Highways and Spatial Location within Cities: Evidence from India

Ejaz Ghani, Arti Grover Goswami, and William R. Kerr

We investigate the impact of the Golden Quadrilateral (GQ) highway project on the spatial organization and efficiency of manufacturing activity. The GQ project upgraded the quality and width of 3,633 miles of roads in India. We use a difference-in-difference estimation strategy to compare non-nodal districts based upon their distance from the highway system. For the organized portion of the manufacturing sector, we find that GQ led to improvements in both urban and rural areas of nonnodal districts located 0–10 km from GQ. These higher entry rates and increases in plant productivity are not present in districts 10–50 km away. The entry effects are stronger in rural areas of districts, but the differences between urban and rural areas are modest relative to the overall effect. For the unorganized sector, we do not find material effects from the GQ upgrades in either setting. These findings suggest that in the time frames that we can consider—the first five to seven years during and after upgrades—the economic effects of major highway projects contribute modestly to the migration of the organized sector out of Indian cities but are unrelated to the increased urbanization of the unorganized sector.

JEL codes: L10, L25, L26, L60, L80, L90, L91, L92, M13, O10, R00, R10, R11, R14

Adequate transportation infrastructure is an essential ingredient for economic development and growth. Rapidly expanding countries like India and China face severe constraints on their transportation infrastructure. Business leaders, policy makers, and academics describe infrastructure as a critical hurdle for sustained growth that must be met with public funding, but to date there is a limited understanding of the economic impact of those projects. We study how proximity to a...
major new road network affects the organization of manufacturing activity. Additionally, we also compare how GQ influenced the manufacturing operations and entrepreneurship rates in the urban and rural portions of the districts through which it passed.

Our setting is the Golden Quadrilateral (GQ) Project, a large-scale highway construction and improvement project in India, using plant-level data from 1994 to 2009. The GQ project sought to improve the connection of four major cities of India—Delhi, Mumbai, Chennai, and Kolkata. Comprising 3,633 miles of road upgrades and new construction, the GQ network connected many of the major industrial, agricultural and cultural centers of India after its construction began in 2001. We find that districts located within 10 km to the GQ network experienced substantial increases in entry levels and higher organized sector productivity. The study suggests a significant increase in the overall output for the average district located on the GQ network, compared to no response in their nearby peers. We find that although entry effects for the organized sector are present in both urban and rural areas, the effects are much stronger and significant in rural settings. The differences are most substantial where new plants with major output levels are locating. By contrast, the productivity gains for the organized sector are very similar in both urban and rural locations.

Ghani et al. (2012) found that district level infrastructure is partly facilitating the relocation of organized manufacturing to rural locations while the unorganized manufacturing is migrating to urban locations. Our work on GQ suggests that such movement in the organized sector seems to be partially explained by national level highways as well, especially with regard to output of young plants. Since the unorganized sector is driving urbanization of Indian manufacturing, it is important for policy makers to understand the dynamics of this sector in the design of policies to promote urbanization. Our work suggests a very limited impact of the GQ upgrades on unorganized manufacturing outside of the nodal districts. We see traces of evidence of the organized sector findings repeating themselves in the unorganized sector but the results are substantially diminished in economic magnitudes. We confirm that these basic patterns are true in both urban and rural settings; they also hold true regardless of the gender of the business owner in the unorganized sector.

Our project contributes to the literature on the economic impacts of transportation networks and infrastructure investments in developing economies, which is unfortunately quite small relative to its policy importance (e.g., Ghani et al., 2012; Datta, 2011). Beyond India, several recent studies find positive economic effects in nonnodal locations due to transportation infrastructure in China (e.g., Banerjee et al. 2012; Baum-Snow et al. 2012; Roberts et al. 2012), Africa (e.g., Jedwab and Moradi 2015), and the United States (e.g., Fernald 1998; Chandra and Thompson 2000; Lahr et al. 2005; Baum-Snow 2007; Michaels 2008; Duranton and Turner 2011). This study also contributes to a literature seeking to understand the development of the manufacturing sector in India (e.g., Ahluwalia 2000; Besley and Burgess 2004; Kochhar et al. 2006) and especially
those emphasizing the importance of infrastructure constraints (e.g., Mitra et al. 1998; Gupta et al. 2008; Gupta and Kumar 2010). Finally, our work also links to studies seeking to understand the allocation of activity across regions and the productivity of firms (e.g., Desmet et al. 2015, Hsieh and Klenow 2009).

