0163	-0.0081	-0.0141	0.0102	0.1155	-0.0160	0.1277	-0.0187	0.0072	-0.0023	-0.0137	0.0160	-0.0678	-0.0073

Notes: Table IX shows the effect of partial changes in wages on allocative efficiency across the different Indian states. In particular, the element in row i and column j shows the effect on allocative efficiency of state j of changing the equilibrium wage in state i (according to the new equilibrium under the presence of the GQ) and keeping the rest of wages unchanged.

Table X

% Change in Total Trade between i and j (Model)

	mean	median	sd/mean	N
both i and j in GQ	5.14	4.36	3.23	78
either i or j in GQ	1.41	-1.12	10.08	208
neither i nor j in GQ	1.93	0.41	7.29	120

Table X shows the mean, median, and coefficient of variation of the % change in total trade between states i and j after the construction of the GQ; "both i and j in GQ" refers to state pairs in which both of them are crossed by the GQ; "either i or j in GQ" refers to state pairs in which only one of them is crossed by the GQ; "neither i nor j in GQ" refers to state pairs in which none of them are crossed by the GQ. The states crossed by the GQ are Delhi, Bihar, Orissa, Rajasthan, Jharkhand, Haryana, West Bengal, Andhra Pradesh, Karnataka, Uttar Pradesh, Tamil Nadu, Gujarat, and Maharashtra. N is the number of state pairs that fall into the different categories.

Table XI

Changes in Real Income ACR (given trade flows) vs baseline model

state	baseline model welfare	ACR welfare	difference
Arunachal Pradesh	-1.34%	-1.39%	0.05%
Mizoram	-1.09%	-1.01%	-0.08%
Sikkim	6.00%	5.74%	0.26%
Nagaland	-0.62%	-1.02%	0.40%
Manipur	-1.49%	-1.92%	0.43%
Tripura	-1.54%	-1.96%	0.42%
Meghalaya	2.12%	1.91%	0.21%
Jammu and Kashmir	0.46%	0.48%	-0.02%
Bihar	7.24%	8.26%	-1.02%
Assam	1.56%	1.25%	0.31%
Goa	11.20%	11.70%	-0.50%
Uttaranchal	1.65%	1.50%	0.15%
Delhi	1.03%	1.24%	-0.21%
Kerala	1.79%	1.72%	0.08%
Chattisgarh	0.00%	-0.04%	0.04%
Himachal Pradesh	1.42%	1.23%	0.19%
Orissa	3.35%	2.75%	0.61%
Punjab	1.47%	1.40%	0.07%
Madhya Pradesh	0.53%	0.56%	-0.03%
Rajasthan	3.60%	3.46%	0.14%
Jharkhand	8.29%	5.35%	2.93%
West Bengal	6.86%	6.04%	0.83%
Andhra Pradesh	1.91%	1.80%	0.11%
Karnataka	4.05%	3.51%	0.54%
Uttar Pradesh	2.09%	1.90%	0.19%
Tamil Nadu	2.42%	1.97%	0.44%
Gujarat	3.04%	2.27%	0.77%
Maharashtra	1.77%	1.42%	0.35%
India	2.71%	2.28%	0.43%

Table XI shows, for each state, the welfare gains predicted by our baseline model, the welfare gains predicted by the ACR formula computed for the trade flows predicted by our baseline model, and the difference between the two.

Table XII
Prices and the Golden Quadrilateral: Differences-in-Differences

	OLS	OLS	IV	IV
	(1)	(2)	(3)	(4)
Dep. Variable: Log change input prices 2001 - 2006				
District within 25 km from GQ	-0.5932***	-0.6011***	-0.6737***	-0.6178***
	(0.1489)	(0.1544)	(0.1970)	(0.1876)
District within 50 km from GQ	-0.5588***	-0.5649***	-0.5122**	-0.4727**
	(0.1358)	(0.1391)	(0.2104)	(0.1986)
District within 100 km from GQ	-0.4139***	-0.4184***	-0.8373***	-0.7383***
	(0.1499)	(0.1542)	(0.2371)	(0.2128)
District within 150 km from GQ	0.0217	0.0349	-0.5550	-0.3210
	(0.1476)	(0.1517)	(0.3740)	(0.2231)
District within 200 km from GQ	-0.0137	-0.0040	-0.2593	-0.2448
	(0.1612)	(0.1683)	(0.2244)	(0.2050)
District within 300 km from GQ	-0.1113	-0.1061	0.1445	-0.0874
	(0.1583)	(0.1646)	(0.2007)	(0.1893)
Input Fixed-Effects	YES	YES	YES	YES
Nodal Districts	YES	NO	NO	NO
Instrument	-	-	Straight-line	Straight-line with Bangalore
Observations	5,123	5,037	5,037	5,037
Average R-Squared	0.42	0.42	-	-
Number of Products	920	912	912	912

Table XII shows the estimation of equation (26). The dependent variable is the log change in the price of input j between 2001 and 2006 in district d. The variable of interest is the connectivity of the district, defined as whether the district is within a certain distance from the GQ in 2006 and 2001. Each row correspons to a different regression, where different distances are considered. Columns (1) includes all districts whereas column (2) excludes nodal districts, both columns displaying OLS regressions. Column (3) instruments the distance to the GQ with the distance to the straight line connecting the four vertices of the GQ (Delhi, Chennai, Mumbai, and Calcutta). Column (4) instruments with the distance to a 5 vertices straight line (adding Bangalore). Input fixed effects are included in all specifications. Robust standard errors are in parenthesis, clustered at the district level. Significance levels: *: 10%; **: 5%; ***: 1%.

