# 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

Université Libre de Bruxelles (ECARES)

Roberto Ramos

Bank of Spain

October 10, 2014

**Abstract**

In this paper, we quantitatively evaluate the benefit of improving transportation infrastructure. We do so by developing a model of internal trade in which asymmetric states trade with each other. Firms compete oligopolistically at the industry level, allowing for markups to change with changes in transportation costs. We apply the model to measure the welfare effects of building a large road infrastructure project in India: the Golden Quadrilateral (GQ). After calibrating our model to rich plant-level and geospatial data, we find large gains: benefits exceed the initial investment in just two years. We also find that: (i) pro-competitive gains are approximately 20% of total gains and (ii) the size of welfare gains are very heterogeneous across states.

*JEL classifications*: F1, O4.

*Keywords*: Internal Trade, Welfare, Infrastructure, Misallocation.

*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, Thomas Holmes, Timothy Kehoe, Virgiliu Midrigan, Scott Petty, John Romalis, James Schmitz, and seminar participants at ECARES, Midwest Macro (Minneapolis), Spanish Economics Association Meeting (Santander), the University of Namur, CEMFI, Tsingua Macro Workshop, SED (Toronto), North American Meeting of the Econometric Society (Minneapolis), the RIEF network annual meeting, ANU in Canberra, WAMS in Melbourne, the University of Queensland, the Indian Statistical Institute, the Delhi School of Economics, the European Economic Association in Toulouse, and the REDg Workshop (Madrid). 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: msantana@ulb.ac.be; Ramos: roberto.ramos@bde.es.

# 1 Introduction

Poor transportation infrastructure is a common feature in low-income countries. For example, in 2000, it would take a truck four to five days to drive the 1,500 km distance between Delhi and Calcutta, which is five times longer than it would in the United States. International organizations and policymakers have not overlooked this fact: between 1995 and 2005, upgrades to the transportation network constituted around 12% of total World Bank lending. Out of this, 75% was allocated to the upgrading of roads and highways. Hence, understanding the impact of large-scale transportation infrastructure projects is a matter of great importance.

In this paper, we develop a model of internal trade that allows us to quantitatively evaluate the welfare gains which stem from improving the transportation infrastructure within a country. Our main contribution is to quantify the impact of improved transportation networks in a setting which allows to distinguish between different types of welfare gains. That is, we determine to what extent reductions in transportation costs improve productive efficiency (Ricardian gains), allocative efficiency (less misallocation due to an increase in competition), and the terms of trade for every trade partner. We use this model to study the welfare impact of building one of the biggest highway networks in the world: the Golden Quadrilateral (GQ) in India. The GQ project upgraded and expanded the roads connecting the four major cities in the country, providing India with around 6,000 km of modern highway roads.

In our model, the states of India trade with each other. There is a continuum of sectors and each sector has firms of heterogeneous productivity competing à la Cournot. In states where transportation costs are high, local firms enjoy market power that allows them to charge high markups in the local market. This creates a wedge between the markups of firms operating at home and the markups of firms operating outside their state. When there is an improvement in infrastructure, transportation costs decline, reducing the geographical advantage of local firms and equalizing markups across all producers, which translates into an improvement in the allocative efficiency of the economy. Our model is similar to the one used by Atkeson and Burstein (2008) and Edmond, Midrigan, and Xu (2012), extended to include multiple non-symmetric economies.

Relative to the literature on transportation and development, our framework allows us to separate the standard Ricardian channel from the pro-competitive and the terms of trade channels to account for the welfare gains stemming from lower transportation costs. Our decomposition of the welfare gains follows the methodology developed by Holmes, Hsu, and Lee (Forthcoming). The Ricardian component is simply the gains in real income if all firms charged their marginal cost. The pro-competitive gains relate to the misallocation arising from the heterogeneous markups charged by firms. This misallocation arises due to the fact that the consumption of goods produced by firms with high markups is inefficiently low. The last component is the terms of trade, which compares the average markup of the goods sold with the average markup of the goods purchased by the state. Ceteris paribus, states with high markups on the goods that they sell relative to the goods that they buy, will enjoy a higher real income.

In order to discipline the key parameters of the model, we make use of a rich micro data set of Indian manufacturing plants and geospatial data on the Indian road network at several points in time. First, we combine two separate sets of micro-level data - the Annual Survey of Industries (ASI) and the National Sample Survey (NSS) - in order to construct a very detailed description of the Indian manufacturing sector over time, covering both formal and informal firms. From these data, among other things, we derive measures of internal trade and prices paid across destinations. Second, we use GIS information on the entire Indian road network in order to compute measures of effective distance across destinations, taking into account the quality of the roads and the evolution of the transportation network over time.

We derive a set of structural equations from the model that allow us to estimate the key parameters. One implication of the model is that transportation costs can be identified by comparing the prices charged by monopolistic firms across destinations. This is the case because the prices charged by these firms only depend on transportation costs across locations, as the level of competition they face is constant. To implement this strategy, we first identify all the goods that are produced by only one plant in India. We then regress the prices paid for these goods across locations against the effective distance between the location of the monopolistic producer and the location of the plant that uses it as an intermediate input. Our measure of effective distance takes the least costly path along the Indian road network into account, incorporating differences in road quality caused by the presence of the GQ. Using the coefficients of the regression we construct a matrix of bilateral transportation costs between Indian states for both 2001 (before the GQ) and 2006 (after the GQ).

Our next step is to identify the elasticity of substitution across sectors. This parameter governs the price elasticity of the demand curve of a sector, and hence it determines the market power of firms that are monopolists in their sector. We use intermediate input usage data to construct trade flows for goods produced by monopolists. For these goods, the model implies a gravity equation that relates bilateral flows to transportation costs. We use internal trade flows and estimated transportation costs to measure how trade flows decline with increases in transportation costs. We set the elasticity of substitution across sectors to match the gravity equation of monopolist trade flows in the data.

We next estimate the elasticity of substitution across products within the same sector. This parameter determines the elasticity of demand faced by firms with a small market share in their sectors. In order to do that, we exploit a linear relationship between sectoral shares and labor shares implied by the model. In the model, firms with higher sectoral shares also charge higher markups, and hence have lower labor shares. The strength of this relationship depends on the gap between the elasticity of substitution both across and within sectors. Given our estimate of the elasticity of substitution across sectors, we set an elasticity of substitution within sectors that matches the slope coefficient of an OLS cross-sectional regression of the labor shares of plants against their sectoral shares. The values of these elasticities are crucial to quantify the size of the pro-competitive gains, since they determine the amount of misallocation that can potentially occur in the economy. In particular, the gap between the two elasticities determine to what extent firms with higher sectoral shares enjoy more market power in their sector and charge higher markups. We set the rest of parameters of the model such that, in equilibrium, the model reproduces some important features of the Indian manufacturing sector.

Next, we quantify the effects of the construction of transportation infrastructure. To do so, we measure the impact of changing the transportation costs to those estimated for 2001, before the construction of the GQ, in the model. We find that the aggregate gains for India derived from the construction of the GQ are 2.04% of real income. Because we only considered the manufacturing sector in our model, the result is in terms of manufacturing value-added. Putting the welfare gains from the model into dollar amounts yields a gain of *$3.1 billion per year*. Since the GQ cost $5.6 billion to build, our model predicts that it would take only two years for India to recover the initial cost.

Importantly, we also find wide heterogeneity in terms of welfare effects across states. States closest to the GQ gained the most, while those farthest had modest or even negative welfare gains. The negative effects stemming from the construction of the GQ come from the interplay of two forces at work. First, these states benefited from lower transportation costs. Despite their location, shipments can still travel for at least part of the route on the GQ, allowing them to import goods at a lower price. Second, the states that are closer to the GQ start trading more intensively with each other, which implies increases in wages in these states. This translates into an increase of the cost of purchasing goods from these states. Some states which are far from the GQ lost because this higher cost of purchasing goods from other states was not compensated for by the decrease in prices due to lower transportation costs. Interestingly, these states actually became less open after the construction of the GQ (they reduced the value of exports as a fraction of state income). This is a result of *trade diversion*: states close to the GQ diverted their trade towards states that experienced a greater decline in transportation costs.

We find that, on aggregate, pro-competitive gains account for 19% of the total gains from the construction of the GQ. These pro-competitive gains are positive in all but one state, and can reach up to 24% in some of them. This means that the GQ helped reduce the misallocation arising from variation in the market power of firms. We also find wide heterogeneity in the effects of the terms of trade. In fact, some states lost more from the changes in the terms of trade than they gained through pro-competitive effects. In the aggregate, welfare changes in the terms of trade sum to zero. Thus, although the terms of trade do not have an aggregate impact, they can have important effects on the distribution of income across states.

Lastly, we apply a difference-in-difference strategy to our data in order to isolate the effect of the GQ on prices and compare it with the outcome of the calibrated model. To do so, we compare the prices paid for intermediate goods by firms close to the GQ and by those that are further away before and after the construction of the highway. This strategy accounts for the potential endogeneity of infrastructure development, by focusing on price changes in non-nodal districts close and further away from the road network. We find that, in the data, the change in prices in non-nodal districts crossed by the GQ was around 36 percentage points lower than those in districts further away, implying a 1.57 times bigger decrease in prices in districts crossed by the GQ. We find a similar effect in magnitude when computing the equivalent differential effect with our calibrated model. The model predicts that the decrease in prices in states crossed by the GQ is 2.70 times larger than in average state.