I. DATA AND ESTIMATION METHODOLOGY

We employ repeated cross-sectional surveys of manufacturing establishments carried out by the government of India. Data for organized sector surveys are sourced from Annual Survey of Industries conducted in 1994–95 onwards to 2009–10. In all cases, the survey was undertaken over two fiscal years (e.g., the 1994 survey was conducted during 1994–1995), but we will only refer to the initial year for simplicity. This time span allows us two surveys before the GQ upgrades (1994 and 2000) began in 2001 and several years post the upgrade. For unorganized manufacturing, we employ plant-level data from the years 1994, 2000, and 2005. Ghani et al. (2013, 2015) provide additional details on the data sources and preparation.

To introduce the spatial impact of GQ on Indian manufacturing across urban and rural regions, we estimate a difference-in-difference estimation, where a pre-post analysis is conducted with explanatory variables being interactions of indicator variables for how far a district is from the GQ highway network with an indicator variable for the post-GQ upgrades (equal to one in 2005 and 2007). Indexing districts with $i$ and years with $t$, the specification takes the form:

$$ Y_{i,t} = \sum_{d \in D} \beta_d \cdot GQ Dist_{i,d} \cdot PostGQ_t + \eta_i + \gamma_t + \epsilon_{i,t}. $$

(1)

The set $D$ contains three distance bands with respect to the GQ network: a nodal district, 0–10 km from the GQ network, and 10–50 km from the GQ network. The excluded category includes districts more than 50 km from the GQ network. The $\beta_d$ coefficients measure, by distance band, the average change in outcome $Y_i$ over the post period compared to the reference category. Most outcome variables $Y_i$ are expressed in logs, with the exception of TFP, which is expressed in unit standard deviations. District fixed effects control for the overall levels of entry rates in each district across the sample; these fixed effects also control for the main effects of distance from the GQ network. In a similar manner, the year fixed effects control for aggregate annual changes in the Indian economy, including the main effects of the post-GQ upgrades period. Thus, the interactions and their $\beta_d$ coefficients quantify differences in outcomes after the GQ upgrades by spatial band compared to the excluded group that comprises districts located more than 50 km from the GQ network.

Estimations report clustered standard errors, weight observations by log total district population in 2001, and have 312 observations representing the included districts. We winsorize outcome variables at the 1%/99% level to guard against
outliers. Our district sample is constructed such that employment, output, and establishment counts are continuously observed.

II. IMPACT OF GQ UPGRADERS ON ORGANIZED INDIAN MANUFACTURING

Table 1 presents the results from difference-in-difference estimations of equation 1 for organized manufacturing. Column headers provide the outcome variables studied. Column 1 considers the entry of young plants by their log count in the district, Column 2 considers the log employment in these young plants, and Column 3 considers the log output in these young plants. Columns 4–6 repeat the same for urban areas while Columns 7–9 reproduce the results for rural sections of a district.

The top row of Column 1 shows a very substantial increase in young firm counts in nodal cities after the upgrades in comparison to districts more than 50 km from GQ. We are very cautious about interpreting these results much given that the upgrades were built around the connectivity of the nodal cities. The imprecision in these estimates is mostly due to the fact that there are only nine nodal districts. As effects for our other distance categories are being measured for each band relative to districts more than 50 km from the GQ network, the inclusion or exclusion of the nodal districts does not impact our core results regarding nonnodal districts.