Table XIII

Allocative efficiency and the Golden Quadrilateral: Differences-in-Differences

	OLS	OLS	IV	IV
	(1)	(2)	(3)	(4)
Dep. Variable: Change in Covariance Term 2001 - 2006				
District within 25 km from GQ	0.0143	0.0196	0.0247	0.0231
	(0.0123)	(0.0122)	(0.0174)	(0.0163)
District within 50 km from GQ	0.0139	0.0180	0.0192	0.0164
	(0.0111)	(0.0111)	(0.0159)	(0.0149)
District within 100 km from GQ	0.0249**	0.0291***	0.0138	0.0158
	(0.0110)	(0.0109)	(0.0162)	(0.0148)
District within 150 km from GQ	0.0131	0.0168	0.0266	0.0238*
	(0.0111)	(0.0111)	(0.0163)	(0.0141)
District within 200 km from GQ	0.0161	0.0201*	0.0210	0.0233*
	(0.0107)	(0.0107)	(0.0149)	(0.0130)
District within 300 km from GQ	0.0023	0.0057	-0.0087	0.0013
	(0.0112)	(0.0111)	(0.0146)	(0.0130)
Input Fixed-Effects	YES	YES	YES	YES
Nodal Districts	YES	NO	NO	NO
Instrument	-	-	Straight-line	Straight-line with Bangalore
Observations	6,832	6,721	6,721	6,721
Average R-Squared	0.02	0.02	-	-
Number of Industries	117	117	117	117

Notes: Table XIII shows the estimation of equation (27). The dependent variable is the change between 2001 and 2006 of the Olley and Pakes (1996) within-industry cross-sectional covariance between labor share and labor productivity in industry (NIC) j and district d. The variable of interest is the connectivity of the district, which takes value one if the district was within a certain distance of the Golden Quadrilateral in 2006 but not in 2001, and zero otherwise. Each row correspons to a different regression, where different distances are considered. Columns (1) includes all districts whereas column (2) excludes nodal districts, both columns displaying OLS regressions. Column (3) instruments the distance to the GQ with the distance to the straight line connecting the four vertices of the GQ (Delhi, Chennai, Mumbai, and Calcutta). Column (4) instruments with the distance to a 5 vertices straight line (adding Bangalore). Industry fixed effects are included in all specifications. Robust standard errors are in parenthesis, clustered at the district level. Significance levels: *: 10%; **: 5%; ***: 1%.

References

Adamopoulos, T. (2011): "Transportation Costs, Agricultural Productivity, and Cross-Country Income Differences," *International Economic Review*, 52(2), 489–521.

Aghion, P., R. Burgess, S. Redding, and F. Zilibotti (2005): "Entry Liberalization and Inequality in Industrial Performance," *Journal of the European Economic Association*, 3(2/3), 291–302.

Alder, S. (2014): "Chinese Roads in India: The Effect of Transport Infrastructure on Economic Development," *Working Paper*.

Arkolakis, C., A. Costinot, D. Donaldson, and A. Rodríguez-Clare (2015): "The Elusive Pro-Competitive Effects of Trade," *Working Paper*.

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012): "New Trade Models, Same Old Gains?," *American Economic Review*, 102(1), 94–130.

Atkeson, A., and A. Burstein (2008): "Pricing-to-Market, Trade Costs, and International Relative Prices," *The American Economic Review*, 98(5), 1998–2031.

Banerjee, A. V., E. Duflo, and N. Qian (2012): "On the Road: Access to Transportation Infrastructure and Economic Growth in China," *Working Paper*.

Bartelsman, E., J. Haltiwanger, and S. Scarpetta (2013): "Cross-Country Differences in Productivity: The Role of Allocation and Selection," *American Economic Review*, 103(1), 305–334.

Bayoumi, T., and B. Eichengreen (1998): "Is Regionalism Simply a Diversion? Evidence from the Evolution of the EC and EFTA," in *Regionalism Versus Multilateral Trade Arrangements*, ed. by T. Ito, and A. Kruger. University of Chicago Press.

Bernard, A. B., J. Eaton, J. B. Jensen, and S. Kortum (2003): "Plants and Productivity in International Trade," *The American Economic Review*, 4, 1268–1290.

Bollard, A., P. Klenow, and G. Sharma (2013): "India's Mysterious Manufacturing Miracle," *Review of Economic Dynamics*, 16, 59–85.