## 2 Related Literature

Our work builds on papers that quantify the gains from building transportation infrastructure using general equilibrium models of trade. The pioneering work of Donaldson (Forthcoming) studies the benefits from the construction of railroads in colonial India. Herrendorf, Schmitz, and Teixeira (2012) study the impact of transportation improvements during the 19th century United States transportation revolution on the regional distribution of population and welfare. Adamopoulos (2011) and Sotelo (2014) study the income losses due to high transportation costs for agricultural products in developing countries. Allen and Arkolakis (Forthcoming) use a novel theoretical framework to calculate the welfare effects of the construction of a interstate highway system in the United States. In a more recent paper, Donaldson and Hornbeck (2014) develop the "market access" approach in order to assess the gains from the construction of the railroads in the United States. Alder (2014) uses this approach to study the impact of the hypothetical construction in India of a highway network similar to that of China. All the papers in this literature do not consider how reductions in transportation costs affect markups. To our knowledge, our paper is the first attempt to evaluate how improvements in infrastructure impact welfare through the pro-competitive channel.

Our paper is also related to a large set of work that studies the pro-competitive effects of international trade. These papers study how trade affects the markups that firms charge and the resulting impact on welfare. Markusen (1981) is an example of early work in this area. Bernard, Eaton, Jensen, and Kortum (2003), de Blas and Russ (2010), Devereux and Lee (2001), Epifani and Gancia (2011), Licandro and Impullitti (2013), and Holmes, Hsu, and Lee (Forthcoming) study the pro-competitive effects of trade in a setting with oligopolistic competition. Melitz and Ottaviano (2008) study these effects in a setting with monopolistic competition. We differ from these papers in that our aim is to *quantify* the pro-competitive effects of reducing transportation costs. Such quantification is useful since theory is ambiguous as to whether pro-competitive effects are quantitatively significant. In fact, theory indicates that pro-competitive effects can be negative, as stressed by Arkolakis, Costinot, Donaldson, and Rodríguez-Clare (2012).

In this sense, our paper builds on Edmond, Midrigan, and Xu (2012). They quantify the pro-competitive gains channel by using a model in which Taiwan trades with a symmetric partner, which represents the rest of the world. We extend this analysis to a non-symmetric multi-country setting. This setting accounts for changes in labor income and terms of trade, which are not present in the symmetric case. Furthermore, the extended case allows for effects such as trade diversion and heterogeneity of the pro-competitive effects across Indian states.

Our paper is also related to Arkolakis, Costinot, and Rodriguez-Clare (2012) and the set of commonly used trade models that they consider in their paper. In these models, all firms charge the same markup or operate under the assumption of perfect competition. Our paper is different because it also considers the effects of changing markups after a reduction in trade costs.

Our work contributes to a large literature on the misallocation of resources across firms. Papers from this literature include Restuccia and Rogerson (2008), Guner, Ventura, and Yi (2008), and Hsieh and Klenow (2009).[1] We contribute to these papers by evaluating how improvements in transportation infrastructure alleviate the misallocation.

## 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 the overwhelming majority of the system was two lanes or single lane.[3] In addition to the limited lane capacity, more than 25% of national highways were considered to be in poor surface condition.

[1]There are many recent papers that emphasize the misallocation of resources across firms as a source of income difference across countries. Buera, Kaboski, and Shin (2011), Midrigan and Xu (2014), Moll (Forthcoming), Caselli and Gennaioli (2013), Erosa and Allub (2013), and Lopez-Martin (2013) focus on financial frictions. Gourio and Roys (2013), Garicano, Lelarge, and Reenen (2012) and Garcia-Santana and Pijoan-Mas (2014) study the marginal effect of size-dependent policies in France and India respectively. Peters (2013) calibrates a model of imperfect competition with heterogeneous firms to Indonesian data to investigate the impact of misallocation on growth. See Restuccia and Rogerson (2013) and Hopenhayn (2014) for nice surveys of the literature.

[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 decreased over the decades. 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.

[3]Only 3,000 km of the national highway system was four lanes.

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 include 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.

**Geospatial data** We have geospatial data that consists of all the National Highways of India.[4] 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.[5] Figure I shows the evolution of the GQ (in red) in 2001 and 2006.[6] 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 immobile across states.[7]

[4]Data on the Indian National Highway system was supplied by ML Infomap.

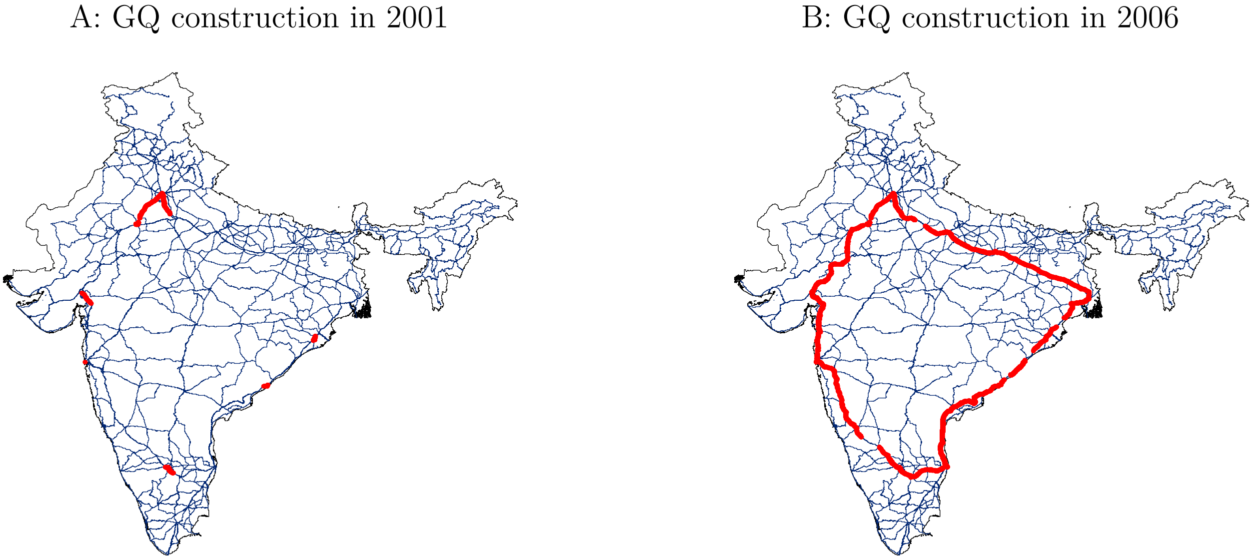
[5]See nhai.org/completed.asp and the Annual Reports of NHAI.

[6]There were seven phases projected in the NHDP. The second phase 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.

[7]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.

# Figure I

Road Network in India and the GQ




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 (around 95% of the total project).

## 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.

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 $\Pi_n$ 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 any 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 faced by that firm. Equations (1), (2), and (3) generate 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 country $n$ 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 country $n$.

Firms within sectors compete à la Cournot. Firm $k$ takes as given the demand characterized by equation (4) and the quantity supplied by competitor firms in the sector 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 $j$ in sector $k$ in state $o$, $\tau_d^o$ 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 of 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 $k$ in sector $j$ in state $d$:

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

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 depend on the market structure of the sector.[8] 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 elasticiy will be equal to $\theta$. 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 $\gamma$. 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:

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

### 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).

[8] In particular, the markup of a firm is an increasing convex function of its market share. This function will be bounded by $\gamma/(\gamma-1)$ and $\theta/(1-\theta)$, which are the smallest and biggest possible markup in the model, respectively.

### 4.5 Misallocation in the Model

Misallocation in this setting arises due to *dispersion* in markups across producers: the marginal revenue product of labor (MPRL) 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.[9] The model is hence relevant to think about the cross-firms misallocation emphasized by Restuccia and Rogerson (2008), Hsieh and Klenow (2009), and Guner, Ventura, and Yi (2008), among others.

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 variation 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 is proportional to the markup charged by that firm.[10] 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), in 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 they would be in the case of perfect competition. Thus, India's aggregate welfare would increase by reallocating labor from firms with low productivity draws (low-markup firms) to firms with high productivity draws (high-markup firms).

## 5 Plant-Level Data on Indian Manufacturing

In this section, we describe the construction of the data set use in the paper. We link firm-level data on the Indian manufacturing sector with geospatial data in order to construct two snapshots in time (2001 and 2006) with detailed manufacturing data and road quality data. The data provides the necessary information to analyze how changes in infrastructure quality affect the manufacturing sector.

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 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.[11]

[9] 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.

[10] It is straightforward to show that the MRPL of a firm operating in state $o$ is $W_o{}^{\epsilon(j,k)}/_{(\epsilon(j,k)-1)}$.

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

This consists of 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: sales, employment, capital stock, wage payments, and expenditures on intermediate goods. The NSS covers all informal establishments in the Indian manufacturing sector. "Informal" refers to all manufacturing enterprises not covered by the ASI. The survey is conducted every five years by the Indian Ministry of Statistics, as one of the modules in the Indian National Sample Survey.

The process of merging the data from the ASI and NSS is straightforward since very similar questions are used to collect the data. Thus, we can create a representative sample of manufacturing plants in India using the weights provided. After merging the ASI and NSS, we have around 190,000 observations for the fiscal year 2000-2001 and 140,000 observations for the fiscal year 2005-2006. Once these observations are properly weighted, each year we have around 17 million manufacturing plants in our data, which employ around 45 million workers.