The key pattern emphasized in Ghani et al. (2015) is shown in the second and third rows. There is a substantial increase in organized sector entry in districts within 10 km of GQ that is not reflected in districts farther away at 10–50 km. Our sample includes 76 districts within 10 km, and 42 districts in the 10–50 km band. To some degree (substantiated further in Ghani et al. 2015), the upgrades of the GQ network can be taken as exogenous for these districts since neither group is in the nodal district set around which the upgrades were based. The districts within 0–10 km of GQ have a 0.4–0.9 log point increase in entry activity after the GQ upgrade compared to districts more than 50 km away.

The remainder of this table applies the methodology from Columns 1–3 to urban and rural areas of districts independently. Some districts do not contain an urban or rural area. Specifically, of our 312 district, 278 have urban areas and 291 have rural areas. In the estimations that consider levels of activity in urban areas, we will restrict our samples to districts that have manufacturing establishments in urban areas in all the years surveyed. In a similar manner, the estimates for rural areas require that the district have rural areas across the full sample period. This provides a consistent panel for each group.

The results for urban areas in Columns 4–6 are more muted than the total effects document in Columns 1–3. We find economically and statistically significant increases in plant entry rates for Column 4. By contrast, Columns 5 and 6 find positive point estimates for employment and output, respectively, but these coefficients are not precisely estimated like the base results in Columns 2 and 3.
TABLE 1. Pre-post estimations of the impact of GQ improvements in urban vs. rural areas for organized sector entrants

<table>
<thead>
<tr>
<th></th>
<th>Total GQ effect for entrants</th>
<th>Columns 1–3 in urban areas</th>
<th>Columns 1–3 in rural areas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Plants (1)</td>
<td>Employment (2)</td>
<td>Output (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post GQ upgrades*</td>
<td>0.702</td>
<td>1.167</td>
<td>1.647+</td>
</tr>
<tr>
<td>Nodal district</td>
<td>(0.662)</td>
<td>(0.814)</td>
<td>(0.951)</td>
</tr>
<tr>
<td>Post GQ upgrades*</td>
<td>0.436++</td>
<td>0.471+++</td>
<td>0.928+++</td>
</tr>
<tr>
<td>District 0-10 km from GQ</td>
<td>(0.172)</td>
<td>(0.239)</td>
<td>(0.346)</td>
</tr>
<tr>
<td>Post GQ upgrades*</td>
<td>-0.012</td>
<td>-0.056</td>
<td>-0.263</td>
</tr>
<tr>
<td>District 10-50 km from GQ</td>
<td>(0.240)</td>
<td>(0.357)</td>
<td>(0.537)</td>
</tr>
<tr>
<td>District and year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1248</td>
<td>1248</td>
<td>1248</td>
</tr>
</tbody>
</table>

Notes: Districts are local administrative units that generally form the tier of local government immediately below that of India’s subnational states and territories. These are the smallest entities for which data is available with ASI. Nodal districts include Delhi, Mumbai, Kolkata, and Chennai and their contiguous suburbs (Gurgaon, Faridabad, Ghaziabad, and NOIDA for Delhi; Thane for Mumbai). The indicator variable for District 0–10 km from GQ takes a unit value for non-nodal districts that have minimum straight-line distance from the GQ network of less than 10 km; other distance-related indicator variables are defined analogously. Estimations consider the logged values of economic activity for young plants of organized sector manufacturing activity in 312 Indian districts for 1994, 2000, 2005, and 2007 from the Annual Survey of Industries. Young plants are those that are less than four years old. The table estimates effects of GQ upgrades for nearby districts relative to districts more than 50 km from the GQ network. The Post GQ upgrades variable takes unit value for the years 2005 and after, once the GQ upgrades commenced in 2001. Outcome variables are winsorized at their 1% and 99% levels. Estimations report standard errors clustered by district, include district and year fixed effects, and weight observations by log total district population in 2001. * is used to denote the interaction of two variables, while +, ++, and +++ denote statistical significance at the 10%, 5%, and 1% levels, respectively.
The output coefficients are also substantially smaller, between one-third and a half the size of what is evident in Column 3 for the full sample.