Chari, A. V. (2011): "Identifying the Aggregate Productivity Effects of Entry and Size Restrictions: An Empirical Analysis of License Reform in India," *American Economic Journal: Economic Policy*, 3(2), 66–96.

Datta, S. (2012): "The Impact of Improved Highways on Indian Firms," *Journal of Development Economics*, 99(1), 46–57.

David, J. M., H. A. Hopenhayn, and V. Venkateswaran (2015): “Information, Misallocation and Aggregate Productivity,” *Working Paper.*

de Blas, B., and K. N. Russ (2015): “Understanding Markups in the Open Economy,” *American Economic Journal: Macroeconomics*, 7(2), 157–80.

De Loecker, J., P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik (2016): “Prices, Markups, and Trade Reform,” *Econometrica*, 84(2), 445–510.

De Loecker, J., and F. Warzynski (2012): “Markups and Firm-Level Export Status,” *American Economic Review*, 102(6), 2437–2471.

Dhingra, S., and J. Morrow (2014): “Monopolistic Competition and Optimum Product Diversity Under Firm Heterogeneity,” *Working Paper.*

Donaldson, D. (Forthcoming): “Railroads of the Raj: the Economic Impact of Transportation Infrastructure,” *American Economic Review.*

Donaldson, D., and R. Hornbeck (Forthcoming): “Railroads and American Economic Growth: A "Market Access" Approach,” *Quarterly Journal of Economics.*

Eaton, J., and S. Kortum (2002): “Technology, Geography, and Trade,” *Econometrica*, 70(5), 1741–1779.

Edmond, C., V. Midrigan, and D. Y. Xu (2015): “Competition, Markups, and the Gains from International Trade,” *American Economic Review*, 105(10), 3183–3221.

Epifani, P., and G. Gancia (2011): “Trade, Markup Heterogeneity and Misallocations,” *Journal of International Economics*, 83(1), 1–13.

Faber, B. (2014): “Trade Integration, Market Size, and Industrialization: Evidence from China’s National Trunk Highway System,” *Review of Economic Studies*, 81(3), 1043–1070.

Feenstra, R. C. (2014): “Restoring the Product Variety and Pro-competitive Gains from Trade with Heterogeneous Firms and Bounded Productivity,” *NBER Working Paper No. 19833.*

Feenstra, R. C., and D. E. Weinstein (Forthcoming): “Globalization, Competition, and U.S. Welfare,” *Journal of Political Economy.*

Garcia-Santana, M., and J. Pijoan-Mas (2014): “The Reservation Laws in India and the Misallocation of Production Factors,” *Journal of Monetary Economics*, 66(0), 193 – 209.

Ghani, E., A. G. Goswami, and W. R. Kerr (2016): “Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing,” *The Economic Journal*, 126(591), 317–357.

Guner, N., G. Ventura, and Y. Xu (2008): "Macroeconomic Implications of Size-Dependent Policies," *Review of Economic Dynamics*, 11(4), 721–744.

Head, K., and T. Mayer (2014): "Gravity Equations: Workhorse, Toolkit, Cookbook," in *Handbook of International Economics*, ed. by G. Gopinath, E. Helpman, and K. Rogoff, vol. 4, chap. 3, pp. 131–195.

Herrendorf, B., J. A. Schmitz, Jr., and A. Teixeira (2012): "The Role of Transportation in U.S. Economic Development: 1840-1860," *International Economic Review*, 53(3), 693–716.

Holmes, T. J., W.-T. Hsu, and S. Lee (2014): "Allocative Efficiency, Mark-ups, and the Welfare Gains from Trade," *Journal of International Economics*, 94(2), 195 – 206.

Hsieh, C.-T., and P. Klenow (2009): "Misallocation and Manufacuring TFP in China and India," *Quaterly Journal of Economics*, 124(4), 1403–1448.

Kothari, S. (2013): "The Size Distribution of Manufacturing Plants and Development," *Working Paper*.

Krueger, A. O. (1999): "Trade Creation and Trade Diversion Under NAFTA," *National Bureau of Economic Research Working Paper 7429*, (7429).

Melitz, M. J., and G. I. P. Ottaviano (2008): "Market Size, Trade, and Productivity," *The Review of Economic Studies*, 75(1), 295–316.

Olley, S., and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64(6), 1263–1297.

Peters, M. (2013): "Heterogeneous Mark-Ups, Growth, and Endogenous Misallocation," *Working Paper*.

Redding, S., and M. A. Turner (2015): "Transportation Costs and the Spatial Organization of Economic Activity," in *Handbook of Urban and Regional Economics*, ed. by G. Duranton, J. V. Henderson, and W. Strange, vol. 5, chap. 20, pp. 1339–1398.

Restuccia, D., and R. Rogerson (2008): "Policy Distorsions and Aggregate Productivity with Heterogeneous Establishments," *Review of Economic Dynamics*, 11(4), 707–720.

Van Leemput, E. (2015): "A Passage to India: Quantifying Internal and External Barriers to Trade," *Working Paper*.

World-Bank (2002): "India's Transport Sector: The Challenges Ahead," *World Bank Reports Series*, 1, 1–65.