It is important to note the huge differences in productivity between formal and informal plants in India. Informal plants account for around 80% of employment and around 20% of total value-added.[12] Thus, it is crucial to merge these data sets to have an accurate picture of the Indian manufacturing sector.

**Prices and the consumption of intermediates** The ASI and NSS contain detailed information about production and intermediate good usage. For each plant in our data, we observe the value and physical quantity of production and intermediate input usage broken down by product.[13] This means that we can compute the output prices charged by plants and the input prices paid by plants.[14] 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 around 5,400 different classified products, which are very narrowly defined. For instance, the ASICC distinguishes between different types of black tea: leaf, raw, blended, unblended, dust, etc. In the processed mineral category, for example, the ASICC distinguishes between around 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 $\{\tau_d^o\}_{d=1}^{d=N}$

[12] See the Technical Appendix (1.1) for details.

[13] All plants report intermediate inputs imported from outside India separately from those which are not imported. This is crucial for our analysis, since we abstract from international trade in this paper.

[14] Although these data sets are becoming widely used, not much attention has been paid to the price information. A notable exception is Kothari (2013).

for all $o$, (ii) the elasticity of substitution across sectors $\theta$, (iii) the elasticity of substitution within sectors $\gamma$, (iv) a number of producers for each state-sector $K_{ij}$ for $i$ and $j$, (v) a set of labor endowments $\{L_n\}_{n=1}^{n=N}$ of the states, and (vi) the parameters governing the productivity distribution of the 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 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. Equation (8) shows that the prices charged by firms depend both on transportation costs and market shares in the destination market.[15] In order to identify transportation costs, we exploit one implication of the model: variation in prices for monopolists (i.e. firms with market shares equal to one) are due solely to variation in transportation costs across destinations. To see this formally, equation (8), along with the fact that a monopolist firm faces a demand elasticity given by $\theta$, implies that the firm will charge:

$$p_d^o(j,k) = \frac{\theta}{\theta-1}\frac{W_o}{a_o(j,k)}\tau_d^o. \tag{14}$$

Then, the relative price charged by a monopolist 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, the prices charged by monopolists across states reveal differences in transportation costs.

Empirically, we define a monopolist firm as a plant selling at least 95 percent of the value of each 5-digit ASICC product nationally. Using the ASI and NSS for the years 2001 and 2006, we identify 261 products that are manufactured by monopolists.[16] The largest category is "Manufacture of chemicals and chemical products," which contains around 40 percent of the products identified. This is consistent with the nature of the chemical industry, in which production is often concentrated in one plant due to economies of scale and then shipped to many locations.[17]

[15] Atkin and Donaldson (2014) also issues regarding the fact that dispersion in prices contain information about markups and transportation costs.

[16] We exclude goods that are not used as intermediate inputs in at least five districts.

[17] 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

Once we have identified the products manufactured by monopolists, the strategy is to use the variation in prices across locations where they were used as intermediate inputs to identify transportation costs. We regress variation in prices on a measure of transportation costs that we call effective distance. This measure takes into account the least costly path to go from origin to destination given the road structure. Furthermore, the varying road quality is also incorporated into this measure. The price of an input in each district is computed as the weighted average of the prices paid by all the plants that use that input in that district.[18]

We estimate equation (14) as follows:

$$\log p^{o}_{d,t}(j,k) = \beta \log \text{ Effective Distance}^{o}_{d,t} + \sum_{o} \delta_o + \sum_{j} \alpha_j + \sum_{t} \eta_t + \epsilon^{o}_{d,t}(j,k) \tag{15}$$

where $p^{o}_{d}(j,k)$ is the average price in district $d$ paid for product $j$ produced by a monopolist located in district $o$, $\delta_o$ are a set of districts of origin fixed effects, $\alpha_j$ a set of product fixed effects, $\eta_t$ are time dummies, 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.[19]

In order to compute the 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 national road that connects them. An arc is referred to as being GQ or non-GQ, depending on whether it was completed in the specific year.[20] We then use Dijkstra's shortest-path algorithm to construct a matrix of lowest-cost distances between all the districts for the years 2001 and 2006. The transportation costs in these two years are different since this algorithm takes into account the fact that traveling on a better quality road (i.e. across the Golden Quadrilateral) is less costly. Specifically, we assume that:

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

where $n_1$ and $n_2$ are nodes, and $\alpha$ 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.[21] This value of $\alpha$ indicates that if a given stretch is GQ, the effective distance

[18]Some districts between 2001 and 2006 are split in one or several new districts. We take into account those splits and construct a set of districts that is homogeneous between 2001 and 2006.

[19]Note that, although we are calibrating the model to the year 2006, we exploit the cross-sectional variation using the two years in our sample to estimate the relationship between prices and effective distance. We proceed in this way in order to have a bigger power in our estimations of transportation costs.

[20]The National Highways Authority of India (NHAI) provides information on the start and completion date for all the stretches of the GQ. See the Technical Appendix for details.

[21]The value of $\alpha$ is based on the fact that the average speed on a national highway is between 30 and 40 km/h

is roughly half of what it is if it is not GQ. The effective distance used to estimate equation (15) is the sum of the effective distance along all the arcs traveled along the shortest path.

Table I presents the results from estimating equation (15). In column (1), we show that a 10 percent increase in the effective distance is associated with a 0.86 percent increase in the price of the good.[22] In column (2), we use a more flexible specification, in order to incorporate potential non-linearities in transportation costs.[23] We include ten deciles of log effective distance, and find that the highest deciles are associated with large increases in the price of the good. We find, for instance, that the prices paid at destinations falling in the second decile of effective distance (around 280 km) are 37% higher than the prices paid at destinations within the first decile (70 km on average). The effect is particularly strong for destinations that are very far from the location where production takes place: the prices are around 52% higher when the effective distance to the destination is in the 10th percentile of the distribution. The 10th decile includes districts located more than 1,800 kilometers away in effective distance, which is roughly the road distance from New York City to Des Moines, Iowa. Although the overall pattern is increasing, the effect seems to be non-monotonic. For example, the coefficient associated with the third decile is 8 percentage points lower than the second decile coefficient. In order to avoid having non-monotonic transportation costs to effective distance in the model, we assume that the relationship between iceberg costs and effective distance is given by a discrete cubic function $g(\text{Coeff. of Effective Distance}_d^o)$, and set the parameters that better fit the coefficients implied by the regression.[24] Lastly, we assume that the iceberg cost for all destinations in the first decile is equal to one. The iceberg cost predicted for all other deciles becomes:

$$\hat{\tau}_d^o = e^{\mathrm{g}(\text{Coeff. of Effective Distance}_d^o)}. \tag{17}$$

**Direct measures of transportation prices** In order to have additional estimates of the transportation costs across Indian locations, we have assembled an additional dataset that contains information on prices charged for shipping a container by truck within India. In particular, we have collected pricing data for shipping a container of size 20 ft x 8 ft x 8.5 ft for around 900,000 origin-destination Indian city pairs. Using this dataset we construct measures of bilateral iceberg costs as a function of effective distance and compare them with the ones we obtain from Equation

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.

[22]Costinot and Donaldson (2012) estimate a similar regression for the price of agricultural goods and their distance to the nearest wholesale market over time in the United States. They find coefficients for distance of a similar magnitude during the 1880-1920 period (0.09 to 0.14). Note that the effective distance is exactly equal to the road distance before the construction of the GQ.

[23]This flexible specification is commonly used to estimate the parameters of trade models using gravity equations. Examples include Eaton and Kortum (2002) and Waugh (2010).

[24]See Section (2) of the Technical Appendix for details.

TABLE I

IMPACT OF ROAD DISTANCE AND INFRASTRUCTURE QUALITY ON PRICES

| | (1) | (2) |
|---|---|---|
| Dep. Variable: Log price at district of destination | | |
| Log Effective Distance | 0.086*** | |
| | (0.023) | |
| Log Effective Distance $2^{nd}$ decile | | 0.371*** |
| | | (0.115) |
| Log Effective Distance $3^{th}$ decile | | 0.298*** |
| | | (0.114) |
| Log Effective Distance $4^{th}$ decile | | 0.137 |
| | | (0.112) |
| Log Effective Distance $5^{th}$ decile | | 0.168 |
| | | (0.131) |
| Log Effective Distance $6^{th}$ decile | | 0.398*** |
| | | (0.121) |
| Log Effective Distance $7^{th}$ decile | | 0.355*** |
| | | (0.133) |
| Log Effective Distance $8^{th}$ decile | | 0.445*** |
| | | (0.142) |
| Log Effective Distance $9^{th}$ decile | | 0.341** |
| | | (0.141) |
| Log Effective Distance $10^{th}$ decile | | 0.516*** |
| | | (0.136) |
| District of Origin Fixed Effects | YES | YES |
| Product Fixed Effects | YES | YES |
| Year Fixed Effects | YES | YES |
| Observations | 2,235 | 2,235 |
| R-squared | 0.876 | 0.881 |

Table I shows the estimation of equation (15). 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, relatives 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. Column (1) uses a linear specification of effective distance, whereas column (2) estimates a non-linear specification, using 10 deciles of effective distance. District of origin, product and year -2001 and 2006- fixed effects are included. Robust standard errors are in parenthesis. Significance levels: *: 10%; **: 5%; ***: 1%.

(17).[25] The overall patern of the two sets of iceberg costs is remarkably similar. Importantly, in

[25]See Section (3) of the Technical Appendix for further details.

both cases, transportation costs increase more than linearly at higher levels of the distribution of effective distance.