When isolating rural areas of districts in Columns 7–9, the outcomes have several key differences. First, the rural entry responses are generally larger than those estimated for urban areas, regardless of the metric. Second, unlike the emphasis on plant counts in urban areas, the rural estimates place more emphasis on employments and outputs of the organized sector. That is, the GQ effort appears to have increased the entry rates of organized sector plants in both urban and rural settings, but the largest increases in terms of employments and outputs are seen in rural settings. In fact, the development of large employment and output bases for organized manufacturing after the GQ upgrades plays the strongest role in the aggregate gains experienced on these dimensions. Ghani et al. (2013) provide a full set of results and those across extended distance bands.

In a similar manner, table 2 considers labor productivity and TFP estimations, using the Sivadasan (2009) methodology to calculate the levinsohn-Petrin (L-P) style TFP estimates. Columns 1 and 2 show a core increase of productivity associated with the GQ implementations in districts nearby the highway system. These increases are primarily driven by adjustments in the incumbent establishments of the districts. We do not quantify the labor productivity and TFP changes of new entrants, as much of the impact of new entrants comes from the extensive margin and these plant-level traits are not defined in these cases. In contrast to the differences observed for entry rates across urban and rural areas, the productivity effects appear quite uniformly realized.

| Table 2. Pre-post estimations of manufacturing productivity for organized sector |
|---------------------------------|-----------------|----------------|-----------------|
| | Total GQ organized sector effect, entrants | Columns 1–2 in urban areas | Columns 1–2 in rural areas |
| | Labor | TFP | Labor | TFP | Labor | TFP |
| Base spatial horizon measuring effects relative to districts 50+ km from the GQ network |
| Post GQ upgrades* | 0.084 | −0.022 | 0.037 | 0.010 | 0.126 | 0.011 |
| Nodal district | (0.135) | (0.033) | (0.145) | (0.043) | (0.264) | (0.167) |
| Post GQ upgrades* | 0.177+ | 0.086++ | 0.165 | 0.098 | 0.198++ | 0.099+ |
| District 0–10 km from GQ | (0.093) | (0.042) | (0.139) | (0.066) | (0.099) | (0.056) |
| Post GQ upgrades* | 0.043 | −0.005 | −0.000 | −0.021 | 0.115 | 0.068 |
| District 10–50 km from GQ | (0.132) | (0.074) | (0.186) | (0.085) | (0.175) | (0.099) |
| District and year fixed effects | Yes | Yes | Yes | Yes | Yes |
| Observations | 1248 | 1244 | 1108 | 1100 | 1160 | 1160 |

Notes: See Table 1. Labor productivity is calculated as total output per employee, and TFP is calculated as the residual of value added over capital and labor inputs in a standard Cobb-Douglas production function using the L-P Sivadasan technique. See Ghani et al. (2015) for details on the methodology. * is used to denote the interaction of two variables, while ++, +++ denote statistical significance at the 10%, 5%, and 1% levels, respectively.
although the coefficients remain significant for rural regions vis-à-vis urban areas. The labor productivity elasticity is slightly stronger in rural areas, likely reflective of the larger-scale production functions that can be used in these locations.

Ghani et al. (2015) confirm these results for overall organized manufacturing using a long-difference estimation technique. Additionally, there are natural concerns about the endogenous placement of the GQ highway system. For example, policy makers might have known about the latent growth potential of regions and attempted to aid that potential through highway development. Ghani et al. (2015) address these concerns by providing three alternative robustness checks described next.

Placebo Estimations

Ghani et al. (2015) compare districts proximate to the GQ network to districts proximate to the North South-East West (NS-EW) highway network (see figure 1 for a map of both GQ and NS-EW). The idea behind this comparison is that

Figure 1. GQ and NS-EW Highway route structure
districts that are at some distance from the GQ network may not be a good control
group if they have patterns of evolution that do not mirror what districts immedi-
ately on the GQ system would have experienced had the GQ upgrades not oc-
curred. The null results observed for districts close to the NS-EW corridor in
Ghani et al. (2015) provide a stronger foundation in this regard, especially as its
upgrades were planned to start at the same time as those of the GQ network
before being delayed. The identification assumption is that unobserved condi-
tions such as regional growth potential along the GQ network were similar to those for
the NS-EW system (conditional on covariates).