**What do transportation costs look like?** As a starting point, we will take the district of New Delhi (located in the National Capital Territory of Delhi) in the year 2001. Panel A of Figure II shows a map of the transportation costs to all districts from New Delhi. The legend on the map shows transportation costs divided into quartiles. The figure also shows that only a small portion of the GQ was upgraded by this point (depicted in red). The first thing to notice is the concentric circles that surround New Delhi. This means that the further the destination, the higher the transportation costs. The concentric circles also show that straight-line distances are highly correlated to the shortest path on the highway system. The reason is that the highway system is dense, as can be seen in Figure I. The second thing to notice is the general level of transportation costs. The map shows iceberg costs of 1.43-1.50 for transporting goods from New Delhi to the southern tip of India.

Our next step is to look at transportation costs from New Delhi in the year 2006 (panel B of Figure II), after a large part of the upgrade of the GQ had been completed. The color categories for the map have not changed from panel A, so that the colors are comparable across maps. The lighter colors reflect a general decrease in transportation costs.

### 6.2 Estimating the Across-Sector Elasticity of Substitution ($\theta$)

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

Formally, we derive a gravity equation implied by the model for the trade flows of monopolist firms. Combining equations (4) and (14), we derive the following condition for the trade flow values:

$$\begin{aligned}\log c_d^o(j,k)p_d^o(j,k) =& (1-\theta)\log W_o + (\theta-1)\log a_o(j,k) + \log P_d^\theta Y_d \\ &+ (1-\theta)\log\tau_d^o + (1-\theta)\log\frac{\theta-1}{\theta}.\end{aligned} \tag{18}$$

The model predicts that higher transportation costs reduce trade flows, and the strength of this relationship depends on the value of $\theta$. 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 $\theta$ 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 (18) as follows:

$$\log \text{Sales}_{d,t}^o(j,k) = \beta\log\hat{\tau}_{d,t}^o + \sum_o \delta_o + \sum_j \alpha_j + \sum_d \lambda_d + \sum_t \eta_t + \epsilon_{d,t}^o(j,k) \tag{19}$$

where $\text{Sales}^o_{d,t}(j,k)$ is the value of sales of product $j$ in year $t$ consumed in district $d$ and produced by a monopolist located in district $o$, $\hat{\tau}^o_d$ is the predicted iceberg transportation cost between districts $o$ and $d$ (obtained from equation (17)), $\delta_o$ is a set of districts of origin fixed effects, $\alpha_j$ is a set of product fixed effects, $\lambda_d$ is a set of districts of destination fixed effects, $\eta_t$ is a set of year fixed effects, and $\epsilon^o_{d,t}(j,k)$ is the error term. The origin fixed effect controls for local wages. The product fixed effect controls for firm productivity. The destination fixed effect controls for market size and aggregate prices at the destination.

Table II presents the results of estimating equation (19). We find that higher transportation costs are associated with lower trade flows at statistically significant levels. The empirical specification indicates that transportation costs which increase by 10 percent are associated with an 8.3 percent decrease in trade flows. This relationship implies that the value of $\theta$ is 1.83.

TABLE II

GRAVITY EQUATIONS FOR MONOPOLISTS

| | (1) |
|---|---|
| Dep. Variable: Log value of sales at destination | |
| $\hat{\tau}^o_d$ | -0.840** |
| | (0.401) |
| District of Origin Fixed Effects | YES |
| District of Destination Fixed Effects | YES |
| Product Fixed Effects | YES |
| Year Fixed Effects | YES |
| Observations | 2,235 |
| R-squared | 0.538 |

Table II shows the estimates of equation (19). 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 (15), namely the predicted transport costs across districts. Origin, destination, product and year fixed effects are included. Product fixed effects correspond to 5-digit ASICC products. Robust standard errors are in parenthesis. Significance levels: *: 10%; **: 5%; ***: 1%.

### 6.3 Estimating the Within-Sector Elasticity of Substitution ($\gamma$)

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:

$$\frac{W_o l^o_d(j,k)}{\tilde{p}^o_d(j,k)c^o_d(j,k)} = 1 - \frac{1}{\gamma} - \left(\frac{1}{\theta} - \frac{1}{\gamma}\right)\omega^o_d(j,k) \tag{20}$$

where $\tilde{p}_d^o(j,k)$ is the factory gate price of the good.[26] 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.

In the data, we do not observe the market share of any given firm by destination. However, a similar condition can be derived for goods that are only produced in one state. In these sectors, the market shares of firms are constant across destinations.

We find that approximately 15% of sectors are operated only in one state. These sectors comprise 30,000 firms. Using data from these firms, we estimate equation (20) as follows:

$$LS_o(j,k) = \beta\omega^o(j,k) + \sum_o \delta_o + \sum_j \alpha_j + \epsilon^o(j,k) \tag{21}$$

where $LS_o(j,k)$ and $\omega^o(j,k)$ are the labor and sectoral shares respectively in state $o$, $\delta_o$ is a set of fixed effects to control for the state where the firm operates, $\alpha_j$ is a set of product fixed effects, and $\epsilon^o(j,k)$ is the error term.

TABLE III

LABOR SHARES VS SECTORAL SHARES

| | (1) Labor Share | (2) Cap+Labor Share |
|---|---|---|
| Sectoral Share | -0.416*** | -0.493*** |
| | (0.094) | (0.105) |
| Constant | 0.707*** | 0.807*** |
| | (0.042) | (0.052) |
| Sector FE | YES | YES |
| Year FE | YES | YES |
| Observations | 1,181 | 1,009 |
| R-squared | 0.870 | 0.893 |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Column (1) of table III shows an OLS regression of firms' labor shares against sectoral shares for sectors that are operated only in one state. Column (2) shows the same regression but including also capital remuneration on the left hand side. Product fixed effects correspond to 5-digit Indian sectoral codes (ASICC). Robust standard errors are in parenthesis: *: 10%; **: 5%; ***: 1%.

We present the results in Table III. Column (1) shows the results when including only labor remuneration on the left-hand side of the equation. In column (2), we also include capital remuneration on the left-hand side of the equation. The second specification controls for across-firm variations in capital intensity. We choose this second specification as our preferred one. An OLS slope coefficient of -0.49 together with an across-sector elasticity of substitution $\theta$ of 1.84 implies a value of $\gamma$ of 19.77.

[26] The factory gate price is the price of the good at the origin. In the data, we computed the factory gate price by dividing a plant's sales by the physical units. See the Technical Appendix (4) for the details of the derivation.

### 6.4 Aggregating Transportation Costs to the State Level

In order to exploit all the variations that exist in the data, we use district-level data in the estimates of transportation costs, $\theta$, and $\gamma$. It is necessary to aggregate the district-to-district transportation costs to state-to-state transportation costs since the model that we simulate is based on interstate trade. We do so in two steps. In the first step, for every district we find the average transportation cost to the districts located in a given destination state. This average is weighted by the population of the destination districts. This yields a measure of district-to-state transportation costs. In the second step, we aggregate the district-to-state measure to obtain state-to-state transportation costs. To do so, we find the average transportation cost from the origin districts of the origin state to a given destination state. This average is weighted by the population of the origin districts.

Given this new set of transportation costs, we repeat the exercise above in which we map the transportation costs from the National Capital Territory of Delhi to all of the states in India. Panel C of Figure II shows a map of these transportation costs. The pattern of faraway states having higher transportation costs that we observed at the district level is also visible in this figure. Panel D of Figure II depicts transportation costs in 2006. The fact that the colors are lighter means that there is a decline in transportation costs to most regions.

Importantly, there is a high variation in the decline of transportation costs across locations. As an illustration, Figure III shows the percentage decline in transportation costs from Delhi. As in the previous figure, the colors of the states represent the quartile in terms of decline in transportation costs. States in the top quartile tend to be far away from Delhi and close to the GQ upgrades. The states in the top quartile underwent a decline of 3.9 to 8.7%. The states with the smallest decline in transportation costs are the ones that are far from the GQ upgrades. For example, the northern state of Jammu and Kashmir and the northeastern states of Arunachal Pradesh, Assam, Manipur, Tripura, and Mizoram. The percentage decline in transportation costs for the bottom quartile ranges from 0 to 0.59%.

### 6.5 Calibrating the Remaining Parameters

**Labor endowment** For the labor endowments of each state, $L_n$, we first normalize the labor endowment of the smallest state to 1. We then set the labor endowments of the remaining states so that the model matches the ratio of manufacturing value-added observed in the data across states.

**Parameters that govern within-industry productivity across states** We will now calibrate the parameters that relate to the number of firms that operate and the productivity distribution. These parameters are crucial for the size of the Ricardian and pro-competitive effects of reducing transportation costs.

The number of firms in sector $j$ of country $n$, $K_{nj}$, is set to match the number of plants observed

## TABLE IV

### PARAMETER VALUES

| Param. | Definition | Value |
|---|---|---|
| (A) PARAMETERS ESTIMATED WITH STRUCTURAL EQUATIONS | | |
| $\tau_d^o$ | Iceberg transportation costs between states | varies by state pair |
| $\theta$ | Elasticity of substitution across sectors | 1.83 |
| $\gamma$ | Elasticity of substitution within sector | 19.77 |
| (B) PARAMETERS TAKEN DIRECTLY FROM DATA | | |
| $K_{ij}$ | Number of firms operating in sector $j$ of country $i$ | varies by state |
| (C) PARAMETERS CALIBRATED IN EQUILIBRIUM | | |
| $L_i$ | Labor endowment of the states | varies by state |
| $\alpha$ | Shape parameter Pareto | 2.05 |

Notes: Table IV refers to a calibration in which productivity draws across states are independent.

in the data. Since there is no operating cost in the model, all firms operate and there is no entry and exit of firms even after changes in transportation costs. Abstracting from firm entry and exit in these kinds of models does not quantitatively affect the final results. The reason is that a reduction in transportation costs will lead to the entry and exit of low-productivity firms. These firms do not significantly affect the markups that large firms charge. This is consistent with the findings of Atkeson and Burstein (2008) and Edmond, Midrigan, and Xu (2012). Furthermore, the data does not show a significant change in terms of the firms across sectors in each state. For example, the auto-correlation of the number of producers per sector-state between 2001 and 2006 is 0.98.[27]

The distribution of the number of firms across state-sectors is important in determining the nature of gains from lower transportation costs. As a simple example to illustrate this idea, consider a two-state example. Suppose that these two states go from autarky to trading with each other. If there is no overlap in the sectors that these two states produce in, the effects from trade will be purely Ricardian. This is true since trading with another state will not change markups. However, if two states produce very similar goods, then there is room for pro-competitive effects from trade.

Another important factor to consider is the correlation of productivity draws across regions. The correlation determines the extent to which local firms with market power face new competition

[27] The number of active sectors across states remained stable over this period. The change in the percentage of active sectors within states is around 3% on average. The total number of firms did not vary significantly either. The percentage change in the total number of firms within states is around 2% on average.

when the economy opens to trade. Edmond, Midrigan, and Xu (2012) show that the correlation in productivity draws is important to determine the size of pro-competitive gains from trade. In a situation in which productivity draws across states are independent, the pro-competitive gains from trade are zero or even negative. Furthermore, they pin down a very high correlation (0.93) for the Taiwanese case. We assume that firms across states, conditional on being operating in a given sector, have perfectly correlated draws in our benchmark calibration. Importantly, note that this does not imply that cross-state productivity is perfectly correlated within a given sector. The reason is that, as mentioned above, the number of firms operating and the number of active sectors vary across states.

We use a Pareto distribution for the productivity draws. This is a commonly used distribution in the trade literature. The tail parameter, $\alpha$, is calibrated in equilibrium to match the fact that the top 5% of firms in manufacturing value-added account for 89% of value-added in this sector.

# 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). Lastly, we use a difference-in-difference strategy to estimate the decline in prices for districts close to the GQ compared to those that are further away. We compare these results with the predictions of the model.

## 7.1 Simulating the Construction of the GQ

In order to quantify the effects of the GQ, we begin with our baseline calibration described in Table IV. We change the transportation costs to reflect the absence of the GQ. To do so, we change the cost to travel on roads that were upgraded by the GQ as described in equation (16). Given these new costs, we re-compute the shortest path using Dijkstra's algorithm. Finally, we re-aggregate the district-to-district transportation costs to state-to-state transportation costs as described in Section 6.4.

**Benefits from the GQ** First, we consider the aggregate change in real income resulting from the GQ. Table V shows that real income increases by 2.04% for India. Changes in aggregate real income are calculated as the mean percentage change of all states weighted by real income. The increase in real income is in terms of the manufacturing value-added, since this is the only sector considered in our model. The value-added of the manufacturing sector was $152.8 billion in 2006.

This implies that the static benefit of the construction of the GQ is \$3.1 billion. These are the benefits that accrue to India *each year* as a result of the construction of the GQ. We can compare these benefits to the cost of the construction of the GQ. Estimates indicate that the government spent approximately \$5.6 billion on the GQ. Thus, the benefits over a two-year period exceed the initial construction costs.

**A framework to decompose the Ricardian and pro-competitive effects of the GQ** We apply the framework developed by Holmes, Hsu, and Lee (Forthcoming) to decompose the changes in real income in a way that highlight the various mechanisms at work in the model. The framework allows us in particular to distinguish between Ricardian, pro-competitive, and terms of trade effects from lowering transportation costs.

We now introduce some notations for the purpose of the decomposition. We define the aggregate markups on the goods sold. This reflects how much market power firms producing in a state have when selling to other states. First, the revenue-weighted mean labor cost 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 income at $d$ that is spent on the goods produced by firm $j$ in sector $k$ from state $n$. 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 country's total income.[28]

We next define the aggregate markups on the goods purchased by state $n$, which reflects 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:

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

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:

[28]The analogous expression at the firm level is that the firm's markup is equal to the reciprocal of the labor share.

$$Y_n = \underbrace{W_n L_n}_{\text{Labor income}} * \underbrace{\frac{1}{P_n^{pc}}}_{\text{Prod. efficiency}} * \underbrace{\frac{P_n^{pc}}{P_n}\mu_n^{buy}}_{\text{Allocative efficiency}} * \underbrace{\frac{\mu_n^{sell}}{\mu_n^{buy}}}_{\text{Markup ToT}}. \tag{22}$$

The first component is the aggregate *labor income*. The second component is the *productive efficiency* component of welfare. The component is simply the inverse of the price index if all firms charge the marginal cost. The third component is *allocative efficiency*. 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. In a situation with no misallocation, i.e. no variations on markups across firms, this index is equal to one. As misallocation increases, this index decreases. The last component is *terms of trade*. This component compares the aggregate markup charged for the goods a country sells with those that it purchases.

Combining the first two terms leads to an expression that is equal to real income if firms charge the 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. We consider changes in the allocative efficiency to be pro-competitive effects as this directly maps back to the welfare losses due to dispersion in markups. Given the expression in equation (22), we decompose the changes in real income into the following terms:

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

**Quantifying the decomposition** Table V shows these three components at the aggregate and state level. We find that, for India as a whole, the pro-competitive component accounts for approximately 19% of the aggregate gains (0.39% of the 2.04% total gains). The pro-competitive component can be up to 29% of the gains at the state level (0.46% of the 1.61% of the gains for Maharashtra).

The welfare effects of the GQ are very heterogeneous across states. Overall, large states gain more from the reduction in transportation costs. Small states see modest gains and in some cases even lose. This is driven by the fact that, due to its placement, the GQ has lowered transportation costs primarily for large states. Many of the small states are located in northeastern India, which is far from the GQ.[29] The states in the Northeast that gained less than 1% include: Assam and Meghalaya. The states in the Northeast that experienced losses include: Arunachal Pradesh,

[29] Northeastern Indian states include: Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim, and Tripura.

Manipur, Mizoram, Nagaland, and Tripura. Figure IV shows a map of the welfare effects across states, including the states that lost.

## Table V

### Quantitative Results

| state | size | income change | Descomposition | | | |
|---|---|---|---|---|---|---|
| | | | $\eta_w$ | $\eta_{PE}$ | $\eta_{ToT}$ | $\eta_{ae}$ |
| India | | 2.04 | 1.47 | -0.46 | 0.00 | 0.39 |
| Maharashtra | 100.00 | 1.61 | 1.94 | -0.96 | 0.17 | 0.46 |
| Gujarat | 65.40 | 2.74 | 2.60 | -0.48 | 0.08 | 0.55 |
| Tamil Nadu | 41.67 | 1.32 | 1.81 | -0.79 | -0.01 | 0.32 |
| Uttar Pradesh | 29.91 | 2.22 | 2.14 | -0.36 | -0.07 | 0.51 |
| Karnataka | 26.97 | 2.87 | 2.61 | -0.21 | -0.10 | 0.57 |
| Andhra Pradesh | 21.23 | 1.78 | 1.84 | -0.34 | -0.11 | 0.39 |
| West Bengal | 18.97 | 3.67 | 3.04 | 0.16 | -0.14 | 0.62 |
| Haryana | 18.80 | 0.97 | 1.55 | -0.62 | -0.17 | 0.21 |
| Jharkhand | 17.10 | 4.97 | 3.44 | 1.22 | 0.03 | 0.29 |
| Rajasthan | 12.64 | 2.84 | 2.46 | 0.43 | -0.30 | 0.25 |
| Madhya Pradesh | 11.39 | 1.40 | 1.55 | -0.25 | -0.11 | 0.20 |
| Orissa | 10.74 | 2.31 | 2.06 | -0.02 | 0.01 | 0.26 |
| Punjab | 10.16 | 0.41 | 0.95 | -0.69 | 0.04 | 0.10 |
| Himachal Pradesh | 9.49 | 0.29 | 1.07 | -0.70 | -0.17 | 0.09 |
| Chattisgarh | 9.32 | 0.28 | 1.02 | -0.84 | -0.01 | 0.11 |
| Kerala | 7.42 | 0.75 | 1.07 | -0.60 | 0.06 | 0.22 |
| Uttaranchal | 4.62 | 0.76 | 1.15 | -0.57 | 0.04 | 0.14 |
| Delhi | 4.58 | 0.95 | 1.15 | -0.53 | 0.12 | 0.21 |
| Assam | 3.72 | 0.28 | 0.86 | -0.72 | 0.04 | 0.09 |
| Goa | 3.57 | 7.82 | 4.89 | 3.51 | -0.60 | 0.02 |
| Bihar | 2.73 | 5.35 | 3.54 | 2.11 | -0.37 | 0.07 |
| Jammu and Kashmir | 2.59 | -0.41 | 0.47 | -0.74 | -0.11 | -0.04 |
| Meghalaya | 0.62 | 0.53 | 0.99 | -0.67 | 0.07 | 0.13 |
| Tripura | 0.26 | -1.60 | -0.25 | -1.63 | 0.25 | 0.03 |
| Manipur | 0.11 | -1.51 | -0.20 | -1.70 | 0.34 | 0.05 |
| Nagaland | 0.08 | -1.06 | 0.06 | -1.45 | 0.26 | 0.07 |
| Sikkim | 0.03 | 3.96 | 2.47 | 0.89 | 0.39 | 0.22 |
| Mizoram | 0.02 | -1.07 | 0.38 | -1.68 | 0.19 | 0.05 |
| Arunachal Pradesh | 0.01 | -1.29 | 0.00 | -1.80 | 0.46 | 0.05 |

Table V shows the % change in real income and the decomposition of the Holmes, Hsu, and Lee (Forthcoming) index for the 29 Indian states; $\eta_w$ represents the % change in labor income component of the index; $\eta_{PE}$ represents the % change in the *Prod. efficiency* component; $\eta_{ToT}$ represents the % change in the *terms of trade* component; and $\eta_{ae}$ represents the % change in the *allocative efficiency* component.