**Instrument Variable (IV) Estimations**

Continuing with potential identification challenges, Ghani et al. (2015) also
consider if the GQ planners were better able to shape the layout of the network
to touch upon India’s growing regions vis-à-vis the NS-EW planners. The dis-
tricts 0–10 km from the GQ network are instrumented with being 0–10 km
from a (mostly) straight line between the nodal districts of the GQ network.
The identifying assumption in this IV approach is that endogenous placement
choices in terms of weaving the highway towards promising districts (or strug-
gling districts) can be overcome by focusing on what the layout would have
been if the network was established based upon minimal distances only. Ghani
et al. (2015) find that the first-stage relationship of this IV estimation is quite
strong. The IV specifications generally confirm the OLS findings, and, in most
cases, the null hypothesis that the OLS and IV results are the same cannot be
rejected.

**Dynamic Estimations**

Dynamic patterns around these reforms provide additional assurance about the
role of the GQ upgrades in these economic outcomes and insight into their
timing. By separately estimating effects for each year, it is feasible to observe
whether the growth patterns appear to follow the GQ upgrades hypothesized to
cause them. As an example, we present results for dynamics estimations of
output for young plants. Figure 2 plots the coefficient values for log new output
and their 90% confidence bands. Vertical lines in the figure marks when the GQ
upgrades began and when they reached the 80% completion mark. Effects are
measured relative to 1994, and we see no differences in 1999 or 2000 for non-
nodal districts within 10 km of the GQ compared to those 10–50 km apart.
Once the GQ upgrades commence, the patterns for output of young plants is
pretty dramatic. These patterns confirm that the timing of the GQ upgrades coin-
cides with the timing of growth in output of young plants.

**III. Impact of GQ on Unorganized Indian Manufacturing**

This section summarizes the connection (or lack thereof) between GQ upgrades
and the broad development of unorganized manufacturing activity. One effort in
this project was to implement the GQ empirical strategy used in Ghani et al. (2015) in the unorganized sector to compare against the large, rapid effects observed for the organized sector. Our unorganized sector data stops in 2005, compared to 2007 for difference-in-difference estimations of the organized sector. Although we hope in future work to extend this time frame to 2010 for both sectors, even by 2005 the organized sector shows a strong response in terms of young firm activity and productivity (Ghani et al. 2015) and sourcing/inventory management (Datta 2011). It is thus useful to quantify whether the unorganized sector has a similarly rapid response.

Table 3 reports one of the portions of our basic framework for the unorganized sector from Ghani et al. (2013) to show differences and similarities with the organized sector. We highlight the following key lessons here. First, unlike the organized sector, difference-in-difference estimations find that non-nodal districts close to the GQ network behave similarly to those located farther away with respect to aggregate levels of unorganized manufacturing. In Ghani et al. (2013) we show that this comparability of districts 0–10 km from the GQ network with those that are 10–50 km apart holds irrespective of the whether

![Figure 2. Dynamics of log new Output Growth](https://academic.oup.com/wber/article-abstract/30/Supplement_1/S97/2897427)

**Notes:** The figure illustrates the dynamics of output growth among young plants for non-nodal districts located 0–10 km from the GQ network relative to districts 10–50 km from the GQ network. The solid line quantifies the differential effect for the GQ upgrades by year, with 1994 as the reference year. Dashed lines present 90% confidence intervals, with standard errors clustered by district.
the highway development close to the district was an upgrade or a new construction (although there are traces of a stronger response for new construction). Likewise, we do not find differences in unorganized activity depending upon the completion date of the GQ upgrade (e.g., considering areas completed prior to 2002) or when examining the gender balance of plant owners. In sum, it appears that the highway improvements had limited aggregate effects for the unorganized sector, especially in comparison to the organized sector. **Khanna (2014)** provides complementary evidence outside of manufacturing using nighttime lights data.