Next, we examine the pro-competitive and terms of trade effects across states. These two terms are the result of the variable markup feature of the model. First, we find that the pro-competitive effects are positive across all states. This means that lower transportation has led to welfare-enhancing changes in markups. As mentioned before, theory is ambiguous as to whether declines in transportation costs lead to gains in allocative efficiency. The range of gains from changes in allocative efficiency is -0.04 to 0.62%. Larger states also see the greatest gains in terms of allocative efficiency.

Secondly, there is a wide dispersion in the effects of the terms of trade component. For example, Goa lost 0.60%, while Sikkim gained 0.39%. Thus, although allocative efficiency improves for all but one state, the changes in the terms of trade can result in some states suffering losses due to changing markups. For example, Goa lost more from the changing terms of trade than it did from the improved allocative efficiency. Thus, although almost all states gain from increases in the allocative efficiency, the changes in markups lead to a significant re-shuffling of income across states through changes in the terms of trade.

Next, we turn to the Ricardian components across states. These terms are generally positive and large across all states. This term also explains the modest or negative effects for the states in the Northeast. The only two factors that affect a firm's marginal cost to serve a destination are the transportation costs that it faces and the wages that it pays its workers. First, we know that the GQ lowers transportation costs for some destinations and leaves the transportation costs for others unchanged. Thus, changes in transportation costs increase the productive efficiency component. Destinations closer to the GQ benefited from a higher increase in productive efficiency. The fact that the Ricardian term is negative for the Northeastern states implies that the effect of wages in other states outweighed the benefits of the GQ in terms of lower transportation costs. Indeed, we find that there is a general rise in wages across almost all states.

### 7.2 The GQ and Trade Diversion and Creation within India

We now study 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 gives room for trade diversion/creation. To study this possibility, we compute the whole set of bilateral trade flows across Indian states before and after the GQ. We define total trade between state $i$ and state $j$ as

$$\text{Total Trade}_{i,j} = \text{exports}_j^i + \text{exports}_i^j \tag{23}$$

where $\text{exports}_j^i$ and $\text{exports}_i^j$ are the total exports from state $i$ to state $j$ and the total exports from state $j$ to state $i$ respectively.

Table VI shows the percentage changes of this variable for different sets of state pairs. In Panel

## TABLE VI

### % CHANGE IN TOTAL TRADE BETWEEN $i$ AND $j$ (MODEL)

| | | mean | median | sd/mean | N |
|---|---|---|---|---|---|
| | both $i$ and $j$ in GQ | 5.35 | 5.03 | 1.55 | 78 |
| (A) | either $i$ or $j$ in GQ | 1.42 | -0.79 | 7.11 | 208 |
| | neither $i$ nor $j$ in GQ | 3.25 | 3.11 | 3.54 | 120 |
| | both $i$ and $j$ in (N-EAST+JM) | 3.89 | 2.73 | 2.40 | 190 |
| (B) | either $i$ or $j$ in (N-EAST+JM) | 1.21 | 0.31 | 9.72 | 180 |
| | neither $i$ nor $j$ in (N-EAST+JM) | 4.08 | 4.04 | 1.47 | 36 |

Table VI 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; "both $i$ and $j$ in (N-EAST+JM)" refers to state pairs in which both of them belong to the group (N-EAST+JM); "either $i$ or $j$ in (N-EAST+JM)" refers to state pairs in which one of them belong to the group (N-EAST+JM); "neither $i$ nor $j$ in (N-EAST+JM)" refers to state pairs in which none of them belong to the group (N-EAST+JM). N-EAST+JM includes Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim, Tripura, and Jammu and Kashmir. N is the number of state pairs that fall into the six different categories.

(A) we have divided the state pairs according to their access to the GQ.[30] In the first row we include state pairs in which both of the 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 none of the states is crossed by the GQ. The Panel (B) is equivalent to Panel (A), but classifying the state pairs according to whether they are way far from the GQ. This set of states include the northeastern states (Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim, and Tripura) plus Jammu and Kashmir. With this additional classification, we try to capture the fact that, even if some states were not crossed by the GQ, they were relatively close so that their market access increased.[31]

We find that, on average, trade increases considerably more between state pairs in which both states are cross by the GQ or not crossed by the GQ. In other words, the construction of the GQ resulted in an increase in trade flows between states located far away and between states located close to the GQ. For instance, trade flows between state pairs crossed by the GQ increased on average a 5.35%. For state pairs not crossed by the GQ, the increase in trade was of 3.25%. On the other hand, trade between state pairs in which only one of the states was crossed by the GQ increased much less, 1.40% on average. In fact, the median percentage change in trade between these states is negative: -0.79%.

This means that the trade flows in India becomes more regionally concentrated. For instance, the trade between northeastern states increased by 3.89%. The pattern is similar for the non-northeastern states, where trade increased by 4.08% on average.

[30] This analysis is similar in spirit to Krueger (1999) and Bayoumi and Eichengreen (1998).

[31] Good examples of these states are Goa and Madhya Pradesh.

### 7.3 The GQ and the Evolution of Prices

The model has an important implication which is behind the differential effect of the construction of the GQ across regions: prices fall down more in states crossed by the GQ. We use the time dimension of the data to evaluate the ability which the model has to predict the different responses of prices in GQ vs non-GQ locations. In order to do this, we examine the impact of the GQ on prices using a reduced form approach and we compare it with the outcome of the model.[32] In the model, the prices paid for goods are endogenous and depend on both changes in transportation costs, wages, and markups. These changes in wages and markups depend on complex general equilibrium effects. Thus, it is necessary to simulate the model in order to compare the changes in prices paid for goods with the data.

One of the major issues to tackle when studying the impact of transportation infrastructure is the fact that the placement of infrastructure is not random. In the case of India, the GQ was built with the goal of connecting the major urban centers. In order to deal with this identification problem, we use the strategy adopted by several authors such as Atack, Bateman, Haines, and Margo (2010) and Banerjee, Duflo, and Qian (2012) who have exploited the fact that the goal of infrastructure projects is usually to connect historical cities or large economic centers. The causal effect of transportation infrastructure is identified by applying a difference-in-difference approach comparing non-nodal areas that differ in their distance to the transportation network before and after the infrastructure is constructed. We follow this approach in order to study the impact of transportation costs on prices, making use of the natural experiment provided by the GQ. We run the following difference-in-difference regression in particular:

$$\Delta \log P_{jd} = \sum_j \alpha_j + \beta_1 \Delta \mathrm{GQ}_d + \sum_s \delta_s + \epsilon_{jd}, \tag{24}$$

where $P_{jd}$ is the price of input $j$ in district $d$ between 2001 and 2006, and $\mathrm{GQ}_d$ is a dummy variable taking the value 1 if district $d$ is within a certain distance of the GQ, and $\epsilon_{jd}$ is an error term.[33] Thus, $\Delta\mathrm{GQ}_d = 1$ if a district is within the specified distance of a treated portion of the GQ in 2006 and not in 2001. We use the following categories for distance: 15, 50, 100, 150, 200, and 300 km. We include input fixed effects and state fixed effects in order to account for input-specific price trends and aggregate shocks affecting prices at the state level. Standard errors are clustered at the district level in order to account for possible serial correlation of price shocks within districts.

The estimates of equation (24) can be found in Table VII. We find that districts located within

[32] The difference-in-difference strategy we use to identify how the GQ affected prices is similar to the one used in recent work that investigates the impact of transportation infrastructure in development. These papers include Atack, Bateman, Haines, and Margo (2010), Banerjee, Duflo, and Qian (2012), and Faber (Forthcoming). This strategy has also been applied to investigate the impact of the GQ by Datta (2012) and Ghani, Goswami, and Kerr (2013).

[33] Distance is calculated as the shortest straight-line distance between the district and a treated portion of the GQ.