These results are reasonable given the structure of unorganized manufacturing in India. By definition, unorganized enterprises are small establishments. Most employees in these enterprises are not full time, and many of these establishments are household based. Given the informality of this sector, it is less likely that such establishments and their new location decisions would depend as strongly on materials and products and connectivity to the market through national highways. The small-scale nature of these establishments makes it more likely that they target local product and labor markets. Thus, it is likely that they are more dependent on state or district roads vis-à-vis national highways.

### Table 3. Pre-post estimations of the impact of GQ improvements on unorganized activity

<table>
<thead>
<tr>
<th></th>
<th>Log levels of total activity</th>
<th>Log levels of young firm activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Plants (1)</td>
<td>Employment (2)</td>
</tr>
<tr>
<td>Base spatial horizon measuring effects relative to districts 50+ km from the GQ network</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post GQ upgrades*</td>
<td>0.184</td>
<td>0.259</td>
</tr>
<tr>
<td>Nodal district</td>
<td>(0.177)</td>
<td>(0.146)</td>
</tr>
<tr>
<td>Post GQ upgrades*</td>
<td>-0.086</td>
<td>-0.096</td>
</tr>
<tr>
<td>District 0-10 km from GQ</td>
<td>(0.105)</td>
<td>(0.109)</td>
</tr>
<tr>
<td>Post GQ upgrades*</td>
<td>-0.031</td>
<td>0.006</td>
</tr>
<tr>
<td>District 10-50 km from GQ</td>
<td>(0.116)</td>
<td>(0.122)</td>
</tr>
<tr>
<td>District and year fixed effects</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1089</td>
<td>1089</td>
</tr>
</tbody>
</table>

Notes: Estimations consider the location of unorganized sector manufacturing activity in 312 Indian districts for 1994, 2000, and 2005 from the National Sample Statistics. Young plants are those that are less than four years old. The table estimates effects of GQ upgrades for unorganized sector in nearby districts relative to districts more than 50 km from the GQ network; Ghani et al. (2013) also includes extended spatial rings to measure effects relative to districts 200 km away from the GQ network. The Post GQ upgrades variable takes unit value for the year 2005 after the GQ upgrades commenced in 2001. Outcome variables are winsorized at their 1% and 99% levels, and entry variables are coded at the 1% level where no entry is observed to maintain a consistent sample. Estimations report standard errors clustered by district, include district and year fixed effects, and weight observations by log total district population in 2001. * is used to denote the interaction of two variables. See Ghani et al. (2013) for more details.
IV. Conclusions

Many discussions of India’s future highlight the need and expected growth of urbanization. Another set of discussions, which often overlap extensively, highlight the importance of enhancing India’s infrastructure (e.g., McKinsey 2010, 2012; World Bank 2012). In this paper, we have quantified the degree to which one such infrastructure project—the massive upgrades to the GQ highway network that connects four of India’s major cities— influenced the urbanization of the manufacturing sector in nonnodal locations.

The patterns are quite intriguing. The upgrades are connected to enhancements in organized sector activity in both urban and rural environments. Across the full span of results from Ghani et al. (2013), we conclude that increases in entry rates and plant productivity are fairly balanced between the two settings, with the main exception being that rural areas receive relatively more stimulus in terms of net output growth. By contrast, the unorganized sector of manufacturing is not closely linked to the GQ developments in either location. Thus, the GQ experience suggests that major interdistrict projects are likely to continue and accelerate the spatial adjustments of the organized sector and its general move towards rural locations. On the other hand, the GQ experience does not appear connected to the general movement of the unorganized manufacturing sector into Indian cities.

Our work on India’s GQ highway system provides an important input into policy choices. It builds a framework for estimating the likely quantitative impact of infrastructure development projects and also provides estimates of the relative impacts across districts by distance to the network, thereby offering additional insights into the distributional consequences of large scale infrastructure projects. This methodology would be applicable to similar settings, where, for instance, poor transportation infrastructure severely hinders economic activity. On the whole, the paper speaks to the severe constraints that inadequate infrastructure can have for the development of manufacturing in emerging economies and the potential growth that may follow from alleviating that constraint.

References