# Table VII

## Prices and the Golden Quadrilateral: Differences-in-Differences

| | (1) | (2) |
|---|---|---|
| Dep. Variable: Log change in input prices between 2001 and 2006 | | |
| District within 15 km from GQ | -0.3219** | -0.3293** |
| | (0.1395) | (0.1406) |
| District within 50 km from GQ | -0.3484** | -0.3604*** |
| | (0.1363) | (0.1367) |
| District within 100 km from GQ | -0.2036 | -0.2171 |
| | (0.1659) | (0.1697) |
| District within 150 km from GQ | 0.0711 | 0.0768 |
| | (0.1357) | (0.1416) |
| District within 200 km from GQ | 0.0916 | 0.0973 |
| | (0.1591) | (0.1767) |
| Input fixed-effects | YES | YES |
| State fixed-effects | YES | YES |
| Nodal districts | YES | NO |
| Observations | 5,123 | 5,037 |
| Average R-squared | 0.44 | 0.44 |
| Number of products | 929 | 912 |

Table VII shows the estimation of equation (24). 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. The treatment variable at distance x takes value 1 if district $d$ is within x km from the GQ in 2006 and was not in 2001, and zero otherwise. Columns (1) includes all districts whereas column (2) excludes nodal districts. Input and state fixed effects are included in all specifications. Robust standard errors are in parenthesis, clusterd at the district level. Significance levels: *: 10%; **: 5%; ***: 1%.

15 km and 50 km of the GQ in 2006 experienced statistically significant declines in input prices. For districts located within 15 km, input prices were 33 percentage points lower relative to districts located further away from the GQ. The first coefficient of column (1) includes nodal districts and column (2) excludes nodal districts. For districts within 50 km of the GQ in 2006, we find an even stronger effect, a decrease of 36 percentage points in input prices relative to districts further away.[34] This implies that prices in "GQ" districts decreased 1.57 times as much as in the average district. The model predicts a stronger effect. The decrease in prices charged in states through which the GQ passes are 2.70 times bigger than in the average state. Although we find this evidence supportive for the differential evolution in prices predicted by the model, these comparisons have to be taken with caution for several reasons. First, while we calibrate the model at the state level, we carry out the difference-in-difference exercise at the district level. We do this in order to exploit as much as possible the variation in the data. Second, we are looking at a short span of time in the data (2001-2006). However, one could think that it would take longer for the Indian economy

[34] In the data, extending the treatment group beyond 50 km makes the effect disappear. The evolution of input prices was not significantly different between districts that within 100 km of the GQ and those beyond 100 km.

to converge to the new equilibrium after the construction of the GQ.

# 8 Alternative Scenarios

In this section, we evaluate the aggregate and state-level welfare under various scenarios.

## 8.1 Perfect Competition

We first examine the implications of changing the market structure to perfect competition for all firms. Under perfect competition, there is no misallocation or dispersion in MRPL across firms.

The allocative efficiency component of welfare gives us a sense of the welfare losses due to the misallocation resulting from market power. The allocative efficiency component ranges from 0.967 to 0.925, meaning that real income would increase by 3.3-7.5% under marginal cost pricing. Furthermore, we find that larger states consistently have more misallocation than smaller states. Overall, the levels of misallocation that the model generates are low compared to the ones found by Hsieh and Klenow (2009).

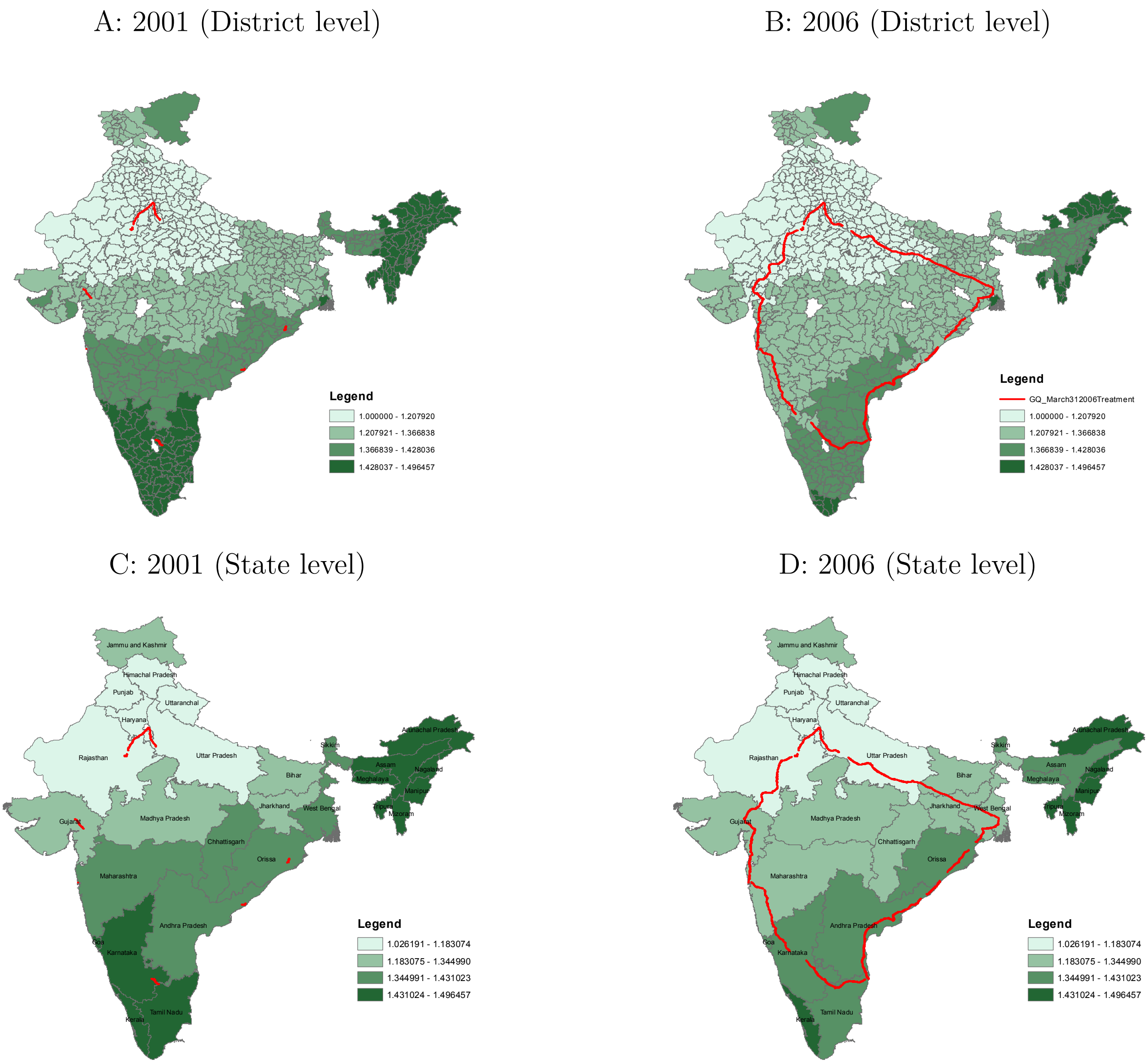
To understand the quantitative importance of market power on firm size distribution, we simulate an equilibrium in which all firms charge marginal cost. In the new equilibrium, all parameters remain the same except firms charge marginal cost. Since productive firms charge higher markups in the calibration, they will be larger in the new equilibrium with perfect competition. We find that the top 5% of firms comprise 97% of all sales in India compared to 89% previously. In addition, the top 1% of firms go from having 42% to 55% of sales. In the case of perfect competition, there is no variance in the MRPL.

# 9 Conclusions

The goal of this paper is to quantitatively evaluate the welfare effects of improving the transportation infrastructure in a setting of internal trade and variable markups. Hence, we determine the extent to which misallocation is driven by high transports costs and decompose the welfare effects into Ricardian and pro-competitive gains, and we can thus gauge the distribution of gains across locations. We apply this framework to the construction of the Golden Quadrilateral in India, a major highway project spanning 5,800 km. We find large gains from the infrastructure project, amounting to more than 2% of real income. Nevertheless, there is wide variation in income gains across Indian states, and even some negative effects after the project. Those locations closer to the GQ reap the main benefits, whereas states further away suffer from trade diversion, which more than compensates for the decrease in transportation costs.

# Figure II

## Estimated Transportation Costs from Delhi




Panel A of Figure II shows the estimated transportation costs from Delhi at the district level for 2001; Panel B of Figure II shows the estimated transportation costs from Haryana at the district level for 2006; Panel C of Figure II shows the estimated transportation costs from Delhi at the state level for 2001; Panel D of Figure II 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.

# Figure III

## Percentage change in transportation costs from Delhi

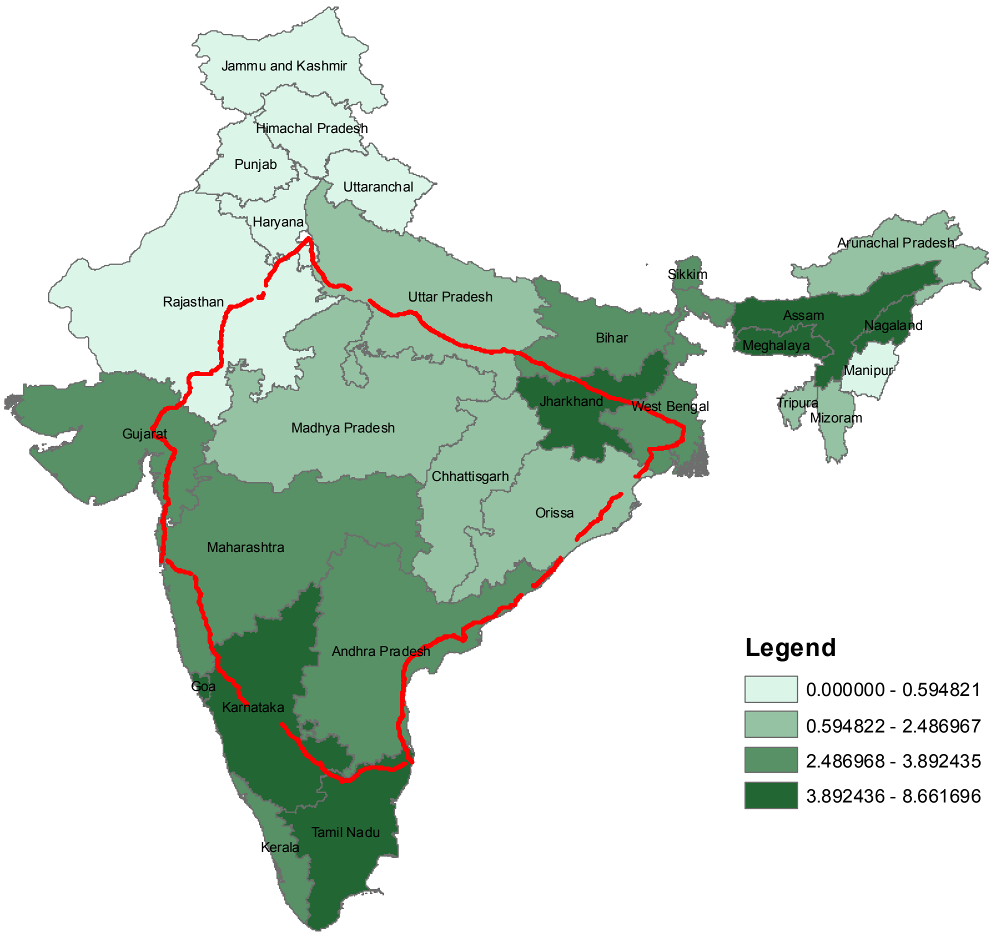



Figure III shows the % change in transportation costs due to the construction of the GQ at the state level.

# Figure IV

## Percentage change in real income after GQ

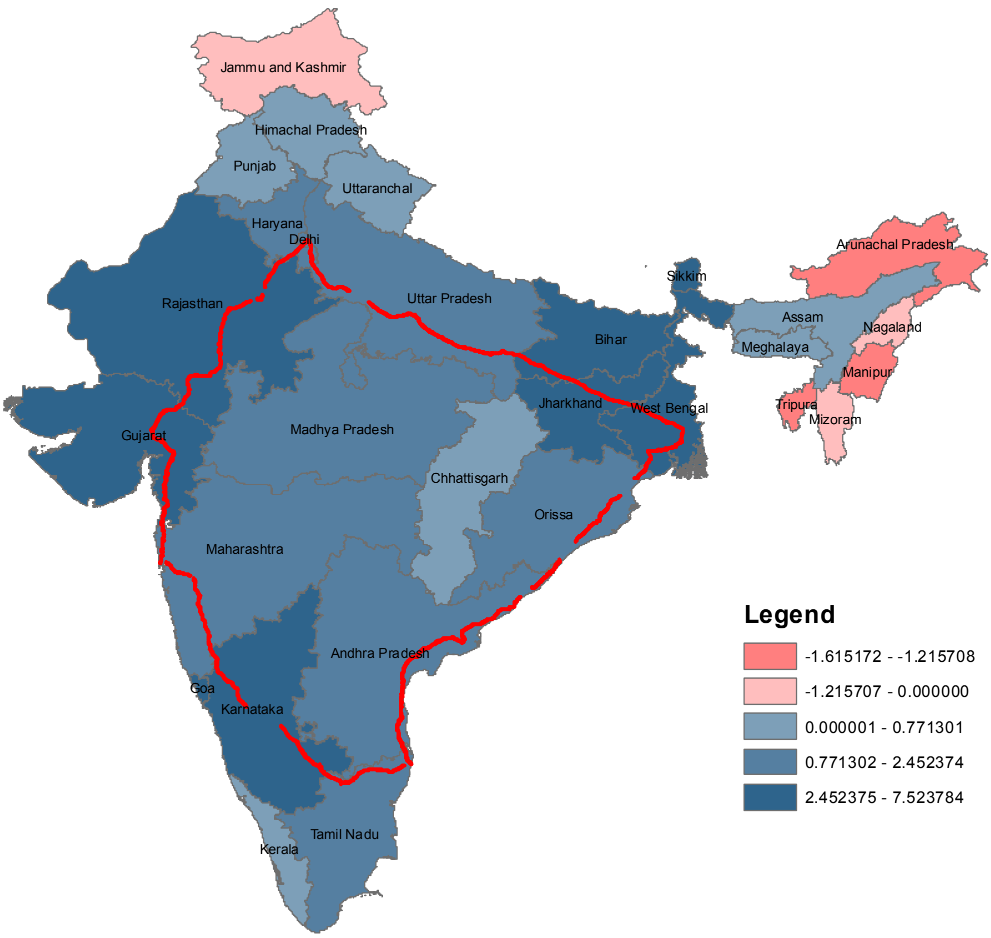



Figure IV shows the % change in real after the decrease in transportation costs due to the construction of the GQ. The numbers represented in this map correspond to the ones presented in column 2 of Table V.

## 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*.

Allen, T., and C. Arkolakis (Forthcoming): "Trade and the Topography of the Spatial Economy," *Quarterly Journal of Economics*.

Arkolakis, C., A. Costinot, D. Donaldson, and A. Rodríguez-Clare (2012): "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.

Atack, J., F. Bateman, M. Haines, and R. A. Margo (2010): "Did Railroads Induce or Follow Economic Growth?: Urbanization and Population Growth in the American Midwest, 1850-1860," *Social Science History*, 34(2), 171–197.

Atkeson, A., and A. Burstein (2008): "Pricing-to-market, trade costs, and international relative prices," *The American Economic Review*, 98(5), 1998–2031.

Atkin, D., and D. Donaldson (2014): "Who's Getting Globalized? The Size and Nature of Intranational Trade Costs," *Working Paper*.

Banerjee, A. V., E. Duflo, and N. Qian (2012): "On the Road: Access to Transportation Infrastructure and Economic Growth in China," pp. 1–50.

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.

Buera, F., J. Kaboski, and Y. Shin (2011): "Finance and Development: A Tail of Two Sectors," *American Economic Review*, 101, 1964–2002.

Caselli, F., and N. Gennaioli (2013): "Dynastic Management," *Economic Enquiry*, 51(1), 971–996.

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.

Costinot, A., and D. Donaldson (2012): "How Large are the Gains from Economic Integration? Theory and Evidence from U.S. Agriculture, 1880-2002," *Working Paper*.

Datta, S. (2012): "The impact of improved highways on Indian firms," *Journal of Development Economics*, 99(1), 46–57.

de Blas, B., and K. Russ (2010): "Understanding Markups in the Open Economy under Bertrand Competition," Discussion paper, National Bureau of Economic Research.

Devereux, M. B., and K. M. Lee (2001): "Dynamic gains from international trade with imperfect competition and market power," *Review of Development Economics*, 5(2), 239–255.

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

Donaldson, D., and R. Hornbeck (2014): "Railroads and American Economic Growth: A "Market Access" Approach," *Working Paper*.

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

Edmond, C., V. Midrigan, and D. Y. Xu (2012): "Competition, markups, and the gains from international trade," Discussion paper, National Bureau of Economic Research.

Epifani, P., and G. Gancia (2011): "Trade, markup heterogeneity and misallocations," *Journal of International Economics*, 83(1), 1–13.

Erosa, A., and L. Allub (2013): "Financial Frictions, Occupational Choice, and Economic Inequality," *Working Paper*.

Faber, B. (Forthcoming): "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System," *Review of Economic Studies*.

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.

Garicano, L., C. Lelarge, and J. V. Reenen (2012): "Firm Size Distortions and the Productivity Distribution: Evidence from France," *Working Paper*.

Ghani, E., A. G. Goswami, and W. Kerr (2013): "Highway to success in India: the impact of the golden quadrilateral project for the location and performance of manufacturing," *World Bank Policy Research Working Paper*, (6320).

Gourio, F., and N. Roys (2013): "Size-Dependent Regulations, Firm Size Distribution, and Reallocation," *Working Paper*.

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

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 (Forthcoming): "Allocative Efficiency, Markups, and the Welfare Gains from Trade," *Journal of International Economics*.

Hopenhayn, H. (2014): "Firms, Misallocation, and Aggregate Productivity: A Review," *The Annual Review of Economics*, 6(1).

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).

Licandro, O., and G. Impullitti (2013): "Trade, firm selection, and innovation: the competition channel," *Working Paper*, pp. 1–54.

Lopez-Martin, B. (2013): "From Firm Productivity Dynamics to Aggregate Efficiency," *Working Paper*.

Markusen, J. R. (1981): "Trade and the gains from trade with imperfect competition," *Journal of International Economics*, 11(4), 531–551.

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

Midrigan, V., and D. Xu (2014): "Finance and Misallocation: Evidence from Plant Level Data," *American Economic Review*, 2(104), 422–58.

Moll, B. (Forthcoming): "Productivity Losses from Capital Frictions: Can Self-Financing Undo Capital Misallocation?," *American Economic Review*.

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

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

——— (2013): "Misallocation and productivity," *Review of Economic Dynamics*, 16(1), 1–10.

Sotelo, S. (2014): "Trade Frictions and Agricultural Productivity: Theory and Evidence from Peru"," *Working Paper*.

Waugh, M. E. (2010): "International Trade and Income Differences," *American Economic Review*, 100(5), 2093–2124.

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